ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 9, March 2018

ON MINIMIZING ENERGY COST IN INTERNET SCALE SYSTEM WITH DYNAMIC DATA

A.vel[1]
M.E Student,
Mrk Institute of technology,
Kattumannarkoil.

A.Akilan[2]
Assistant Professor
Mrk Institute of Technology,
Kattumannarkoil.

ABSTRACT

The Online Social Network (OSN) services of dynamic Internet-scale systems by fully exploiting the energy efficiency in geographic diversity, data availability, and fault tolerance, large scale OSN services are often deployed across multiple sites at diverse locations. Intelligently managing and allocating resources among various clients is important for system provider. In our paper we propose a model to minimize the cost spent upon OSN also the quality of service is improved a Genetic Based Data Placement (GBDP) Technique. It is used to balances the tradeoff between energy cost and delay performance. Sometimes the reservation plan is satisfying the consumer's future demand and provider's resource prices and energy cost.

1. INTRODUCTION

1.1 Overview

Using a network of remote servers hosted on the Internet to store, manage, and process data, rather than a local server or a personal computer. Cloud users no longer physically possess their data, so how to ensure the integrity of their outsourced data becomes a challenging task. Recently proposed schemes such as provable data possession and proofs of retrieve ability are designed to address this problem, but they are designed to audit static archive data and therefore lack of data dynamics support. Moreover, threat models in these schemes usually assume an honest data owner and focus on detecting a dishonest cloud service provider despite the fact that clients may also misbehave.

Cloud computing is considered the evolution of a variety of technologies that have come together to change an organizations' approach for building their IT infrastructure. Actually, there is nothing new in any of the technologies that are used in the cloud computing where most of these technologies have been known for ages. It is all about making them all accessible to the masses under the name of cloud computing. Cloud is not simply the latest term for the Internet, though the Internet is a necessary foundation for the cloud, the cloud is something more than the Internet.

The cloud is where you go to use technology when you need it, for as long as you need it. You do not install anything on your desktop, and you do not pay for the technology when you are not using it.

✓Public Cloud

Public clouds are owned and operated by third parties; they deliver superior economies of scale to customers, as the infrastructure costs are spread among a mix of users, giving each individual client an attractive low-cost model. All customers share the same infrastructure pool with limited configuration, security protections, and availability variances. These are managed and supported by the cloud provider. One of the advantages of a Public cloud is that they may be larger than an enterprises cloud, thus providing the ability to scale seamlessly, on demand.

✓Private Cloud

Private clouds are built exclusively for a single enterprise. They aim to address concerns on data security and offer greater control, which is typically lacking in a public cloud.

✓There are two variations to a private cloud:

- ✓ On-premise Private Cloud: On-premise private clouds, also known as internal clouds are hosted within one own data center. This model provides a more standardized process and protection, but is limited in aspects of size and scalability.
- ✓ Externally hosted Private Cloud: This type of private cloud is hosted externally with a cloud provider, where the provider facilitates an exclusive cloud environment with full guarantee of privacy. This is best suited for enterprises that don't prefer a public cloud due to sharing of physical resources.

1.1.1 Cloud Computing Types:

- ✓ Infrastructure as a Service
- ✓ Software as a service
- ✓ Platform as a Service

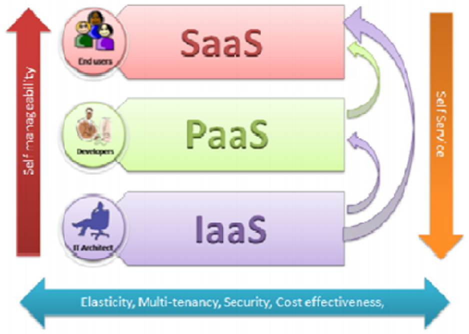

Figure 1.Cloud service models

Fig 1.1.1 Cloud computing types

✓ Infrastructure as a Service:

It is a provision model in which an organization outsources the equipment used to support operations, including storage, hardware, servers and networking components. The service provider owns the equipment and is responsible for housing, running and maintaining it.

✓ Software as a Service:

It is a software distribution model in which applications are hosted by a vendor or service provider and made available to customers over a network, typically the Internet. It is becoming an increasingly prevalent delivery model as underlying technologies that support Web services and service-oriented architecture become increasingly available.

✓ Platform as a Service:

It is an outgrowth of Software as a Service (SaaS). It is a way to rent hardware, operating systems, storage and network capacity over the Internet. The service delivery model allows the customer to rent virtualized servers and associated services for running existing applications or developing and testing new ones.

1.1.2 Essential characteristics of the Cloud

✓ On-demand self-service

A service consumer can automatically make use of the computing capabilities, such as server processing time and network storage without requiring human interaction with each service's provider.

✓ Broad network access

Cloud capabilities are available over the network and accessed through various platforms.

✓ Resource pooling

The provider's computing resources (HW and SW) are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to users' demand. Multi-tenancy is the most important feature of the cloud-based application.

✓ Rapid elasticity

Capabilities can be rapidly and elastically provisioned; it can be quickly scaled out, and quickly scaled in. For the user, the capabilities available for provisioning appear to be unlimited and can be purchased in any quantity at any time.

1.2 OBJECTIVES OF THE STUDY

- ✓ To Cloud computing provides a flexible and convenient way for data sharing, which brings various benefits for both the society and individuals.
- ✓ To exist a natural resistance for users to directly outsource the shared data to the cloud server since the data often contain valuable information.
- ✓ Identity-based encryption is a promising cryptographically primitive to build a practical data sharing system.
- ✓ The performance comparisons indicate that the proposed RS-IBE scheme has advantages in terms of functionality and efficiency cost-effective data-sharing system.

2. LITERATURE SURVEY

2.1 Group formation in large social network E.Goh,H.Shachan, N.Modadugu,and D.Boneh.2011

The paper presents SiRiUS, a secure file system designed to be layered over insecure network and P2P file systems such as NFS, CIFS, Ocean Store, and Yahoo! Briefcase. SiRiUS assumes the network storage is un trusted and provides its own read-write cryptographic access control for file level sharing. Key management and revocation is simple with minimal out-of-band communication. File system freshness guarantees are supported by SiRiUS using hash tree constructions. SiRiUS contains a novel method of performing file random access in a cryptographic file system without the use of a block server. Extensions to SiRiUS include large scale group sharing using the NNL key revocation construction. The implementation of SiRiUS performs well relative to the underlying file system despite using cryptographic operations.

2.2 Improved Proxy Re-Encryption Schemes with Applications to Secure Distributed Storage G. Ateniese, K. Fu, M. Green, and S. Hohenberger, 2013

Atomic proxy re-encryption, in which a semi-trusted proxy converts a cipher text for Alice into a cipher text for Bob without seeing the underlying plaintext. We predict that fast and secure re-encryption will become increasingly popular as a method for managing encrypted file systems. Although efficiently computable, the wide-spread adoption of BBS re-encryption has been hindered by considerable security risks. Present new re-encryption schemes that realize a stronger notion of security, and demonstrate the usefulness of proxy re-encryption as a method of adding access control to a secure file system. Performance measurements of our experimental file system demonstrate that proxy re-encryption can work effectively in practice.

2.3 Characterizing user behavior in online social networks Shucheng Yu, Cong Wang, Kui Ren, and Weijing Lou.2015

The project addresses this challenging open issue by, on one hand, defining and enforcing access policies based on data attributes, and, on the other hand, allowing the data owner to delegate most of the computation tasks involved in fine-grained data access control to un trusted cloud servers without disclosing the underlying data contents. We achieve this goal by exploiting and uniquely combining techniques of attribute-based encryption (ABE), proxy re-encryption, and lazy re-encryption. The scheme also has salient properties of user access privilege confidentiality and user secret key accountability. Extensive analysis shows that the proposed scheme is highly efficient and provably secures under existing security models.

2.4 Data A fast and high quality multilevel scheme for partitioning irregular graphs V. Goyal, O. Pandey, A. Sahai, and B. Waters 2014

As more sensitive data is shared and stored by third-party sites on the Internet, there will be a need to encrypt data stored at these sites. One drawback of encrypting data, is that it can be selectively shared only at a coarse-grained level (i.e., giving another party your private key). It develops a new cryptosystem for fine-grained sharing of encrypted data that we call Key-Policy Attribute-Based Encryption (KP-ABE). The cryptosystem, cipher texts are labeled with sets of attributes and private keys are associated with access structures that control which cipher texts a user is able to decrypt. It demonstrates the applicability of our construction to sharing of audit-log information and broadcast encryption.

3. SYSTEM ANALYSIS

3.1 EXISTING SYSTEM

In existing system data in cloud are randomly placed there by the cost to access the data is increased. However, it ignores two major sources of cost to datacenter operator's reservation plan of allocating resources, and over-provisioning data center capacity to tolerate highly datacenter utilization. Here we use a geography online social network for the cost of cloud storage and inter cloud communication. The request of one cloud by one time for the data availability is complex. Where single service providers compete for resource in a static manner, and show that there is a Nash equilibrium solution which is socially optimal and increases the resource prices, that is not effectiveness of our solution in realistic settings.

3.2 PROPOSED SYSTEM

Cloud services continue to grow to span large numbers of datacenters, making it increasingly urgent to develop automated techniques to place application data across these datacenters. Based on the analysis of month long traces from two large-scale commercial cloud services. By including candidate locations for datacenters in genetic based data placement technique is used, the operator can identify which combination of additional sites improve latency at modest costs in greater inter-datacenter traffic.

3.2.1Advantages for proposed system

- ✓ It supports public auditing on shared data stored in the cloud. In particular, we exploit ring signatures to compute verification metadata needed to audit the correctness of shared data.
- ✓ Then identity of the signer on each block in shared data is kept private from public verifiers, who are able to efficiently verify shared data integrity without retrieving
- ✓ In addition, our mechanism is able to perform multiple auditing tasks simultaneously instead of verifying them one by one.

3.2.2 Block diagram

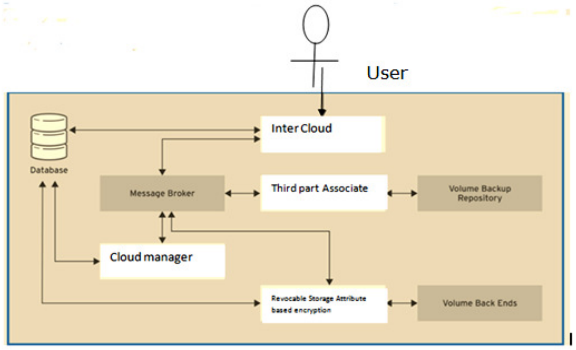

Fig 3.2.1 Block diagram

We first model the cost, the QoS, and the data availability of the OSN service upon clouds. Our cost model identifies different types of costs associated with multi cloud OSN while capturing social locality an important feature of the OSN service that most activities of a user occur between herself and her neighbors. Guided by existing research on OSN growth and our analysis of real-world OSN dynamics, our model approximates the total cost of OSN over consecutive time periods when the OSN is large in user population but moderate in growth, enabling us to achieve the optimization of the total cost by independently optimizing the cost of each period. Our QoS model links the QoS with OSN users' data locations among clouds. For every user, all clouds available are sorted in terms of a certain quality metric therefore, every user can have the most preferred cloud, the second most preferred cloud, etc. The QoS of the OSN service is better if more users have their data hosted on clouds of a higher preference.

4. SYSTEM ARCHITECTURE

4.1 SYSTEM DIAGRAM

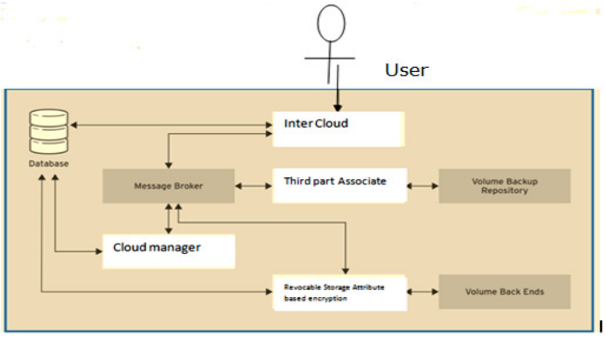

Fig 4.1 system diagram

We first model the cost, the QoS, and the data availability of the OSN service upon clouds. Our cost model identifies different types of costs associated with multi cloud OSN while capturing social locality an important feature of the OSN service that most activities of a user occur between herself and her neighbors. Guided by existing research on OSN growth and our analysis of real-world OSN dynamics, our model approximates the total cost of OSN over consecutive time periods when the OSN is large in user population but moderate in growth, enabling us to achieve the optimization of the total cost by independently optimizing the cost of each period. Our QoS model links the QoS with OSN users' data locations among clouds. For every user, all clouds available are sorted in terms of a certain quality metric therefore, every user can have the most preferred cloud, the second most preferred cloud, etc. The QoS of the OSN service is better if more users have their data hosted on clouds of a higher preference.

5. MODULE DESCRIPTION

5.1Modules

- ✓ Inter cloud Manager Module
- ✓ User Member Module
- ✓ File Security Module
- ✓ Signature verification Module
- ✓ Data Repository Module.
- ✓ Storage and traffic cost module

5.1.1 Inter cloud Manager Module

In this module, we create a local Cloud and provide priced abundant storage services. The users can upload their data in the cloud. We develop this module, where the cloud storage can be made secure. However, the cloud is not fully trusted by users since the ISP are very likely to be outside of the cloud users' trusted domain. Similar to we assume that the cloud server is honest but curious. That is, the cloud server will not maliciously delete or modify user data due to the protection of data auditing schemes, but will try to learn the content of the stored data and the identities of cloud users.

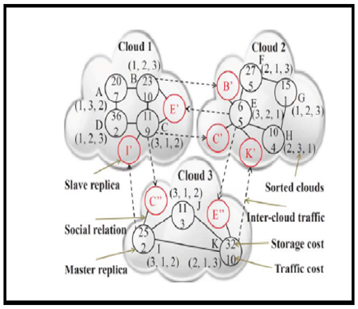

Fig 5.1.1 Inter Cloud Manager Module

5.1.2 User Member Module

Therefore, we assume that the group manager is fully trusted by the other parties. The Group manager is the admin. The group manager has the logs of each and every process in the cloud. The group manager is responsible for user registration and also user revocation too. A user party verifier cannot get any information about the client's data m from the protocol execution. Hence, the protocol is private against third party verifiers. If the server modifies any part of the client's data, the client should be able to detect it furthermore; any third Party verifier should also be able to detect it. In case a third party verifier verifies the integrity of the client's data, the data should be kept private against the third party verifier.

5.1.3 Signature Verification Module

The user membership is dynamically changed, due to the staff resignation and new employee participation in the company. The group member has the ownership of changing the files in the group. We describe a protocol based on this hash function which prevents 'cheating' in a data transfer transaction, while placing little burden on the trusted third party that oversees the protocol. I also describe a cryptographic protocol based on similar principles, through which a proverb can demonstrate possession of an arbitrary set of data known to the verifier. The verifier isn't required to have this data at hand during the protocol execution, but rather only a small hash of it.

5.1.4 File Security Module

Encryption is the process of converting data to an unrecognizable or "encrypted" form. It is commonly used to protect sensitive information so that only authorized parties can view it. We propose an enhanced dynamic proof of retrieve ability scheme supporting public audit ability and communication-efficient recovery from data corruptions. To this end, we split up the data into small data blocks and encode each data block individually using network coding before outsourcing.

1. Encrypting the data file.
2. File stored in the cloud can be deleted by either the group manager or the data owner.

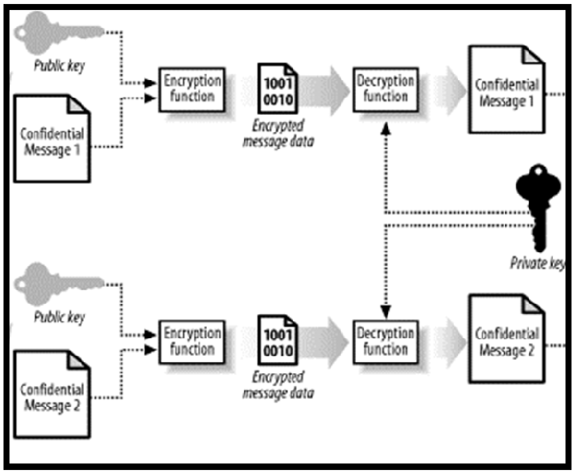

Fig 6.1.4 File security Module

5.1.5 Data Repository Module

A group signature scheme allows any member of the group to sign messages while keeping the identity secret from verifiers. Besides, the designated group manager can reveal the identity of the signature's originator when a dispute occurs, which is denoted as traceability. User revocation is performed by the group manager via a public available revocation list based on which group members can encrypt their data files and ensure the confidentiality against the revoked users.

5.2 METHODOLOGY

5.2.1 Revocable Storage Attribute-Based method

Genetic data is a public key cryptography primitive for one-to-many communications. In KP-ABE, data are associated with attributes for each of which a public key component is defined. User secret key is defined to reflect the access structure so that the user is able to decrypt a cipher text if and only if the data attributes satisfy his access structure.

✓ Key generation

Consider a hash function H : (0, 1)* H : (0, 1)s be a cryptographic hash function. To sign a message M = (0, 1)k and choose 2 * k random numbers Xij with 1 ≤ i ≤ k and j = {0,1}. For each i and j compute Yij= H (Xij). Here Yij are the public by key and the Yij are the private key values for each 2 * k values.

✓ Signing a message

Given a message M = m1, m2, ...,mk with mi∑{0,1} and the private keys Xij with 1≤ i ≤ k and j = {0,1} .We have to check mi equals to 0 or 1 for each i. If it is 0, then sigi = Xi0;otherwise sigi = Xi1 The signature is the concatenation of all sigifor i = {1,...... k} . Therefore, sig is evaluated as sig=(sig1//sig2//......//sigk) with// denotes the concatenation operator.

✓ Signature verification

For a given message M = m1, m2, ...mk with mi ∑{0,1} and the signature of the given message is sig = (sig1 // sig2 // ...... // sigk) and Yij is the corresponding public key of the Lamport One-Time signature scheme. For each 1≤ i ≤ k the hash value H (sigi) gets computed. If on mi = 0 then H (sigi) must be H (sigi) = Yi0 otherwise. H (sigi) must be H (sigi) = Yij to be a valid signature. The Loss scheme is the big size of the public and private key. which we will focus on throughout the rest of this paper. Note that calculating requires the storage cost and the traffic cost of_each user in. For any cost optimization mechanism that runs_at the beginning of a billing period, estimation is required to predict each user's costs during this billing period. Let us for now deem that the costs can be predicted and known.

5.2.2 Genetic Based Data Replacement Method

The load balancing abides by three rules such as the location rule, the distribution rule, and the selection rule. Here the work will process through a dynamic process after doing scheduling server. Firstly the tasks will be fixed a number. Afterward it will auto execute task number and size randomly. Then the task handled from a task slot, where the randomly generator are deposited for processed. Mainly the development of cloud computing is dynamic but for the arrangement purpose it can be expressed as allocating N number of jobs applied by the cloud customers to M number of processing unit in cloud. Every processing unit has a processing unit vector where vector consists of that mean how many million instructions can be processed by the machine per second. R is indicated as delay cost and x is cost of execution of instruction.

6. CONCLUSION

In our project, we study the problem of optimizing the monetary cost spent on cloud resources when deploying an online social network service over multiple geo-distributed clouds. I model the cost for data placement, quality of service with our vector approach, and address of data availability by ensuring a minimum number of replicas for each user. Based on these models, I present the optimization problem of minimizing the total cost while ensuring the QoS and the data availability. I propose as our algorithm. By extensive evaluations with large-scale Twitter data, is verified to incur substantial cost reductions over existing, state-of-the-art approaches.

7. FUTURE ENHANCEMENT

As a further work for a more efficient system, a joint optimization can be conducted to simultaneously control the data query, content placement, and replication at the back-end infrastructures, as well as the request mapping at the front-end servers.

REFERENCES

[1] Amazon, Seattle, WA, USA, “Amazon EC2 pricing,” [Online]. Available:
http://aws.amazon.com/ec2/pricing/

[2] Amazon, Seattle,WA, USA, “Case studies,” [Online].Available: http:// aws.amazon.com/solutions/case-studies/

[3] Microsoft, Redmond, WA, USA, “Pricing overview: Microsoft Azure,” [Online]. Available: http://www.windowsazure.com/en-us/ pricing/details/

[4] Facebook, Menlo Park, CA, USA, “Facebook Newsroom,” [Online]. Available: http://newsroom.fb.com

[5] GitHub, “Twitter/Gizzard,” [Online]. Available: http://github.com/ twitter/gizzard

[6] HubSpot, Inc., Cambridge, MA, USA, “State of the Twittersphere,” 2010.

[7] A. Abou-Rjeili and G. Karypis, “Multilevel algorithms for partitioning power-law graphs,” in Proc. IPDPS, 2006, pp. 1–10.

[8] S. Agarwal et al., “Volley: Automated data placement for geo-distributed
cloud services,” in Proc. NSDI, 2010.