

SECURE SHARING OF MEDICAL DATA IN BIG DATA ANALYSIS

D. Jayakumari¹, D. Suganthi², P. Shebi³, J. Cimiya John⁴

¹Associate Professor, ²Assistant Professor

²Department of Information Technology,

¹Department of ECE,

^{1,2,3,4} RVS College of Engineering and Technology, Coimbatore, Tamil Nadu, India

Abstract— Cross cloud computing is the piece of software that is portable and can run on multiple platforms without human intervention. Cross cloud service composition provides a concrete approach that is capable for large scale big data processing. We propose a privacy-aware cross-cloud service composition method, named Hire Some-II (History record-based Service optimization method) based on its previous basic version Hire Some-I. In our method, to enhance the credibility of a composition plan, the evaluation of a service is promoted by some fits Qos history records. Using Hiresome-II technique time consumption for searching of data can be reduced. Cloud computing is an internet based computing, which shares resources, software and information are provided to computers and other devices on demand, in this security is the major concern.

I. INTRODUCTION

Cloud Computing is a technology that uses the internet and central remote servers to maintain data and applications. [2] Cloud computing allows consumers and businesses to use applications without installation and access their personal files at any computer with internet access. This technology allows for much more efficient computing by centralizing data storage, processing and bandwidth. Cloud computing and big data still in constant evolution, they provide a cost effective and scalable infrastructure to provide big data and business analytics. The rise of cloud computing and cloud data stores have been a precursor and facilitator to the emergence of big data. Cloud computing is the co modification of computing time and data storage by means of standardized technologies. Cloud computing has become a highly demanded service or utility due to the advantages of high computing power, cheap cost of services, high performance, scalability, accessibility as well as availability. Some cloud vendors are experiencing growth rates of 50% per annum, but due to being in a stage of infancy, it still has pitfalls that need proper attention to make cloud computing services more reliable and user friendly.

Cloud computing poses privacy concerns because the service provider can access the data that is in the cloud at any time. It could accidentally or deliberately alter or even delete information. Many cloud providers can share information with third parties if necessary for purposes of law and order even without a warrant. That is permitted in their privacy policies, which users must agree to before they start using cloud services. Solutions to privacy include policy and legislation as well as end users' choices for how data is stored. Users can encrypt data that is processed or stored within the cloud to prevent unauthorized access.

Cloud computing also leverages concepts from utility computing to provide metrics for the services used. Such metrics are at the core of the public cloud pay-per-use models. In addition, measured services are an essential part of the feedback loop in autonomic computing, allowing services to scale on-demand and to perform automatic failure recovery. Cloud computing has evolved by addressing

the QoS (quality of service) and reliability problems. Cloud computing provides the tools and technologies to build data/compute intensive parallel applications with much more affordable prices compared to traditional parallel computing techniques.

II. LITERATURE SURVEY

M. Zhang, R. Ranjan, A. Haller proposed and investigated the cloud infrastructure services landscape advances steadily leaving users in agony of choice. Cloud Recommender is a new declarative approach for selecting Cloud-based infrastructure services. Cloud Recommender automates the mapping of users specified application requirement to cloud service configuration They formally capture cloud service configurations in ontology and provide its implementation in a structured data model which can be manipulated through both regular Expressions and SQL.

M. Li, S. Yu, Y. Zheng, K. Ren investigated Personal health record is maintain in the centralize server to maintain patient's personal and diagnosis information. Personal health record (PHR) is an emerging patient-centric model of health information exchange, which is often outsourced to be stored at a third party, such as cloud providers. However, there have been wide privacy concerns as personal health information could be exposed to those third party servers and to unauthorized parties. The security schemes are used to protect personal data from public access. To assure the patients' control over access to their own PHRs, it is a promising method to encrypt the PHRs before out sourcing. In this paper we propose novel patient-centric framework and suite of mechanism for data access control to PHR's stored in semi trusted servers. To achieve fine-grained and scalable data access control for PHRs, we leverage attribute based encryption techniques to encrypt each patient's PHR file. Data owner update the personal data into third party cloud data centers.

III. PROBLEM DEFINITION

Assume that if a user needs to access a particular data from the cross cloud service composition, the major problem is managing the massive amounts of data that increases the risk and magnitude of a potential data breach. Sensitive, personal data and confidential data can be exposed and violate compliance and data-security regulations. And also because of storing large amount of data, it takes time to search required data.

3.1 OBJECTIVE

To ensure protection of data between the users by providing the keys to users for the authorized access. Security is a very dangerous point in accessing confidential data hence attack can easily be done. To prevent this, authentication is required for sharing the contents securely. In order to send the contents securely a concept known as providing key is implemented hence allows only the authorized users to access the confidential data. To increase the high search performance indexed based searching is used.

IV. SYSTEM ANALYSIS

4.1 EXISTING SYSTEM

Current service optimization approaches often assume that the quality delivered by service providers does not change over time. However, this assumption is often spoiled by the dynamic network environment. Therefore, a service provider may fail to deliver their services with promised quality. It often causes some fluctuations in service performance, and the evaluation of a service will be suspicion, if the evaluation is scored by the QoS values promised in advance. History Record-Based Service Composition Evaluation k-means algorithm is used as data filtering tool to select representative history records. As the latest computing paradigm, cloud is characterized by delivering

hardware and software resources as virtualized services by which users are free from the burden of acquiring the low level system administration details.

4.1.1 DRAWBACKS

- In existing system, it is a challenge to tradeoff the privacy and the time cost in cross cloud service composition.
- Cloud service and its bandwidth probably fails to match to the cloud.
- It often causes some fluctuation in the service performance.

4.2 PROPOSED SYSTEM

In the proposed system presents a tree structure is recruited to specify the service composition context. Concretely, a Task-Service tree is defined, to incorporate a task and a group of candidate services into an integrated application context. Here, the candidate services are the qualified services that can fulfill the task execution's specification in functional and non-functional property. In addition index based searching is introduced for searching efficiency.

The history and indexed based records associated with a service's quality. As the index based searching is the purpose of storing an index is to optimize speed and performance in finding relevant documents for a search query. For example, while an index based searching can search 10,000 documents can be queried within milliseconds, a sequential scan of every word in 10,000 large documents could take hours. The additional computer storage required to store the index based searching hence the time saved during information retrieval.

4.2.1 ADVANTAGES

- It protects the privacy of cloud for cross cloud service composition but also greatly speeds up that calculating process for selecting a optimal service composition plan with higher optimality and precision.
- It is suitable for developing a cross cloud service composition plan over big data of history records with privacy consideration
- The tree mechanism is initiated by a Task Service tree which is a multiple fork tree and is more compatible for real life systems.

V. SYSTEM DESIGN

5.1 METHODOLOGY

Initially a key management system is provided in which Public-key algorithms uses a public key and a private key. The public key is made available to everyone. A sender encrypts data with the public key, only the holder of the private key can decrypt this data. It also uses Hiresome-II along with the index based searching which hires tree structure which increases high search performance .

5.2 SYSTEM ARCHITECTURE

Architecture diagram shows the relationship between different components of system. This diagram is very important to understand the overall concepts of system. Architecture diagram is a diagram of system, in which the principle parts or functions are represented by block connected by lines that shows the relationship of the block.

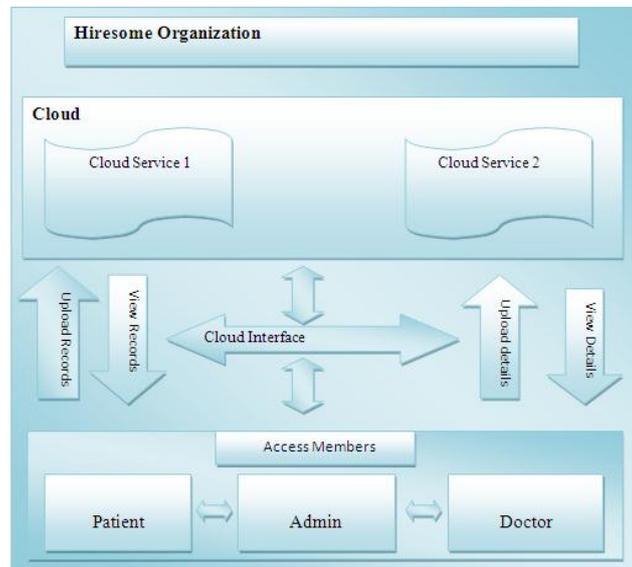
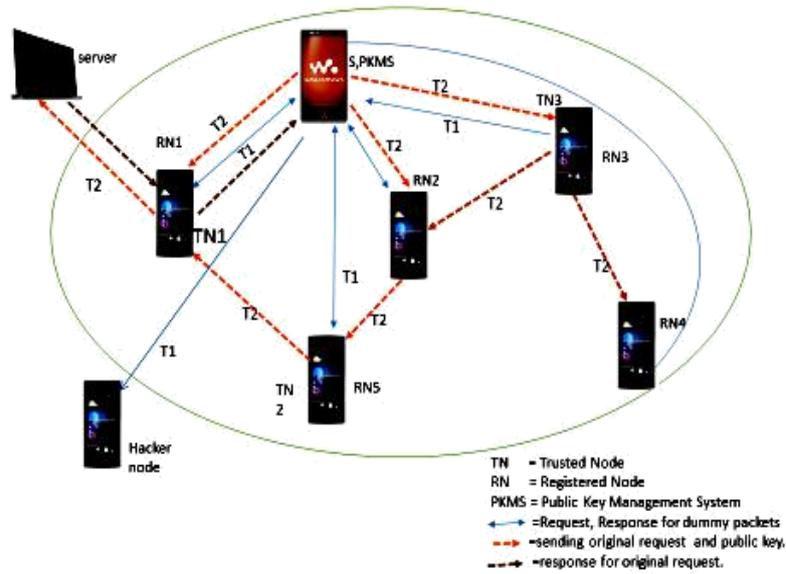
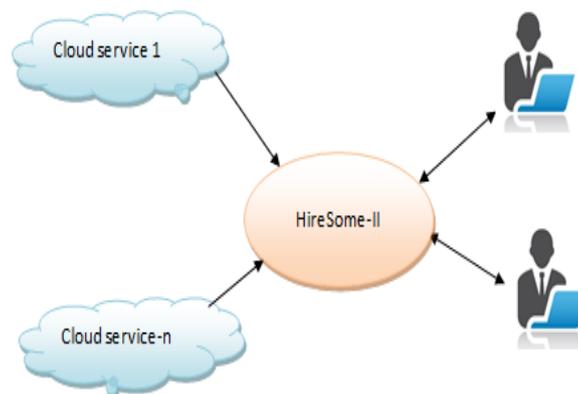


Figure5.1 hiresome-2 architecture

They are heavily used in the engineering world in hardware design, electronic design, software design and process flow diagram.



Cloud is an online storage and it can store large amount of data. While searching a data in a cloud from another cloud due to huge number of data the time taken to retrieve the data will take time. By using Index based searching these required data can be retrieved in less time. And also in a cloud security is very important. The administrator should know whether the data requestor have the authority to access the loud or not. This is done by providing keys to the users during first registration.

VI. SYSTEM IMPLEMENATION

6.1 Authentication

Login:

The user has to provide exact username and password which was provided at the time of registration, if login success means it will take up to main page else it will remain in the login page itself.

Key Generate and Forward:

The admin will generate key to the users thus ensuring privacy to the data from getting accessed by unauthorized users

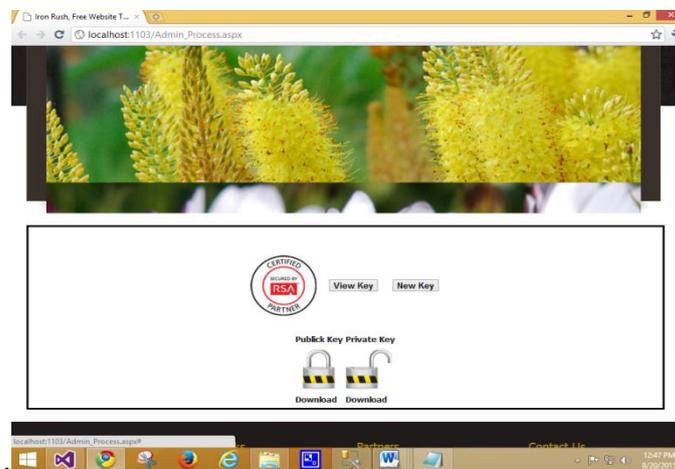


Figure 6.1 Admin key generation

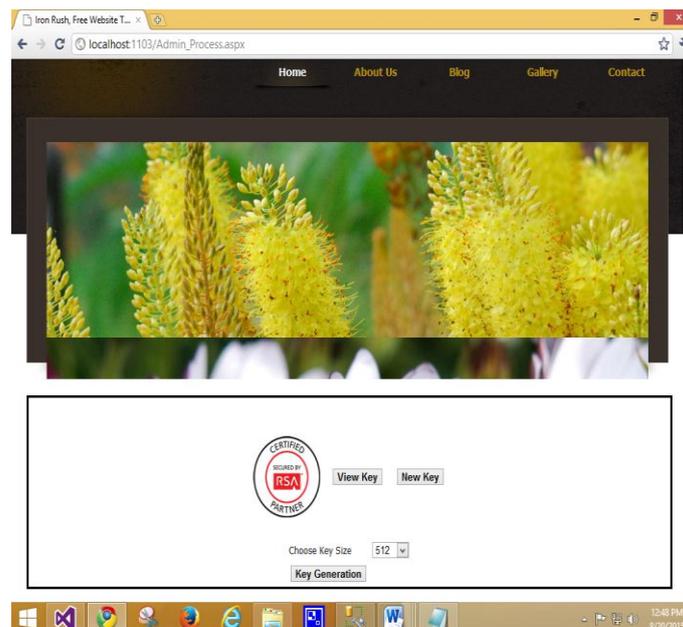


Figure 6.2 Key generation

6.2 Verification and Update data:

If the key used to login is checked by the admin from the database whether it is correct or not. After verification of the key if it is correct the users are allowed to access the data.

6.3 Searching of Data:

During searching of data as there were large number of data in the cloud it will took time to retrieve the required data. So by Indexed based searching methodology the data can retrieved easily in less time.

VII. PERFORMANCE ANALYSIS

In our system whenever an user gets sign in, the admin will provide keys and also during the log in of every users the admin will check the database whether the user is the authorized one or not. During searching of data by using Index based searching the required data can obtain in quick time.

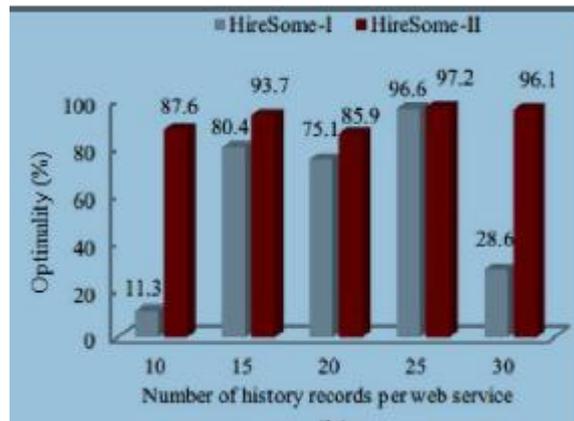


Fig7.1: Number of web services =3.

HireSome-I and HireSome-II performs are close to the Benchmark method, and HireSome-II performs better than HireSome-I in achieving a (close-to-) optimal plan.

VIII. CONCLUSION AND FUTURE ENHANCEMENTS

8.1 CONCLUSION

In our proposed scheme we are ensuring the protection of data between the users by providing the keys to users for the authorized access. In order to send the contents securely a concept known as providing key privacy is implemented hence allows only the authorized users to access the confidential data and also the high search performance is increased using the history and index based searching. Hence the time consumption is also reduced.

8.2 FUTURE ENHANCEMENT

Plan to apply our method to some specific cloud systems for processing big data applications. Besides, as the privacy preservation for big data analysis, share and mining is a challenging research issue due to increasingly larger volume of datasets in cloud, we also plan to investigate the scalability

of privacy preservation in big data applications with cloud service access. For future work, we plan to apply our method to some Specific cloud systems for processing big data applications[1].

REFERENCES

- [1] S. Tai, J. Nimis, A. Lenk, and M. Klems, "Cloud Service Engineering," in Proc. 32nd ACM/IEEE ICSE, May 2010, vol. 2, pp. 475-476.
- [2] V. Nallur and R. Bahsoon, "A Decentralized Self-Adaptation Mechanism for Service-Based Applications in the Cloud," IEEE Trans. Softw. Eng., vol. 39, no. 5, pp. 591-612, May 2013.
- [3] H.Y. Lin and W.G. Tzeng, "A Secure Erasure Code-Based Cloud Storage System with Secure Data Forwarding," IEEE Trans. Parallel Distrib. Syst., vol. 23, no. 6, pp. 995-1003, June 2012.
- [4] A. Gambi and G. Toffetti, "Modeling Cloud Performance with Kriging," in Proc. 34th ICSE, June 2012, pp. 1439-1440.
- [5] R. Buyya, C.S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, "Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility," Future Gener. Comput. Syst., vol. 25, no. 6, pp. 599-616, June 2009.
- [6] J. Cao, K. Hwang, K. Li, and A.Y. Zomaya, "Optimal Multiserver Configuration for Profit Maximization in Cloud Computing," IEEE Trans. Parallel Distrib. Syst., vol. 24, no. 6, pp. 1087-1096, June 2013.
- [7] M. Li, S. Yu, Y. Zheng, K. Ren, "Scalable and Secure Sharing of Personal Healthrecords in Cloud Computing Using Attribute Based Encryption" June 2013.