

Secured Multi Purpose Smart Card Authentication Using RFID in Multi Cloud using Big Data

Lydia Joyce.A¹, Audline Beena.S.P²

Student, Department of Computer Science, Sri Muthukumar Institute of Technology, Chennai, India

Assistant Professor, Department of Computer Science, Sri Muthukumar Institute of Technology,
Chennai, India

ABSTRACT- Cloud computing is a type of computing that depend on distribution of computing resources rather than having local servers or private devices to handle applications. Cloud computing uses networks of great groups of servers classically running low-cost consumer PC technology with particular connections to spread data-processing tasks across them. Big data is a wide-ranging stretch for data sets so large or composite that traditional data handling requests are insufficient. Cloud based Large Scale Big Data incorporation is still in Examination. That too Cross Cloud is the most multifaceted integration. Cross Cloud is applied by allocating tasks to the numerous web service area. Different Tasks or same can be conquered by different Cloud based on its accessibility. Big Data is incorporated in the cloud. Cloud providers containing Amazon Web Services (AWS), Sales force.com, or Google App Engine provide users the decisions to organise their presentation over a system. Two Cloud Servers Drop Box and Google drive with one Big Data Database Storage that is Mongo Data Base is also organised. Also Multi Access Smart Card Application with Ration card, Passport and Hospital Applications for User Access are organized. User Individual Authentication comprises User Name, Password, Primary Key and Radio Frequency Identification (RFID) Card are all stored and proved in Mongo Data Base. Complete Data is splitted and stored in Two Servers parallely. User Request is controlled by the Mongo Data Base and balance part is controlled by Drop Box and Google Drive. The distributed storage system not only supports safe and vigorous data storage and retrieval, but also lets a user forward the data in the storage servers to another user without recovering the information back.

Index Terms- Service Composition, Qos. MapReduce, Cloud, Big Data

I. INTRODUCTION

A model for providing data equipment facilities in which resources are recovered from the internet through web-based implements and applications, rather than a direct association to a server. Data and software packages are deposited in servers. However, cloud computing structure permits access to information as long as an electronic device has entree to the web. This type of administration permits workforces to work distantly.

Cloud computing permits customers and trades to use applications without installation and access

their individual records at some computer with internet access. This knowledge allows for much more effective computing by consolidating data storage, processing and bandwidth.

A simple instance of cloud computing is email, Gmail, Yahoo, or Hotmail etc. An internet connection is needed and can start transferring emails. The email management and server software is all on the cloud (internet) and is completely achieved by the cloud service provider Google, Yahoo etc.

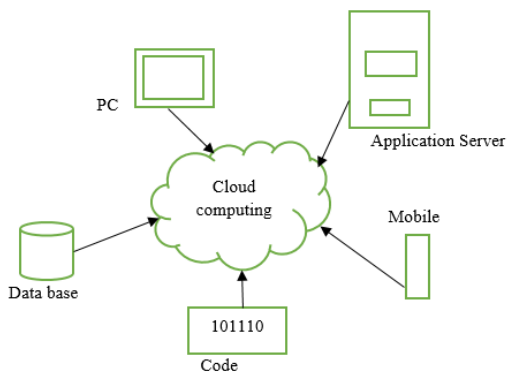


Fig. 1. Cloud computing

Cloud computing also prominences on manipulating the competence of the common possessions. Cloud resources are typically not merely common by several customers but are also vigorously modified per demand. This can work for assigning capitals to users. For instance, a cloud processor facility that serves European users through European business hours with a definite solicitation (e.g., email) may reallocate the similar incomes to serve North American users during North America's business periods with a changed presentation (e.g., a web server). This method helps exploit the use of calculating power while decreasing the overall cost of resources by using fewer power, rack space, air conditioning, etc. to preserve the system. With cloud computing, multiple operators can access a single server to recover and update their information without getting authorizations for different applications.

The name "moving to cloud" indicates to an association moving away after a out-dated Capital Expenditure model (buy the devoted hardware and depreciate it over a period of time) to the Operating Expense model (use a shared cloud organisation and pay as one uses it).

Promoters claim that cloud computing permits companies to avoid upfront organisation charges, and emphasis on tasks that differentiate their businesses instead of an

organisation. Promoters also claim that cloud computing permits initiatives to get their applications up and running faster, with better manageability and fewer maintenance. And also allows IT to more quickly regulate resources to meet changeable and unpredictable business petition. Cloud providers classically use a "pay as you go" model. This can lead to unpredictably great charges if administrators do not adjust to the cloud pricing model.

The current accessibility of high-capacity systems, low-cost mainframes and storage devices as well as the pervasive implementation of hardware virtualization, service concerned with architecture and autonomic utility computing have led to a growth in cloud computing. Businesses can scale up as figuring needs rise and then scale down again as demands decrease.

Cloud computing has now become a highly claimed service or efficacy due to the compensations of high computing power, high performance, scalability, cheap cost of services, accessibility as well as availability. Cloud retailers are experiencing development rates of 50% per annum.

The cloud makes it promising for retrieving the information from anywhere at any time. While a traditional computer arrangement requires to be in the identical location as data storage device, the cloud takes away that step. The cloud eradicates the need for the same physical position as the hardware supplies the data. The cloud supplier can both own and house the hardware and software essential to run the home or commercial applications.

This is particularly helpful for productions that cannot afford the identical quantity of hardware and software space as a bigger company. Small companies can store their statistics in the cloud, eradicating the cost of obtaining and storing memory devices. One obligation is that the need to have an

internet association in order to contact the cloud. This means to gaze at an unambiguous document that have accommodated in the cloud. Create an internet connection either through a wired or wireless internet or a movable broadband association. The advantage is that to access the same document from wherever any device can access the internet. These devices could be a desktop, laptop, phone, or tablet. This can also help the corporate to function more efficiently because anyone who can associate to the internet and the cloud can work on documents, access software, and store data. Visualize picking up the smartphone and transferring a pdf document to evaluate instead of devising to halt by the workplace to design it or upload it to the laptop. This is the liberty that the cloud can deliver for the organization.

Big data can be stored, processed, acquired, and examined in many ways. Every big data source has different characteristics, comprising the frequency, velocity, volume, type, and genuineness of the data. When big data is administered and stored, additional magnitudes come into play, such as ascendancy, policies, and security. Selecting an architecture and structure, a suitable big data solution is stimulating because so abundant influences have to be deliberated.

The "Big data designs and construction" chains grants a structured and pattern-based approach to simplify the task of describing an overall big data architecture. Because it is significant to measure whether a business consequence is a big data problem, pointers are encompassed to benefit control which commercial problems are respectable competitors for big data results. When big data is administered and stored, additional magnitudes come into play, such as ascendancy, policies, and security.

Big Data can be defined by the subsequent

Features

- (i) **Volume** – It is the scope of data which describes the value and potential of the data under deliberation and whether it can actually be considered as Big Data or not.
- (ii) **Variety** – The arrangement to which Big Data belongs to is also very necessary fact that needs to be known by the data analysts. This helps the people, who are carefully observing the data and are associated with it, to effectively use the data to their benefit and thus upholding the consequence of Big Data.
- (iii) **Velocity** – The term velocity signifies to the speed of generation of data or how fast the information is produced and achieved to meet the demands and the experiments which lie forward in the track of development and growth.
- (iv) **Variability** – This refers to the unpredictability which can be exposed by the data at times, thus hindering the process of being able to handle and achieve the data efficaciously.
- (v) **Complexity** – Data association can develop a very complex process, predominantly when enormous volumes of data come from many bases. These data need to be associated and connected in order to be able to understand the material that is supposed to be conveyed by these data.

Examples of Big data

Data is of two forms-

1. Structured Data

2. Unstructured Data

Structured data has a semantic meaning attached to it whereas unstructured data has no latent meaning.

II. RELATED WORK

Hsiao-Ying Lin et al. addressed [14] a threshold proxy re-encryption scheme which supports encoding, forwarding and partial decryption operations in a distributed way. A secure cloud storage organisation that delivers secure data forwarding and secure data storage functionality in a decentralized structure is presented. Storage servers act as storing content addressable blocks and key servers act as contact nodes for providing a front-end cover such as traditional file system interface. The file system and storage systems are highly compatible.

Ming Li et al. proposed [4] a threshold proxy re-encryption scheme and assimilate it with a safe regionalized code to form a secure distributed storage system. The system encounters the necessities that storage servers autonomously perform programming and re-encryption and important servers individually perform partial decryption. Errors are prevented due to systematic process. The data being stored in cloud database is in encoded set-up. Data's are promoted directly to another user by storage server directly under the command of the data owner. Efficient encryption techniques which reduce the time needed for encryption and decryption is failed.

Junwei Cao et al. investigated [24] a pricing model for cloud computing. A different queuing discipline change the distribution of the waiting times and changes the average task response time and the expected service charge. Two simple queuing disciplines, namely,

- **Shortest Task First (STF):** Tasks are arranged in a waiting queue in the increasing demand of their task implementation requests.
- **Largest Task First (LTF):** Tasks are arranged in a waiting queue in the decreasing demand of their task implementation requirements.

There is a lack of analytical result on waiting time distribution for STF and LTF. This can be extended to other service charge functions and also can be applied to other pricing models.

Rajkumar Buyya et al. predicted [21] an architecture for market-oriented distribution of resources within clouds and obtained an image for the formation of world-wide cloud conversation for trading services. Interaction protocols needs to be prolonged to maintain interoperability among diverse cloud service providers. Cloud computing and other connected patterns need to converge so as to produce unified and other related paradigms need to converge so as to produce unified and interoperable platforms for delivering IT services.

III. PRELIMINARY

A. Anonymization

Data anonymization is a type of information modification whose intent is privacy protection. It is the method of either removing personally or encrypting identifiable information from data sets, so that the people whom the data describe remain unidentified. Data anonymization allows the removal of data through a boundary, such as among two divisions inside an activity or between two activities, while reducing the risk of unintentional revelation and in certain atmospheres in a

manner that enables assessment and analytics post-anonymization.

In the framework of medical data, anonymized information denotes to data after which the patient cannot be familiar by the receiver of data. The name, full post code and address must be detached together with any other information which, in aggregation with other data held by or revealed to the recipient, could identify the patient.

B. MapReduce

MapReduce is the soul of Hadoop. The word MapReduce denotes to two distinct and separate tasks that Hadoop courses perform. The map job, takes a set of information and changes it into a different set of data, where distinct elements are fragmented into tuples (key/value pairs). The decrease job receipts the output from a plan as input and groups those data tuples into a smaller set of tuples. The reduce job is always executed after the map job. Cloud computing combined with MapReduce becomes more prevailing and adaptable as cloud offer in infrastructure resources on demand.

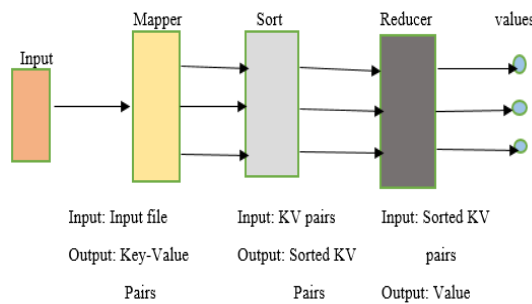


Fig. 2. MapReduce

K-means process

K-means is a clustering method that aims to find the positions $\mu_i, i=1....n$ of the clusters that minimize the square of the space from the data facts to the group.

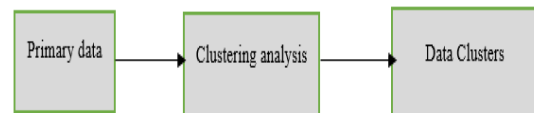


Fig. 3. K-means clustering

If k value is given, then the K-means algorithm can be executed through the following steps:

- The objects are partitioned into k non-empty sets.
- The cluster centroids (mean point) are identified.
- Each cluster is assigned to a specific point.
- The distances from each fact is found and the points are set to the cluster where the distance is minimum from the centroid.
- The centroid is found from the new cluster which has been formed after re-allocating the points.

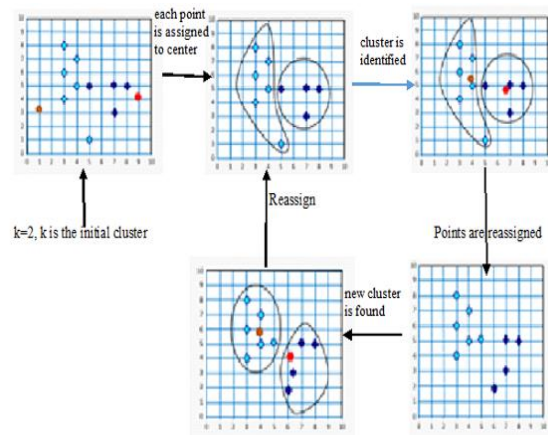


Fig. 4 . Step by step process for K-means

For defining cluster centres that diminish alteration, A key can be found by setting the partial derivative of Distortion with respect to each cluster center to zero.

$$A_j = \text{cluster}(z_i) = \arg_{x \min} \|y_i - \mu_j\|^2$$

$$\text{Distortion} = \sum_{i=1}^n (y_i - a_i)^2 = \sum_{k=1}^k \sum_{i \in k} (\mu_i - \mu_j)^2$$

$$\delta \text{Distortion} = \delta \sum_{i \in k} (\mu_i - \mu_j)^2 = -2$$

$$\Rightarrow \mu_j = 1/|\mu_j| \sum_{i \in k} y_i$$

For any k clusters, the value of k should be such that even if the value of k increases from after several levels of clustering the distortion rests constant. The attained fact is called the “Elbow”.

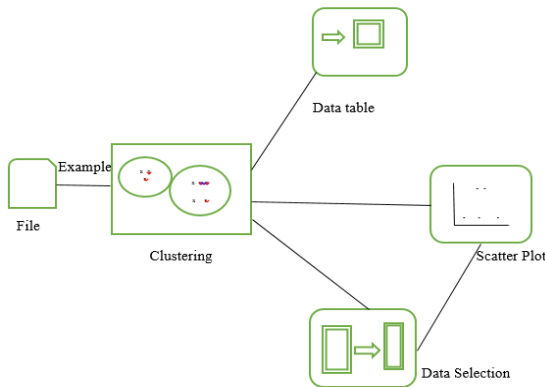


Fig. 5. Overall diagram for K-means Clustering

SYSTEM MODELS

User Interface

The User wants to create an account and then only they are permitted to contact the Network. As soon as the User generates an account, they are to login into their account and mandate the Work from the Service Provider.

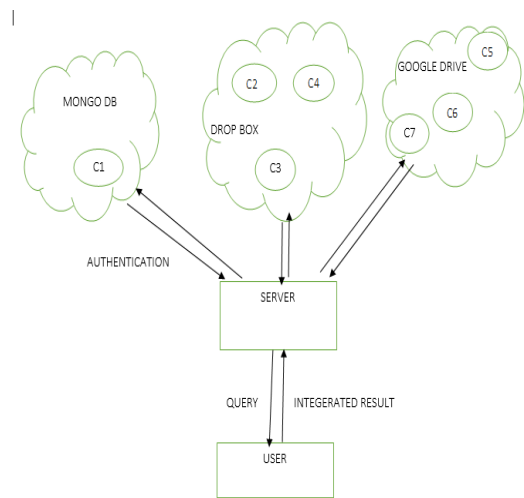


Fig. 6. Structural design

Based on the User’s request, the Service Provider will process the User’s requested Job and reply to them. All the User facts will be kept in the Database of the Data Service Provider. In this Project, the User Interface Frame to Connect with the Data Server through Network Coding using the programming Languages is designed. By sending the request to Server Provider, the User can access the requested data if they are authenticated by the Server.

Mongo DB setup

MongoDB is a cross-platform document-oriented database. Categorized as a NoSQL database, MongoDB avoids the out-dated table-based relational database structure making the combination of data in certain types of presentations easier and faster.

Google Drive and drop box

The two real time cloud one is drop box and another one is Google drive is deployed, a swing like application is created for the integration of cloud both Google drive and drop box, through the application the user get the registered and login to access the real time cloud. All the user information are stored and kept in the above mentioned clouds.

Big Data Setup

Big data is a comprehensive term for any collection of data sets so large and complex that it becomes interesting to practice using traditional data processing applications. The experiments include analysis, capture, duration, search, distribution, storage, transmission, picturing, and privacy defilements. The trend to larger data sets is due to the supplementary information derivable from analysis of a single large set of related data, as compared to separate lesser sets with the similar total amount of data, permitting associations to be found to spot commercial trends, prevent diseases, combat crime and so on. Big data is implemented here because every employ has an instructed information to make analysis on the data.

RFID Authentication and user credentials

Client is an application which is created and installed in the User's machine therefore can perform the activities. The Application's First Page Consist of the User registration Process. The User Login Page is created by RFID and Text Field password to the server. While creating username and password the RFID is stored in the database.

Data poisoning

A concept that user will have the ration card, hospital and passport or organization records is implemented that will maintain the information both private and public data. Therefore employee may contain private data like customer id, customer name, salary and the loan applied and public data like email id and phone number. But poisoning is the goal of the attacker to destroy or stealing client data, the relevant data is put in the two clouds.

Cross Cloud retrieval

Cloud Service Provider will contain the large quantity of information in their Data Storage. Similarly the Cloud Service provider will sustain all the User information to authenticate the User when they login into their account. The User information will be stored in the Database of the Cloud Service Provider. Similarly the Cloud Server will readdress the User demanded job to the Resource Assigning Module to process the User demanded Job. The Request of all the Users will practice by the Resource Assigning Module. To connect with the Client and with the additional elements of the Cloud Network, the Cloud Server will create connection among them. For this Purpose a User Interface Frame is created. Similarly the Cloud Service Provider will send the User Job request to the Resource Assign Module in (FIFO) First in First out manner.

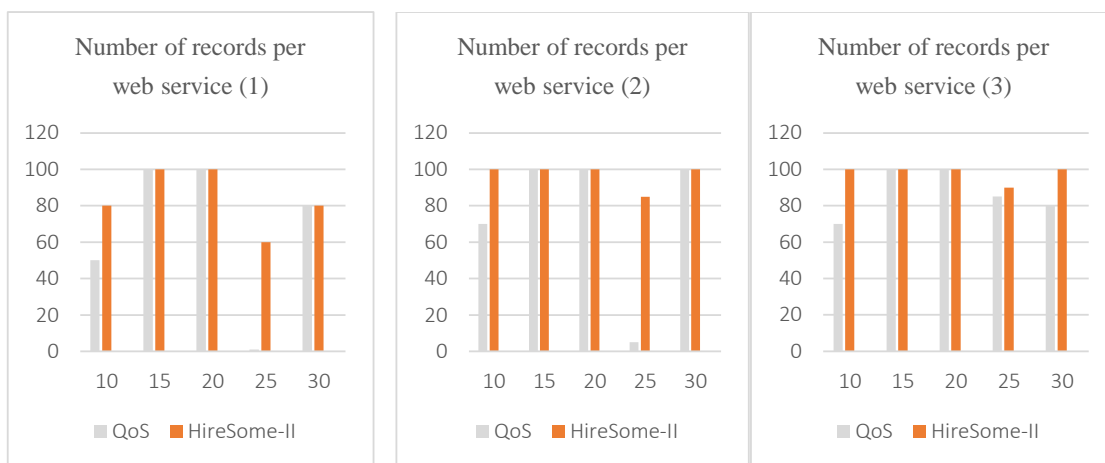
IV. COMPARISON ANALYSIS

For HireSome-II method, each applicant service has one or two illustrative records in a single level. The time when complexity is high, the average height of the history record tree is also high. To shorten the comparison examination without loss of generalisation, the number of tasks is fixed with 4, and the number of QoS standards is fixed with 5. Furthermore, 2, 3, 4, and 5 candidate web services are correspondingly allocated to a task and 10, 15, 20, 25, 30 QoS history records are allocated individually to a candidate service for further consideration. The figures indicate the estimation results with the expectations. With these expectations, the imitation is accompanied for 100 times originated by various set of history records.

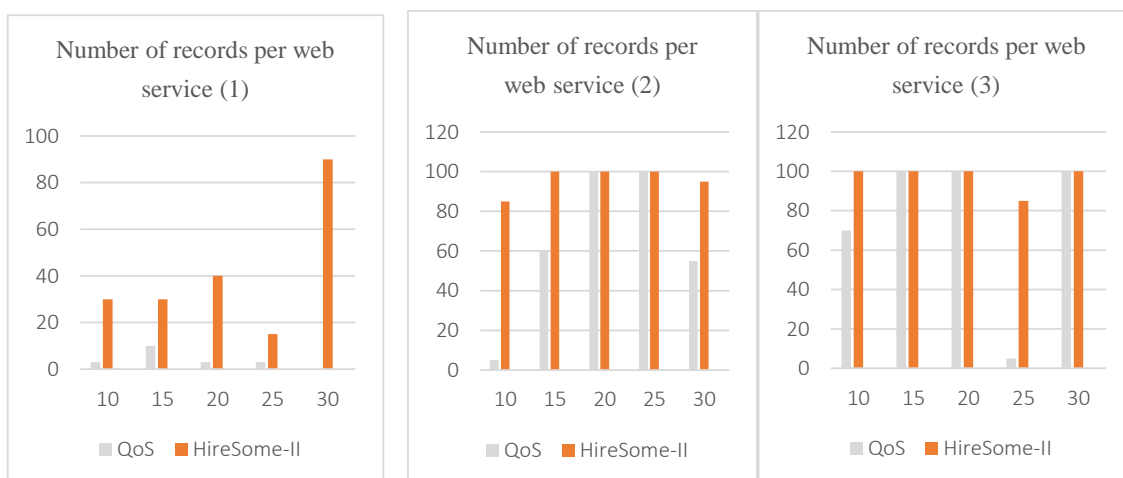
From the Fig. 7, compared with QoS and the Benchmark, HireSome-II have the minimum time to produce the optimum service configuration plan. Besides, in the Fig. 7, the time cost for producing the optimum service configuration plan is affected by the amount of history records per web service once the amount of web services per task is secure. Moreover, associated with the Benchmark process, the time cost of QoS and HireSome-II alters gradually with the increasing of the amount of history records per web service, whereas for the Benchmark, the time cost increases highly with the increasing of the number of history records per web

service. Hence, compared to the Benchmark, QoS and HireSome-II have an improvement in large-scale service composition.

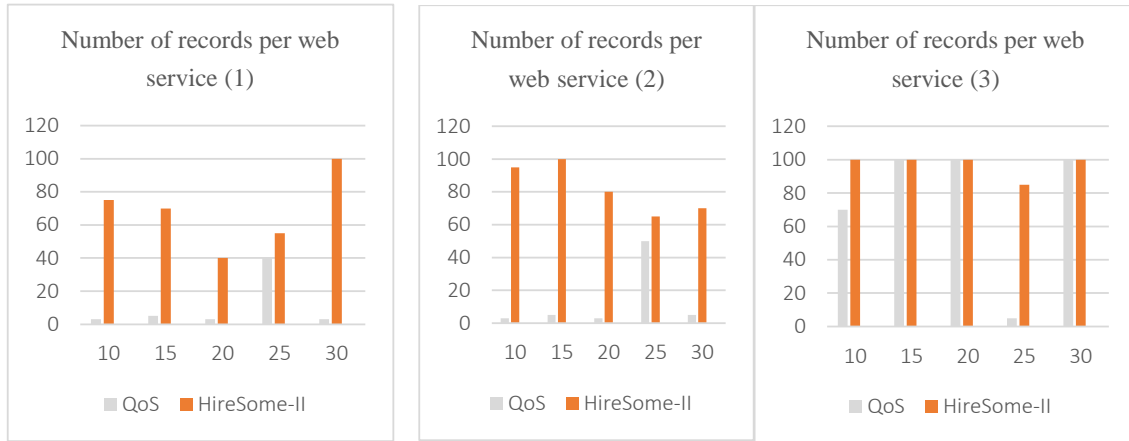
Choosing a set of facilities for a (near-) optimum composition plan in terms of QoS is critical when numerous corresponding functionally services are available [25]. Consequently, service composition is a definitive issue in service computing area. Quality-aware composition of web services has been fully examined in [7], [8], [9], [10], [11].



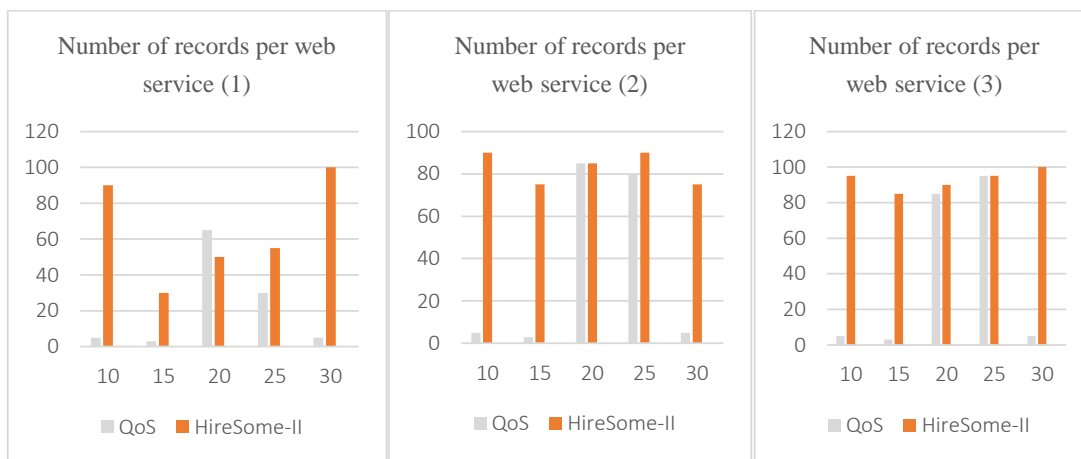
Number of web services= 2



Number of web services= 3



Number of web services= 4



Number of web services=5

Fig. 7. Comparison between QoS and HireSome-I

V. CONCLUSION

The potential of big data is to convert the way healthcare suppliers use refined skills to achieve understanding from their clinical and other data sources and make respectable judgments. In the upcoming process a quick and extensive implementation of big data analytics in health care business is seen. Big data analytics assures security and privacy. The applications of big data analytics are still at promising stage of growth and its application in the health care industry will confidently help its organizations.

The scalability difficulty of large-scale data is examined and obtainable an extremely scalable two-phase TDS method using incremental map reduce clusters. Datasets are subdivided in parallel in the first phase, making transitional results. Data Anonymization is a means of moderating the privacy threat in e-health Accounts. Even if e-health accounts are in place at various hospitals they are not standardized and this data is typically unstructured, just like a log file, for with Map

Reduce is well-matched. Then, the intermediate outcomes are combined and further anonymized to yield reliable k-anonymous data sets in the second phase.

The composition estimation methodology attains two benefits. Firstly, the method considerably decreases the time difficulty as only some demonstrative history records are enrolled, which is highly required for big data applications. Secondly, the technique guards cloud privacy as a cloud is not mandatory to uncover all of its business records, which consequently safeguards privacy in big data. In forthcoming process, there is a plan to examine the scalability of privacy preservation in big data presentations with cloud service contact.

ACKNOWLEDGMENT

Deepest appreciation to all those who provided the possibility to complete the report. I would also like to thank the experts who were involved in the validation survey for the project.

REFERENCES

- [1] 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.
- [2] 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.
- [3] 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.
- [4] M. Li, S. Yu, Y. Zheng, K. Ren, and W. Lou, "Scalable and Secure Sharing of Personal Health Records in Cloud Computing Using Attribute-Based Encryption," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 1, pp. 131-143, Jan. 2013.
- [5] X. Zhang, C. Liu, S. Nepal, and J. Chen, "An Efficient QuasiIdentifier Index Based Approach for Privacy Preservation over Incremental Data Sets on Cloud," *J. Comput. Syst. Sci.*, vol. 79, no. 5, pp. 542-555, Aug. 2013.
- [6] W. Lin, W. Dou, X. Luo, and J. Chen, "A History Record-Based Service Optimization Method for QoS-Aware Service Composition," in *Proc. 2011 IEEE ICWS*, July 2011, pp. 666-673.
- [7] B. Benatallah, M. Dumas, Q.Z. Sheng, and A.H.H. Ngu, "Declarative Composition and Peer-To-Peer Provisioning of Dynamic Web Services," in *Proc. 18th Int'l Conf. Data Eng.*, 2002, pp. 297-308.
- [8] L. Zeng, B. Benatallah, A.H.H. Ngu, M. Dumas, J. Kalagnanam, and H. Chang, "QoS-Aware Middleware for Web Services Composition," *IEEE Trans. Softw. Eng.*, vol. 30, no. 5, pp. 311-327, May 2004.
- [9] L. Zeng, B. Benatallah, M. Dumas, J. Kalagnanam, and Q.Z. Sheng, "Quality Driven Web Services Composition," in *Proc. 12th Int'l Conf. WWW*, May 2003, pp. 411-421.
- [10] G. Canfora, M. Di Penta, R. Esposito, and M.L. Villani, "An Approach for QoS-Aware Service Composition Based on Genetic Algorithms," in *Proc. 2005 Conf. GECCO*, June 2005, pp. 1069-1075.
- [11] M. Alrifai, D. Skoutas, and T. Risse, "Selecting Skyline Services for QoS-Based Web Service Composition," in *Proc. 19th Int'l Conf. WWW*, Apr. 2010, pp. 11-20.
- [12] Cong Wang, Kui Ren, Qian Wang and Wenjing Lou "Towards Secure and Dependable Storage Services in Cloud Computing," *IEEE*

- transactions on Services Computing, vol 5, no. 3, pp 220-232, 2011.
- [13] C.Wang, Qian Wang, Kui Ren, Wenjing Lou, "Ensuring Data Storage Security in Cloud Computing," Proc. IWQoS '09, pp. 1-9, July 2009.
- [14] Hsiao-Ying Lin and Wen-Guey Tzeng "A Secure Erasure Code-Based Cloud Storage System with Secure Data Forwarding," IEEE Trans. Parallel and Distributed Systems, vol 23, no. 6, pp. 995-1003, June 2012.
- [15] L. Wang, J. Zhan, W. Shi and Y. Liang, "In Cloud, Can Scientific Communities Benefit from the Economies of Scale?," IEEE Trans. Parallel Distrib. Syst., vol. 23, no. 2, pp.296-303, 2012.
- [16] H. Takabi, J.B.D. Joshi and G. Ahn, "Security and Privacy Challenges in Cloud Computing Environments," IEEE Security and Privacy, vol. 8, no. 6, pp. 24-31, 2010.
- [17] Guo.Z.H, Fo.G, Zhou.M, Ruan.Y, "Improving Resource utilization in MapReduce". In IEEE cluster'12. pp,402410,2012.
- [18] LeFevre.K, DeWitt.D.J, and Ramakrishnan.R(2008), "Workload-Aware Anonymization Techniques for LargeScale Data Sets," ACM Trans. Database Systems, vol. 33, no. 3, pp. 1-47, 2008.
- [19] White.T, "Hadoop: The Definitive Guide", 3rd version, O'Reilly Media,2012.
- [20] Zaharia.M, Konwinski.A, Joseph.A.D, Katz.R, Stoica.I, "Improving MapReduce Performance in heterogeneous environments". In OSDI'08, pp.29-42, 2008.
- [21] 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.
- [22] R. Jurca, B. Faltings, and W. Binder, "Reliable QoS monitoring based on client feedback," in Proc. 17th Int'l WWW, May 2007, pp. 1003-1012.
- [23] K. Kritikos and D. Plexousakis, "Mixed-Integer Programming for QoS-Based Web Service Matchmaking," IEEE Trans. Services Comput., vol. 2, no. 2, pp. 122-139, Apr./June 2009.
- [24] Junwei Cao, M. Li, S. Yu, and W. Lou, "Optimal Multiserver Configuration in Cloud Computing," in ICDCS 11, Jun. 2012.
- [25] A. Klein, F. Ishikawa, and S. Honiden, "Towards NetworkAware Service Composition in the Cloud," in Proc. 21st Int'l Conf. WWW, Apr. 2012, pp. 959-968.



