ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 8, August 2015

Data Security and Updation of Data Lifecycle in Cloud Computing using Key-Exchange Algorithm

Sangita Rautela[1], Arvind Negi[2], Prashant Chaudhary[3]

M.Tech, CSE, Uttaranchal University, Dehradun(U.K.), India [1,2]

CSE Department, Uttaranchal University, Dehradun(U.K.), India [3]

Abstract: Cloud Computing is an emerging paradigm which has become today's hottest research area due to its ability to reduce the costs associated with computing and provides various types of facilities to the users. In today's era, it is most interesting and enticing technology which is offering the services to its users on demand over the internet and based on based on the concept of storage, virtualization, connectivity and processing power to store resources. Since Cloud Computing stores the data and disseminated resources in the open environment, security has become the main obstacle which is hampering the deployment of Cloud environments. Even though the Cloud Computing is promising and efficient, there are many challenges for data security as there is no vicinity of the data for the Cloud user. To ensure the security of data, in this paper authors proposed a modern data lifecycle model using implementation of DIFFIE-HELLMEN algorithm. It allows two parties to communicate with each other also exchange their secret keys over an unsecure transmission channel without knowing to each other.

Keywords: Encryption, Decryption, Data lifecycle, Key-exchange.

I. Introduction

Cloud computing is a mannequin of know-how processing, storage, and supply in which bodily resources are supplied to purchasers on demand. It can be defined as "administration of resources, applications and knowledge as services over the internet on demand". Cloud computing builds on founded traits for riding the price out of the supply of services at the same time increasing the pace and agility with which services are deployed. In cloud computing, many carriers offer unique cloud services with extraordinary pricing units like Amazon, Google and Microsoft. Carriers have their possess insurance policies and contract. Cloud computing presents the power to store their knowledge over the cloud. Then it is the responsibility of the cloud provider supplier to hinder the info from unauthorized entry. To comfortable the information quite a lot of forms of security mechanism are used.

NIST definition of cloud computing Cloud computing is a model for enabling convenient, on demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

The cloud is described as a three-tier structure, namely Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS) from low-layer to high-layer. Recently PaaS is evolving rapidly as a software development/ deployment environment for enterprise systems. For example, Microsoft Windows Azure provides .NET development environment and SQL service, whereas Google App Engine provides Python and Java development environment with key-value store service. A developer chooses an appropriate service among available ones to build his customized system.

Cloud computing is a form of information technology which is being used where lesser investment in efficient software is needed. Cloud computing consists of Access to applications and services is enabled over the network and it require only access to internet connection. Possibly one can get access of the cloud with the use of an ordinary client simply anywhere and anytime and one needs a certain information facility, without any special software. Cloud computing also facilitates the clients for immediate access to pre-set common but valuable information resources (as access to the network, hardware, storage capacities, software, and special information services) that are eagerly available without a wide agreement making process.

A. Data implementation model

i. SEARCHING or COLLECTION

Starting the information lifecycle and data move, we appear at the start of information: knowledge assortment. Moves taken early within the knowledge lifecycle can eventually pay off. For example, administrative principles can also be situated that limit the gathering of useless data, or information regarded too gigantic a chance to acquire.

Don't forget prohibiting the gathering of Social protection numbers, protected wellness knowledge (as defined in

DOI 10.17148/IJARCCE.2015.4883

HIPAA), complete credit card numbers and different sensitive information unless indispensable for the efficiency of trade. Additionally, if you happen to without doubt need this knowledge, then be certain to encrypt, or at the least appear at truncation practices.

ii. IMPORTANCE or PURPOSE

One more consideration in the information glide is whether the info is main to the business process. Administrative controls must discourage irrelevant knowledge collection because it might be "first-rate to have" or priceless for future initiatives.

Eventually, you ought to feel concerning the consequences to you as CISO and the industry if this information has been ever misplaced or breached. Will you be competent to explain why this knowledge used to be gathered in the first position?

iii. AUTHENTICATION

One of the major issue of public-key encryption is to address the problem of key distribution. The motivation for running the Diffie-Hellman protocol is to implement a secure session over an insecure network connection. Authentication is verifying that the person who requested an access to the information is who he claims to be. It is a process of proving identity. Authentication is a major security measure for cloud computing service providers and users. It is important for service providers to insure that the technologies of authentication are accurate. The main techniques used in authentication are username and password, key-exchange, tokens, biometrics, certificates, and Kerberos. Username and password is the most common user technique. Here authors uses Diffie-Hellman key-exchange algorithm for authentication of users and cloud networks.

iv. SHARING or CLASSIFICATION

A key foundation to this method is information classification. In other words, in case you have information in play, how do you know what controls follow?

The info classification system and development of this discipline is well researched and discussed in many safety boards. Nonetheless, the respect to position into play is simplicity and ease of implementation. Key gamers must comprise the info owner, information custodian, authorized division and the CISO. It's primary to remember a less difficult procedure where there are most effective a constrained quantity of classifications, including:

- Industry sensitive or personal
- Individual identifiable understanding (PII)--some state data breach legal guidelines could also be important in defining this class and typically incorporate identify plus Social safety quantity, driver's license number or credit card/account quantity
- Protected wellbeing know-how (PHI) for HIPAA security
- Unrestricted or public knowledge.

Knowledge owners will have to classify files centered upon company steering. The information custodian ensures most effective right members have access to view and manipulate information headquartered on role and classification. Legal will screen this knowledge for data retention and compliance activities comparable to e-discovery. The CISO, meanwhile, will use this classification to oversee appropriate storage, dealing with and unlock.

The CISO will have to additionally work with legal to arrange a marking standard, which states how a report will have to be marked and the way classifications may also be changed if integral.

With each and every classification, you will have to set up specific dealing with, storage and disposal specifications that weave protection into the information lifecycle.

v. PRESERVING or STORAGE

As data is moved through the lifecycle, it will be stored in databases, processed and handled as required for the trade. This step is meant to make certain that touchy and protection information is thoroughly saved, dealt with and now not given or released to unauthorized members or organizations. Coverage exists to make sure individuals don't digitally or bodily manage or release touchy data until approved. Some ideas to keep in mind listed here are:

- Encryption of right knowledge in transit and at leisure
- Hashing data to be guaranteed of data integrity
- Entry controls to ensure best licensed contributors get to touch, view and manipulate information
- Activation and monitoring of audit logs.

vi. COMMUNICATION or TRANSMISSION

This aspect of the lifecycle entails digital transmission of data as well as bodily.

For illustration, considerations for data security might comprise SSL or Transport Layer security (TLS) tunneling, encryption of e-mail and attachments, and e mail content filtering or blocking off.

Physical transportation screw ups can be minimized with the aid of encrypting all media in transit (i.e., backup tape encryption), monitoring the media as it strikes from point to factor, and receipt management so the organization is assured the data is bought when and where expected. Most state knowledge breach notification legal guidelines also relieve the corporation of the notification mandate if the lost or misplaced know-how is encrypted.

A key consideration right here is to also be certain that contractual controls with the physical transportation enterprise are in play, including indemnification of the enterprise will have to the courier lose the data in transit. Although indemnification is not necessarily a compliance problem, it most likely displays an group's due care and attitude toward its fiduciary responsibilities to defend the enterprise.

vii. ARCHIVAL or CONVERSION

This is generally by using a long way the most important chance subject of the info lifecycle. Right here controls are elaborate to establish to avoid customers from

DOI 10.17148/IJARCCE.2015.4883

copying information, making screenshots of data in method, pasting data into individual spreadsheets and databases, and so on.

For instance "information personalization," aka "personal information assortment initiatives," substantially increases the chance profile for the group. As an example, an employee could also be collecting knowledge from various enterprise databases and screenshots for individual use, equivalent to his or her possess telephone record or roster, or for future initiatives. Right here the info lifecycle is significantly disrupted and sensitive information can finally end up on customers' workstations, USB drives, and even at their homes, regardless if the intent is optimistic and for the good of the corporation.

Controls to bear in mind in this area are technical controls to restrict information flowing outside to the corporation except it is encrypted, such as data leak prevention technology; electronic mail content material evaluation and administration; and a draconian notion of prohibition of in my view owned transportable media.

Administrative controls could also incorporate policies and strategies on how information must and must no longer be used or amassed by means of individual workers. Principles forbidding use of manufacturer computers for worker off-hours movements are also indispensable. Of path, center of attention on step one--data collection--would aid cut back this risk, too.

viii. LEAVE or RELEASE

In 2003, the reputation of Jet Blue was seriously damaged because data held by the airline was improperly released to a TSA contractor. This event demonstrates that how and when sensitive data is released in the lifecycle must be closely controlled. Aspects of this lifecycle element should include:

- Definition as to who has authority to release data outside the organization
- Recognition that not all data can be released upon a simple request
- Data is not released unless appropriate for the business and only if legal and with appropriate controls.

Other subtle elements of data release also need to be examined. For instance, a company may use a vendor to analyze data as part of a contract. What controls are in play relative to data sharing, data control and data breach by the vendor? To reduce the enterprise's risk profile, strong contractual controls need to be established to indemnify it if the vendor loses the data or uses it for inappropriate purposes (e.g., using your data for vendor marketing campaigns).

ix. BACKUP

This area tends to be a fairly mature domain for security. However, there continue to be breaches where unencrypted backup tapes are lost in transit. Look at the data lifecycle to ensure this process contributes to the security and availability of the data.

x. RETENTION or DATA HOLDING

Data being held with the aid of the enterprise is field to discovery for legal process. Litigation holds are becoming increasingly crucial on the grounds that of more focal point on e-discovery because of the new Federal rules of Civil method issued in December 2006.

The information lifecycle wishes to be certain that information is simply and properly retained so that data will also be with no trouble located and held for these discovery standards. Nonetheless, you also need to be certain that data is destroyed at the proper time to ensure that surprises are minimized in the course of the discovery approach when knowledge proposal to be "dead" or gone surfaces.

xi. EVALUATION or REMOVAL

The top of the process the "demise" of information is the information destruction approach.

That is without a doubt another field that may be fraught with issues for the CISO if no longer carried out utterly and with the right controls. If information is to be eliminated, then it must be fully destroyed and not left in any post-destruction residual. You do not need to listen to about your surplus apparatus being stuffed with sensitive data now within the open market. Some key practices to bear in mind:

- Destroy hard drives by physical destruction or shredding--there's too much risk with the incomplete "wipe." Costs run about 25 cents a pound and can be easily witnessed. Of course, you can also use disk-wiping tools, but the diligence required to assure disks are properly wiped and processed may actually cost more than the total cost and assurance of physical destruction.
- Destroy paper by shredding. The process should be periodically witnessed, and the enterprise needs contractual assurance that it is indemnified should the vendor fail to complete the process.
- Do not destroy information on litigation hold--data on hold for legal review or subpoena--or too early in the retention schedule.
- Make sure employees know how to handle waste that is classified as confidential, business sensitive, etc., so such documents do not wind up in a public landfill and are instead shredded or destroyed.

Now, your data may be dead, but that doesn't mean the lifecycle has ended. This new approach to looking at how data lives and dies begs for additional analysis.

One thought experiment is to map the "risk value" of each lifecycle stage. No empirical evidence necessarily supports this mapping, but it can be used to show the relative risks encountered when you look at data lifecycle security in the enterprise.

The biggest risk is the data manipulation, conversion or alteration stage. Since it is so easy for an individual to copy and collect data for other uses, data gets distributed throughout the enterprise and cannot be easily controlled. And the risk can be significant if the data is moved offsite,

DOI 10.17148/IJARCCE.2015.4883

to a home computer, placed on an unencrypted USB drive, etc.

The flow of data is a new way to guide security professionals' focus, time and energy. They can look at new ways to not only protect the data, but also use this as a way to communicate risks and issues to executive management.

Also, this data lifecycle security approach can be a new way to build a security program, procedures and strategy. And it may be a new way to justify expense in critical areas of the organization, including security, legal and operations.

II. Literature Survey

While coming with this paper we have referred the technical paper on Secure Data Access over Cloud Computing and Secure Data Access in Cloud Computing. We also had visited many previous research papers, survey papers, blogs and websites those are related to the data security in cloud computing.

[1] Mahmood identifies that the major issues pertaining to data security in the cloud computing environment are:

□ *Data Location and Data Transmission* — the customers may want that data should reside on a specific territory based on data polices and legislations within the certain country. Similarly, cross border transition of data (from one country to another) may lead to potential risks due to varying policies, regulations and legislations.

□ *Data Availability* — the unavailability of data may lead to service outages.

□ *Data Security* — when the data mobility is at high level, then security risks become the major concern, particularly, when data is transferred to another country with a different regulatory framework.

[2] A recent survey by Cloud Security Alliance (CSA)&IEEE indicates that enterprises across sectors are eager to adopt cloud computing but that security are needed both to accelerate cloud adoption on a wide scale and to respond to regulatory drivers. It also details that cloud computing is shaping the future of IT but the absence of a compliance environment is having dramatic impact on cloud computing growth.

Recently the cloud computing technology became widely used by most of the business companies to increase their productivity and with that there are still some concerns about the security provided by cloud computing. In 2013, cloud computing is still in high demand where the organizations are either already using or intending to use cloud computing infrastructure services, and the share of cloud service will continue to increase as a percentage of total revenue. One of the biggest concerns with cloud data storage is the verification of data integrity at entrusted servers, and how to deal with sensitive data. It is not an easy task to maintain customer’s most sensitive cloud data securely, which is needed in many applications for clients. This makes some companies wary of switching to cloud computing because the user does not know on which server the data is stored and is this server provides secure data or not.

Data life cycle process is used to secure data which is transferred to third party. In this life cycle protection data confidentiality, integrity, availability, access control process, key management and encryption process is done in cloud service provider and client. The life cycle consists of create data, store data, share data, destruct data in cloud. That is between user, data owner and cloud service provider. In this life cycle the encrypted data is used by frequent key changing process which makes complex and costly for data owner.

The process of encrypting the data using D-H key algorithm process in data access makes the data security more efficient. To secure data access transferred to cloud service provider by data owner cryptographic technique is applied. RSA algorithm is used for encrypting key and to make the key more secure digital signature is used with RSA algorithm. The process of adding digital signature makes data more secure. Data transferred by encrypting key with digital signature from data owner to cloud service provider and data owner makes the data available to user after decrypting the key with digital signature.

III.Related Work

Similar to an economic value-add analysis methodology, the data lifecycle security model (Figure below) shows how data is collected, classified, authenticated, stored, used, retained and ultimately destroyed. It shows process, transition and a business flow.

The figure shows the data lifecycle. Data first goes through the SEARCHING or COLLECTION phase, where the risk of losing data or data being manipulated is moderate. Next it goes to the CREATION phase here risk is low, next to the AUTHENTICATION phase, a new modern step and specially using for enhance the security of cloud computing using key exchange algorithm, the risk is high in this phase.

Then data IMPORTANCE or PURPOSE afterfiltering of an authorized user ,the risk here is moderate. The next phase is CLASSIFICATION or SHARING, where there is a low risk of data loss, the next phase is PRESERVING or STORING means handling of data, where there is a high risk of data loss or unprivileged access. Then data communicated or transmitted, the risk here is moderate.

The next phase is the ARCHIVAL or CONVERSION means manipulation of data, this phase has a high risk of losing data. The data then goes through LEAVE or RELEASE, BACKUP, and RETENTION or DATA HOLDING and then EVALUATE or REMOVAL phases, where the risk is moderate.

We can conclude that the phases where the data is threatened the most are the authentication, preserving and storage, and the archival, or manipulation of data.Authentication phase is a proposed or updated field in data lifecycle for cloud computing.

DOI 10.17148/IJARCCE.2015.4883

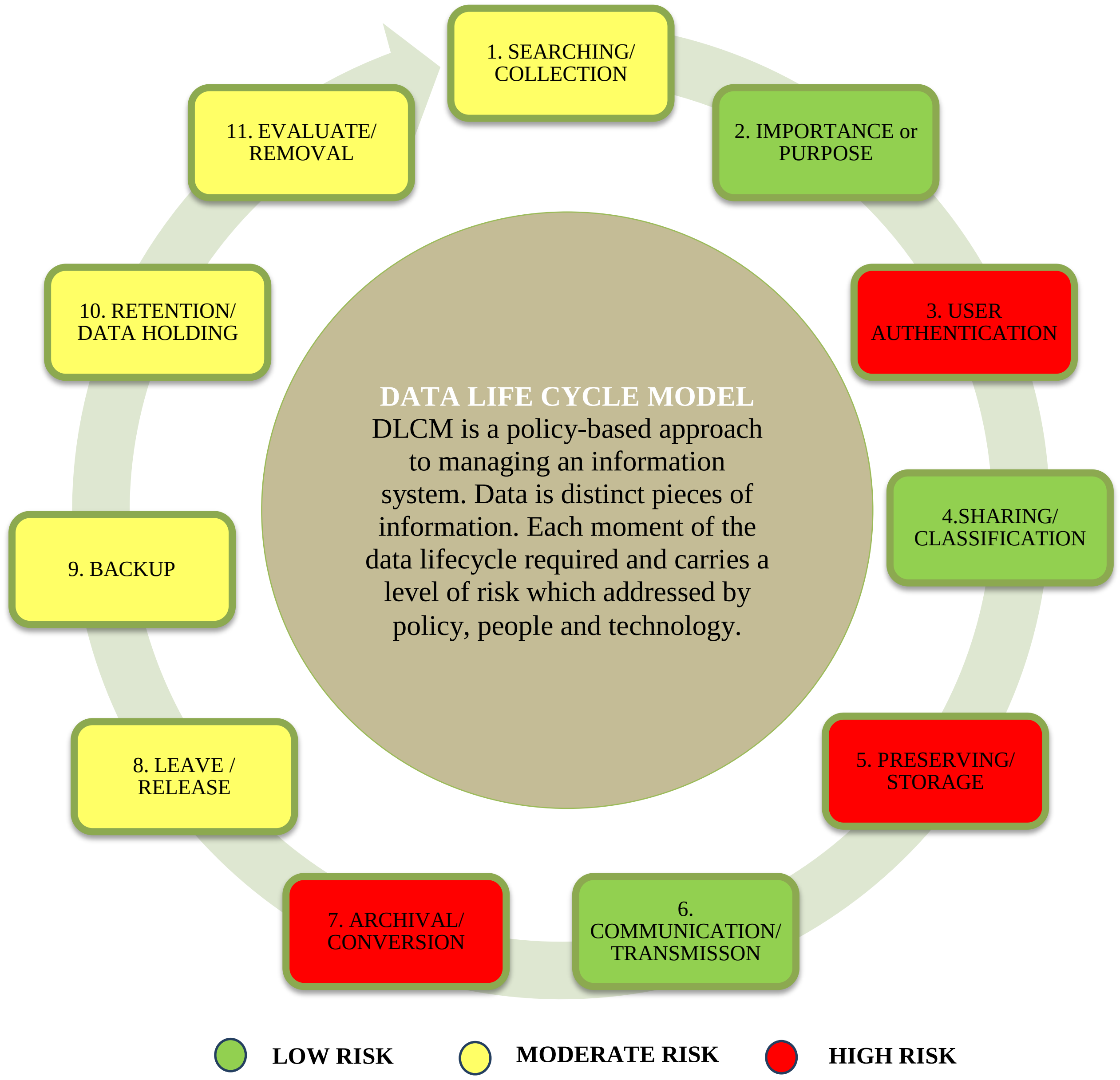
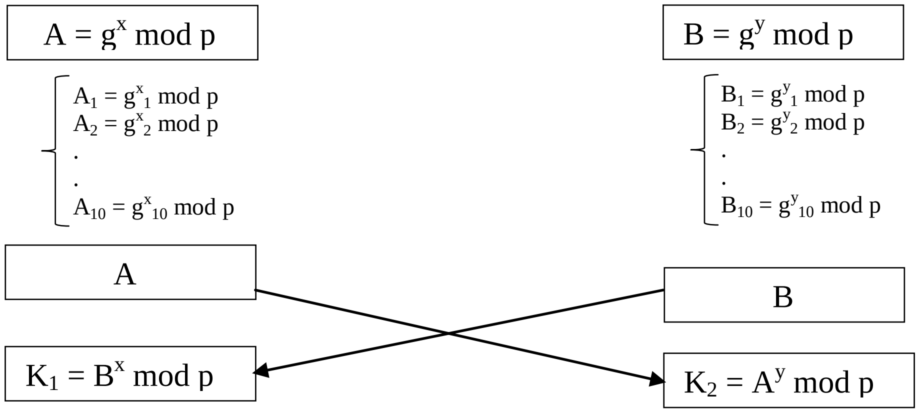

1. SEARCHING/ COLLECTION
2. IMPORTANCE or PURPOSE
3. USER AUTHENTICATION
4.SHARING/ CLASSIFICATION
5. PRESERVING/ STORAGE
6. COMMUNICATION/ TRANSMISSON
7. ARCHIVAL/ CONVERSION
8. LEAVE / RELEASE
9. BACKUP
10. RETENTION/ DATA HOLDING
11. EVALUATE/ REMOVAL
DATA LIFE CYCLE MODEL
DLCM is a policy-based approach to managing an information system. Data is distinct pieces of information. Each moment of the data lifecycle required and carries a level of risk which addressed by policy, people and technology.
🟩 LOW RISK
🟨 MODERATE RISK
🟥 HIGH RISK

Fig 1: Data Lifecycle Model



$A = g^x \bmod p$

$A_1 = g^x{}_1 \bmod p$
$A_2 = g^x{}_2 \bmod p$
.
.
$A_{10} = g^x{}_{10} \bmod p$

$B = g^y \bmod p$

$B_1 = g^y{}_1 \bmod p$
$B_2 = g^y{}_2 \bmod p$
.
.
$B_{10} = g^y{}_{10} \bmod p$

A

B

$K_1 = B^x \bmod p$

$K_2 = A^y \bmod p$

Diffie-Hellman key exchange algorithm gets it security from the diffculty of calculating discrete logorithms in a finite field, as compared with the ease of calculating exponentiation in the same field.

DOI 10.17148/IJARCCE.2015.4883

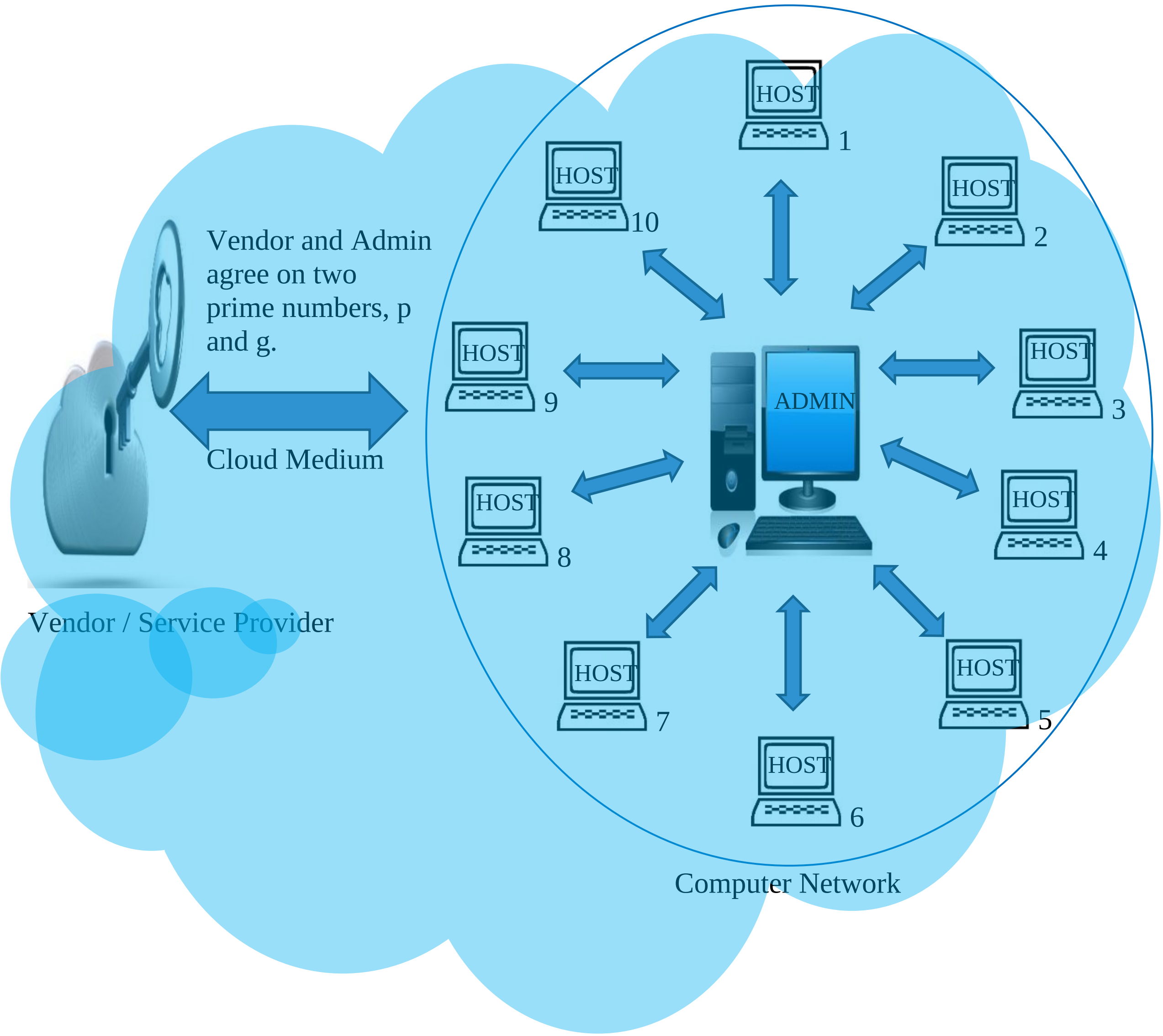


Fig 2: Data Communication and Key Sharing in Cloud Computing

IV. PROPOSED ALGORITHM

Here we are using unique integers for a particular network sub-system (HOST 1 to HOST 10), which chooses random integers, calculate A and sends to Vendor. Vendor also chooses random integers for particluar network sub-systems, calculate B and sends to Admin and agree on two large prime number.

1. Admin and Vendor agree on two large prime numbers, p and g.
2. Admin chooses another large random integer x, and calculate A

$A = g^x \bmod p$

Network sub-systems also chooses random integer numbers (x_1,x_2….. x_{10}), and calculate A.

$$\left.\begin{array}{l} A_1 = g^{x1} \bmod p \\ A_2 = g^{x2} \bmod p \\ . \\ . \\ . \\ A_{10} = g^{x10} \bmod p \end{array}\right\}$$

3. Admin sends the number A to Vendor.
4. Vendor chooses another large random integer y, and calculate B

$B = g^y \bmod p$

Now Vendor also contains random integers (y_1, y2….. y_{10}) for network sub-system, and calculate B.

$$\left.\begin{array}{l} B_1 = g^{y1} \bmod p \\ B_2 = g^{y2} \bmod p \\ . \\ . \\ . \\ . \\ B_{10} = g^{y10} \bmod p \end{array}\right\}$$

5. Vendor sends sends the number B to Admin.
6. A computes the secraete key K_1

$K_1 = B^x \bmod p$

7. B computes the secret key K_2

$K_2 = A^y \bmod p$

DOI 10.17148/IJARCCE.2015.4883

V. Conclusion

This paper presented an independent study of Data Lifecycle in cloud computing. This provides the specific method to secure data on cloud computing with the help of key exchange algorithm. Thus, in our proposed work, only the authorized user can access the data by following the proposed Data Lifecycle Model. Even if some unauthorized user gets the data intentionally if he captures the data also, he cannot decrypt it and get back the original data from it. Hence forth, data security is provided by implementing Diffie-Hellman algorithm.

In this paper we are not using encryption/decryption techniques or algorithms, here are some security issues. The researchers are try to mitigate these issues by building efficient symmetric (AES, DES) or asymmetric (DSA, RSA, ElGamal) algorithms in future.

References

[1] "The NIST Definition of Cloud Computing". National Institute of Science and Technology. Retrieved 24 July 2011.Mudili Soujanya, Sarun Kumar,*Personalized IVR system in Contact Center, Department of Computer Science Engineering* International Institute of Information Technology Bhubaneswar, India.

[2] What is cloud computing.Retrived April 6, 2011, available at: http://www.microsoft.com/business/engb/solutions/Pages/Cloud.aspx.

[3] Chevassut, Olivier, “Authenticated group Diffie-Hellman key exchange: theory and practice” Lawrence Berkeley National Laboratory, 10-03-2002.

[4] Yogita Gunjal, Prof. J.Rethna Virjil Jeny, “Data Security and Integrity of Cloud Storage in Cloud Computing”, in the year of April 2013.

[5] S. K. Sood, “A combined approach to ensure data security in cloud computing”, Journal of Network and Computer Applications, vol. 35, no. 6, (2012), pp. 1831-1838.

[6] T. Dillon, C. Wu and E. Chang, “Cloud computing: issues and challenges, *24th IEEE International Conference on Advanced Information Networking and Applications*, AINA, pp. 27-33, Apr. 2010.

[7] M. A. Vouk, “Cloud computing–issues, research and implementations,” *Journal of Computing and Information Technology*, CIT, vol. 16, no. 4, pp. 235-246, 2008.

[8] Shucheng Yu, Cong Wang. (March 2010) 'Achieving secure Scalabe and Fine grained data access control in Cloud Computing ', IEEE Conference INFOCOM , 1-9.

[9] Damgrd, Ivan, et al. "Secure key management in the cloud." Cryptography and Coding. Springer Berlin Heidelberg, 2013. 270-289.

[10] Kalyani M. ."Cloud Security: E_cient and Reliable Encryption Key Management Crucial for Data Protection". https://spideroak.com/privacypost/cloud-security/secure-encryption-key-management-in-the-cloud/

DOI 10.17148/IJARCCE.2015.4883