Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 5.258

IJCSMC, Vol. 5, Issue. 4, April 2016, pg.105 – 109

Multi-Cloud Data Hosting for Protection Optimization and Security

C. Divya Shaly[1], R.Anbuselvi[2]

[1]Research Scholar, Department of Computer Science, Bishop Heber College Tiruchirappalli, Tamilnadu, India
[2]Asst.Professor, Department of Computer Science, Bishop Heber College Tiruchirappalli - 620017, Tamilnadu, India
[1] shalysweetie@gmail.com [2] r.anbuselvi@yahoo.in@gmail.com

Abstract—The enterprises and organizations are hosting their data into the cloud, in order to reduce the IT maintenance enhance and cost the data reliability., Facing the numerous cloud vendors as well as their heterogenous pricing policies, customers may well be puzzle with which cloud(s) are suitable for storing their data and what hosting strategy is the cheaper. The general status is that customers usually put their data into a single cloud (and then simply trust to the luck. Based on the comprehensive analysis of many state-of-the-art cloud vendors, this paper proposes a novel data hosting scheme named CHARM which integrates two key functions crave. The first is selecting several suitable clouds and an appropriate redundancy strategy to store data with minimized financial cost and guaranteed availability. The second is triggering a transition process to alert the distribution of data according to the variations of data access pattern and pricing of clouds. CHARM not only saves around 20 percent of financial cost but also exhibits sound adaptability to data and price adjustments.

Keywords- Multi Cloud, CHARM, Heterogeneous Cloud, Storage Services, DEPSKY.

I. INTRODUCTION

1.1 HETEROGENEOUS CLOUDS

Existing clouds exhibit great heterogeneities in terms of both working an pricing policies and performances. Different cloud vendors build their respective keep upgrading them with newly emerging gears and infrastructures. They also design diverse system architectures and apply various techniques to make their services competitive. Such system diversity leads to visible performance variations across cloud vendors moreover pricing policies of accessible storage services provided by dissimilar cloud vendors are separate in both pricing levels and charging items.

1.2 MULTI-CLOUD DATA HOSTING

Newly, multi-cloud data hosting has received wide notice from researchers, customers, and startups. The fundamental principle of multi-cloud (data hosting) is to allocate data across multiple clouds to gain improved redundancy and avoid the vendor lock-in risks.

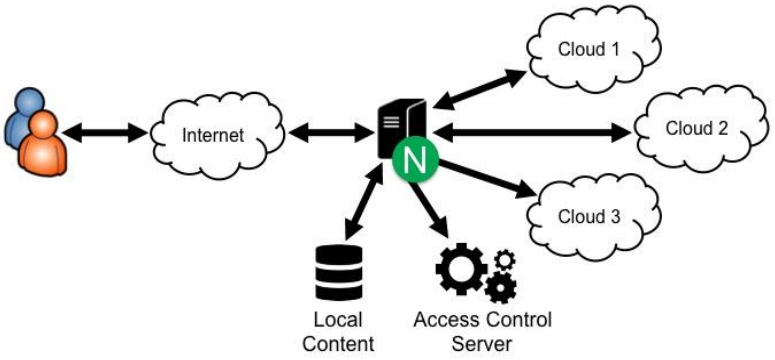

Fig.1.1.Multi-Cloud Architecture

1.3 DATA HOSTING SCHEME

To efficient heuristic-based algorithm to choose proper data storage modes. Moreover, to implement the necessary procedure for storage mode transition by monitoring the variations of data access patterns and pricing policies. In the proposed system and implement CHARM, a novel, efficient, and heuristic-based data hosting scheme for heterogeneous multi-cloud environments. CHARM accommodate different pricing strategies, availability requirements, and data access patterns. It select suitable clouds and an appropriate idleness strategy to store data with minimized monetary cost and guaranteed availability.

To implement a flexible transition scheme for CHARM. It keeps monitoring the variations of pricing policies and, adaptively triggers and data access patterns the transition process between different data storage modes. It starts a data migration process among different clouds if necessary. Data Hosting, Workload Statistic, Storage Mode Switching (SMS), and Predictor. Workload Statistic keeps collecting and tackling access logs to guide the placement of data. It also sends statistic information to interpreter which guides the action of SMS. Data Hosting stores data using replication or removal coding, according to the size and access frequency of the data. SMS decides whether the storage mode of certain data should be distorted from replication to erasure coding or in reverse, according to the output of Predictor.

II. REVIEW OF LITERATURE

2.1 CLOUD STORAGE SERVICES

Cloud storage services such as Microsoft One Drive ,Google Drive and Drop boxs provide users with a convenient and reliable way to share and store data from anywhere, on any device, and at any time. The users' data stored in cloud storage are mechanically synchronized across all the designated devices connected to the cloud in a well timed manner. With multiplicity of devices – particularly mobile devices – that users possess today, such "anywhere, anytime" features significantly simplify consistency maintenance and data management, thus provide an ideal tool for data sharing and association.

File operation includes file creation, file deletion, file modification, and frequent file modifications.

Data update rate denotes how often a file operation happens. Sync deferment. When frequent file modifications happen, some cloud storage services intentionally reschedule the sync process for a certain period of time for batching file updates.

Data sync granularity. A file operation is synchronized to the cloud either in a full-file granularity or in an incremental, chunk-stage granularity. When the former is adopted, the whole updated file is delivered to the cloud; when the latter is adopted, only those file chunks that contain altered bits (relative to the file stored in the cloud) are delivered.

2.2 TRAFFIC OVERUSE PROBLEM

However, although these performance optimizations, we observe that the network traffic construct by cloud storage applications exhibits pathological badly in the presence of frequent, short updates to user data. Each time a synced file is customized, the cloud storage application's inform-trigger real-time synchronization (URS) mechanism is activated . compress the binary diff of the new data ,send and URS compute and update to the cloud with several session maintenance data. unluckily, when there are frequent, short updates to synced files, the amount of session maintenance traffic far exceeds the amount of useful update traffic sent by the client over time. In the behavior the traffic overuse problem. In essence, the traffic overuse difficulty originates from the update sensitivity of URS. It's not find variation problems in the three providers. The charge model are comparable for all providers , based on the number of operations and the size of the blob.

2.3 LOSS AND CORRUPTION OF DATA

There are numerous cases of cloud services losing or corrupting customer data, Lost the contacts, notes, photos, etc. of a large number of users. The data was recovered several days later, but the users is not so lucky, when the company lost half a terabyte of data that it not at all managed to recuperate DEPSKY deals with this problem using Byzantine fault-tolerant duplication to store data on several cloud services, allowing data to be retrieved properly even if some of the clouds corrupt or lose data.

2.4 LOSS OF PRIVACY

The cloud source has access to both the data stored in the cloud and metadata like access patterns. The provider may be trustworthy, but mean insiders are a wide-spread security problem. This is an especial concern in applications that involve keeping personal data like health records. An obvious solution is the customer encrypting the data by storing it, but if the data is accessed by distributed applications this involves running protocols for key distribution process in different machines need access to the cryptographic keys). DEPSKY employs a secret sharing scheme and removal codes to avoid storing plain data in the clouds and to improve the storage efficiency, amortizing the duplication factor on the cost of the solution.

2.5 VENDOR LOCK-IN

There is presently some unease that a few cloud computing providers become leading, the so called vendor lock-in issue level moving from one provider to another one may be exclusive because the cost of cloud usage has a component relative to the amount of data that is written and read. DEPSKY addresses this issue in two customs. First, it does not depend on a single cloud provider, but on a only some, so data access can be balanced among the providers allowing for their practices. Second, DEPSKY uses removal codes to store only a fraction of the whole amount of data in each cloud. In case the need of replace one source by another arises, the cost of migrate the data will be at most a portion of what it would be otherwise.

2.6 CRITICAL DATA STORAGE

Given the overall compensation of using clouds for running large scale systems, around the world many governments are considering the use of this model.

III. PROPOSED WORK

In the proposed work CHARM scheme. In this paper propose a novel cost-efficient data hosting scheme with high availability in varied multi-cloud, named “CHARM”. It cleverly puts data into multiple clouds with minimized financial cost and guaranteed availability. Specifically, its combine the two widely used redundancy mechanisms, i.e., replication and removal coding, into a consistent model to meet the required availability in the presence of different data access patterns. Next, its devise an efficient heuristic-based algorithm to choose proper data storage modes (redundancy mechanisms and involving both clouds and). Moreover, we implement the required procedure for storage mode transition by monitoring the variations of data pricing policies and access patterns. In the evaluate the performance of CHARM using both trace driven simulation and prototype experiments. The traces are composed from two online storage systems, both of which possess of thousands of users. As a holistic storage system, there are several other factors to be considered, geographical data consistency, such as cache strategies. In this proposed work consists of following modules are,

3.1 CLOUD STORAGE

Cloud storage services have become gradually more popular. Because of the importance of privacy, many cloud storage encryption schemes have been proposed to defend data from those who do not have access. All such schemes assumed that cloud storage provider are secure and cannot be hacked; however, in practice, some authorities (i.e., coercers) may force cloud storage providers to expose user secrets or confidential data on the cloud, thus altogether circumventing storage encryption schemes. Given such fake user secrets, outside coercers can only obtained forged data from a user’s stored cipher text. Once coercers think the expected secrets are real, they will be satisfied and more importantly cloud storage providers will not have exposed any real secrets. Therefore, user privacy is still protected. This concept comes from a particular kind of encryption scheme called deniable encryption.

3.2 HOLDER PROCESS

Vendor component is to upload their records using some admittance policy. First they get the public key for particular upload folder after receiving this public key vendor request the covert key for particular upload file. Using that secret key owner upload their file.

3.3 CLIENT PROCESS

This component is worn to help the client to investigate the file using the file id and file name .If the file id and name is inaccurate means we do not obtain the file, if not attendant ask the public key and get the encryption file. If the user wants the decryption file means user have the secret key.

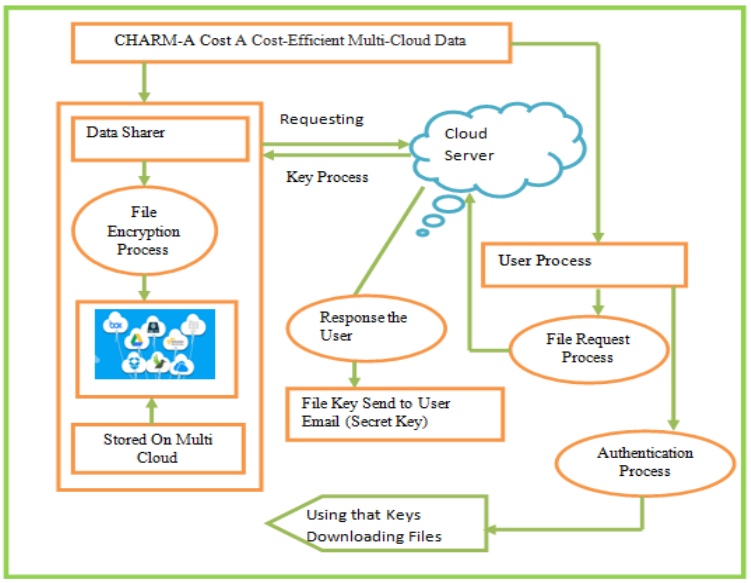

Fig 1.2 Framework of CHARM

IV. ALGORITHM

The key design of this heuristic algorithm can be described as follows:

In the first assign each cloud a value which is calculated based on four factors to indicate the preference of a cloud. We choose the most preferred n clouds, and then heuristically swap the cloud in the favored set with the cloud in the complementary set to search better solution. This is comparable to the thought of Kernighan-Lin heuristic algorithm, which is applied to effectively partition graphs to minimize the sum of the expenses on all boundaries cut. The partiality of a cloud is impacted by the four factors, and they have different weights. The accessibility is the advanced the better, and the price is the lower the better.

V. CONCLUSION

Cloud services are experiencing fast growth and the services based on multi-cloud also become prevailing. One of the nearly everyone concerns, when moving services into clouds, is capital expenditure. So, in this paper, we design a novel cargo space system CHARM, which guides customers to distribute data among clouds cost-effectively. CHARM makes fine-grained decisions about which cargo space mode to utilize and which clouds to place data in. The evaluation proves the efficiency of CHARM.

REFERENCES

[1] Aliyun OSS (Open Storage Service). [Online]. Available: http:// www.aliyun.com/product/oss, 2014.
[2] Gartner: Top 10 cloud storage providers. [Online]. Available: http://www.networkworld.com/news/2013/010313-gartnercloud- storage-265459.html? page=1, 2013.
[3] Z. Li, C. Jin, T. Xu, C. Wilson, Y. Liu, L. Cheng, Y. Liu, Y. Dai, and Z.-L. Zhang, "Towards network-level efficiency for cloud storage services," in Proc. ACM SIGCOMM Internet Meas. Conf., 2014, pp. 115–128.
[4] Z. Li, C. Wilson, Z. Jiang, Y. Liu, B. Y. Zhao, C. Jin, Z.-L. Zhang, and Y. Dai, "Efficient batched synchronization in dropbox-like cloud storage services," in Proc.ACM/IFIP/USENIX 14th Int. Middleware Conf., 2013, pp. 307–327.
[5] C. M. M. Erin Allen, "Library of congress and duracloud launch pilot program using cloud technologies to test perpetual access to digital content," Library Cong., News Releases. [Online]. Available: http://www.loc.gov/today/pr/2009/09-140.html, 2009.
[6] Zhenhua Li, Cheng Jin, Tianyin Xu, Christo Wilson, Yao Liu, Linsong Cheng, Yunhao Liu, Yafei Dai and Zhi-Li ZhangTowards "Network-level Efficiency for Cloud Storage Services" ACM 978-1-4503-3213, http://dx.doi.org/10.1145/2663716.2663747. 2014.

[7] R. Rodrigues and B. Liskov, "High availability in DHTs: Erasure coding vs. replication," in Proc. 4th Int. Conf. Peer-to-Peer Syst., 2005, pp. 226–239
[8] Alysson Bessani, Miguel Correia, Bruno Quaresma, Fernando Andre and Paulo Sousa ," DEPSKY: Dependable and Secure Storage in a Cloud-of-Clouds" ACM 978-1-4503-0634-8/11/04. . , 2011.
[9] Patrick Wendell, Joe Wenjie Jiang, Michael J. Freedman, and Jennifer Rexford," DONAR: Decentralized Server Selection for Cloud Services" , ACM 978-1-4503-0201-2/10/08 ...$10.00. ACM 978-1-4503-0201-2/10/08 ..., 2010.
[10] Apache Libcloud. [Online]. Available: http://libcloud.apache. org/, 2014.
[11] Mohammad Hajjat, Xin Sun, Yu-Wei Eric Sung, David Maltz, Sanjay Ra and Mohit Tawarmalani," Cloudward Bound: Planning for Beneficial Migration of Enterprise Applications to the Cloud", ACM 978-1-4503-0201-2/10/08, 2010.
[12] C. Huang, H. Simitci, Y. Xu, A. Ogus, B. Calder, P. Gopalan, J. Li, and S. Yekhanin, "Erasure coding in windows azure storage," in Proc. USENIX Annu. Tech. Conf., 2012.
[13] H. V. Madhyastha, J. C. McCullough, G. Porter, R. Kapoor, S. Savage, A. C. Snoeren, and A. Vahdat, "scc: Cluster storage provisioning informed by application characteristics and SLAs," in Proc. USENIX Conf. File, Stroage Technol., 2012.
[14] A. Bessani, M. Correia, B. Quaresma, F. Andre, and P. Sousa, "DepSky: Dependable and secure storage in a cloud-of-clouds," in Proc. 6th Conf. Comput. Syst., 2013, pp. 31–46.
[15] H. B. Ribeiro and E. Anceaume, "Datacube: A P2P persistent data storage architecture based on hybrid redundancy schema," in Proc. 18th Euromicro Int. Conf. Parallel, Distrib. Netw.-Based Process., 2010, pp. 302–306.
[16] DuraCloud. [Online]. Available: http://www.duracloud.org/,2014.
[17] S. Liu, X. Huang, H. Fu, and G. Yang, "Understanding data characteristics and access patterns in a cloud storage system," in Proc. 13th IEEE/ACM Int. Symp. Cluster, Cloud, Grid Comput., 2013, pp. 327–334.
[18] M. A. Shah, M. Baker, J. C. Mogul, and R. Swaminathan, "Auditing to keep online storage services honest," in Proc. 11th USENIX Workshop Hot Topics Oper. Syst., 2007.

