An Adaptive Technique to Find Pricing and Repurchasing of Cloud Space in Big Data

Simran Dhussa[1], Kalyani Mule[2], Ritika Borawake[3], Pooja Dixit[4], Prof.R.R.Nikam[5]

BE Students, Information Technology Engineering, SRES COE, Kopargaon, Maharashtra, India[1]

Assistant professor, Information Technology Engineering, SRES COE, Kopargaon, Maharashtra, India[2]

ABSTRACT

Processing streaming big data becomes critical as new diver Internet of Thing applications begin to emerge. Here a cloud service intermediary rents the cloud service from multiple cloud providers and provides streaming processing service to the users with multiple service interfaces. Depending upon the usage access of the user we are going to allocate the package for the particular user depending upon the multi-cloud data downloaded from the Online Server. Clustering algorithm is used for processing huge data using the MapReduce Technology and MapReduce Algorithm. We provide an excellent way for the user to interact with their cloud system using pricing and repurchasing of their cloud usage package. We also evaluate our pricing strategy with extensive simulations and compare its revenue with our pricing strategies. The intermediary can repurchase the sparse capacity with dynamic rate per load which depends on the duration time and the amount of sparse resources that the users hold. The intermediary can choose different prices to users for different users according to the amount and time of computing resources rented.

Keywords: *Streaming Big Data, Cloud Computing, Multiple Cloud*.

1 INTRODUCTION

Now-a-days the very important of big data computing is streaming data processing. Processing streaming big data gets to be basic as new jumper Internet of Thing applications begin to emerge. Usually, for lower maintenance cost, users often use cloud services for processing big data. With cloud services, it is no need to maintain a large scale cluster and users only consider the details of big data computing. Usually, in some cases, cloud providers give users a reasonable price for their services, especially for some long-term users. Here a cloud service intermediary rents the cloud benefit from different cloud suppliers and gives spilling processing service to the users with multiple service interfaces. Depending upon the usage access of the user we are going to allocate the package for the particular user depending upon the multi-cloud data downloaded from the Online Server.

2 LITERATURE SURVEY

1. Estimating the statistical characteristics of remote sensing big data in the wavelet transform domain

In this paper,L. Wang, H. Zhong, R. Ranjan, A. Zomaya, and P. Liu, proposed that,Streaming data processing is becoming a very important part of big data computing in recent years. Usually, for lower maintenance cost, users often use cloud services for processing big data. With cloud services, it is no need to maintain a large scale cluster and users only consider the details of big data computing. Furthermore, some cloud providers also provide computing services based on some popular systems (e.g., Hadoop, etc.). With this type of service, users only need to put their processing programs on the cloud platform and wait for the result.

2. Resource provisioning policies to increase Iaas providers profit in a federated cloud environment, in High Performance Computing and Communications (HPCC)

In the paper,A. Toosi, R. Calheiros, R. Thulasiram, and R. Buyya proposed that,Usually, in some cases, cloud provides give users a reasonable price for their services, especially for some long-term users. However, for most streaming big data computing scenario, their price seems a little expensive especially the rate per load. To reduce the cost for streaming big data computing, an optional method is choose some small cloud providers with lower rate per load. However, small cloud providers have not enough capacity to support large scale workloads.Meanwhile, they also provide simple service without any support of big data computing.

3. Elastic mapreduce over multiple clouds, in The 13th IEEE/ACM International Symposium On Cluster, Cloud and Grid Computing

In this paper, A. Iordache, C. Morin, N. Parlavantzas, E. Feller, and P. Riteau, proposed that Multiple cloud service mode is a better solution that users can deploy their computing in multiple cloud providers with lower rate. However, with multiple cloud providers,users have to considerate about the difficulty of management and the deployment of big data computing systems. Multiple cloud intermediaries can provide flexible services for users deploying their big data computing programs with a lower rate than the main cloud providers.

4. The design of the borealis stream processing engine

In this paper, D. J. Abadi, Y. Ahmad, M. Balazinska, M. Cherniack, J. Hyon Hwang, W. Lindner, A. S. Maskey, E. Rasin, E. Ryvkina, N. Tatbul, Y. Xing, and S. Zdonik proposed that Borealis is a distributed extension of Aurora which can process streaming data through multiple processors and computers. For support distributed architecture, Borealis presents an efficient algorithm for the distribution of jobs between nodes. The STREAM project provides a data base management functionality to support for continuous queries over streaming data. It presents a relatively minor extension to SQL named CQL to offer Stream-to-Relation and Relation-to-Stream Operators.

3 SYSTEM ARCHITECTURE

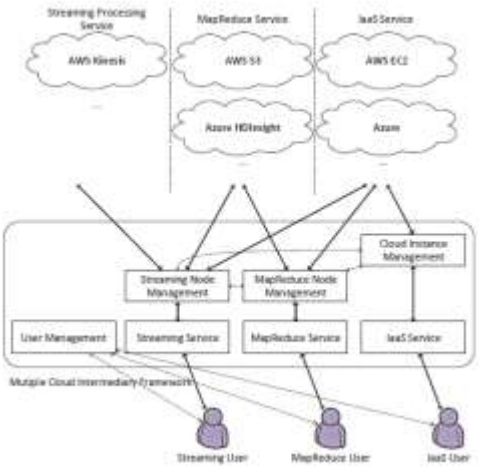

Figure 1: Multiple Cloud Intermediary Framework Structure

As the structure shown in Figure 1, the multiple cloud intermediary framework for streaming computing consists of several modules to meet the design concepts. There are seven main modules in the framework including the cloud instance management, streaming node management, MapReduce node management, streaming service, MapReduce service, IaaS Service and user management modules.Cloud instance management module manages all compute instances at the IaaS service level. This module records all status of the instances and assigns appropriate instances to other modules.Streaming node management module manages the computing resources which are provided to users as streaming computing service level. The streaming computing resources are generated in three types of methods. First type is that the intermediary rents resources from the streaming processing cloud services. Second type is that the module deploys the streaming processing systems on the MapReduce nodes. Third type is that the module deploys the streaming processing systems in the compute instances directly. MapReduce node management module manages the computing resources which are provided to users as MapReduce service level. Similarly with the streaming computing resources, the MapReduce computing resources are generated from two types: the resources rented from the cloud MapReduce services, and the module deploys the MapReduce systems in the compute instances.Streaming service module provides streaming computing services to the users. To provide the required ser- vice interface from users, the streaming service module integrates general streaming processing systems. MapReduce service module provides MapReduce services to the users. Similarly, the MapReduce service module integrates general MapReduce implementations to provide the compatible interfaces to support the streaming systems from users. IaaS service module provides IaaS services to the users. Usually, users can get compute instances from this module with the required version of the operating systems and some necessary software. User management module manages all users in the intermediary framework including access control, usage history, billing, etc.

4 WORKING

4.1 Multiple Cloud Systems

There are some existing work focuses on integrate computing resources from multiple cloud providers.

Apache CloudStack is software to integrate cloud computing resources with resource management, user management, API and graphical user interface. Eucalyptus is also similar software which focuses on building Amazon AWS-compatible private and hybrid clouds.

OpenNebula is a multiple cloud software aims at providing an industry standard solution for creating and managing virtual data centers across multiple cloud provides.

OpenStack is the most famous cloud management system which provides an API and a dashboard to manage pools of computing, storage, and network resources from single or multiple cloud environments.

VMware vCloud is a multiple cloud infrastructure that allowing to organize cloud computing at three levels including infrastructure level, platform level and service level. fog is a Ruby API for providing access to computing and storage resources across multiple cloud provides. It also provides an in-memory cloud resource representation to help developers to test and simulate their deployment.

jcloud is also an API for delivering an abstraction layer over the APIs from cloud providers, which facilitates users using means of templates to describe generic virtual machines and allows deploying and grouping of multiple virtual machines.

Cloud4SOA is a multi-cloud PaaS management which enables software developers to create, deploy, execute, and mange business applications through multiple cloud providers.The multicloud based evacuation services architecture maintains basic monitoring and maintenance sercices during of normal activity but quickly scales up service capacity during an emergency.

5 FUTURE SCOPE

In future, we will plan to implement a complete multiple cloud intermediary solution with modified Open Stack to support streaming big data processing management. Meanwhile, it is signification to find scheduling method to optimize the streaming computing performance in the multiple cloud environments. A deeper experiment with the real word testbed is also needed to evaluate the efficiency of the new multiple cloud intermediary solution.

6 CONCLUSION

Thus, we propose a multiple cloud intermediary framework for streaming big data computing to provide streaming big data processing cloud services to the users. The intermediary rents computer resources from different cloud services and provides different service interfaces to users. We also design a Pricing-Repurchasing strategy to maximum the revenue of the intermediary and decrease the risks by long-term renting contracts with users.We formulate the Pricing-Repurchasing problem as a two-stage leader-follower (Stackelberg) game, and analyze the game equilibrium. We also evaluate our pricing strategy with extensive simulations and compare its revenue with our pricing strategies. From the result of performance evaluation, the Pricing-Repurchasing strategy brings more revenue to the intermediary than other methods.

7 REFERENCES

[1] L. Wang, H. Zhong, R. Ranjan, A. Zomaya, and P. Liu, Estimating the statistical characteristics of remote sensing big data in the wavelet transform domain, IEEE Transactions on Emerging Topics in Computing, vol. 2, no. 3, pp. 324337, Sept 2014.

[2] A. Toosi, R. Calheiros, R. Thulasiram, and R. Buyya, Resource provisioning policies to increase iaas providers profit in a federated cloud environment, in High Performance Computing and Communications (HPCC), 2011 IEEE 13th International Conference on, Sept 2011, pp. 279287.

[3] A. Iordache, C. Morin, N. Parlavantzas, E. Feller, and P. Riteau, Resilin: Elastic mapreduce over multiple clouds, in The 13th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid 2013), May 2013, pp. 261268.

[4] D. J. Abadi, Y. Ahmad, M. Balazinska, M. Cherniack, J. Hyon Hwang, W. Lindner, A. S. Maskey, E. Rasin, E. Ryvkina, N. Tatbul, Y. Xing, and S. Zdonik, The design of the borealis stream processing engine, in In CIDR, 2005, pp. 277289.