JCSE International Journal of Computer Sciences and Engineering Open Access

Survey Paper Vol.-7, Issue-1, Jan 2019 E-ISSN: 2347-2693

Data Security Framework for Data-Centers

Narender Kumar[1*], Avinash Karhana[2]

[1,2]Department of Computer Science and Engineering, H.N.B. Garhwal University, Srinagar Garwal, Uttrakhand,India

[*]*Corresponding Author:* narearawal@gmail*.com,*

Available online at: www.ijcseonline.org

Accepted: 22/Jan/2019, Published: 31/Jan/2019

***Abstract*—**Data-centers play an important role in today's time for storing and maintaining huge amounts of data being generated daily and operated in real-time, to keep this data secure becomes as much important as keeping it stored. To address the current flaws in security policy of data-centers and to eradicate those flaws this paper suggest a new security framework for data-centers. This would provide data-centers with all kind of precautions and instructions to handle most of situations that may arise during management and other data related tasks and even after breaches by defining all data storage rules by classifying data, and defining data access rights with security policies for authentication and data verification. These all things make sure that data is secured and more reliable at both aspect storage as well as data authenticity in this framework for Data-Centers.

Keywords—Data-center, Encryption, Data security, Immutable data, Policy, Real-time backup

I. Introduction

In earlier times data was considered to be an office term as only used in business houses or government offices. And all that was kept and maintained on paper records but from 70s came digital storage media that started to convert these manual paper records to digital documents [1]. But still this data was mainly limited to business houses and offices, till the personal computers came into market in mid 90s that gave blast to data generation as at this time data was not generated only by offices and businesses but also by common people [2]. In addition to this internet just boosted this data generation rate graph exponentially, now data was generated by people as well as other devices as sensors and this huge amount of data brought the term “Big Data”. To handle this huge amount of data the need to store, maintain and provide security to this data may have confidential information that has to be kept hidden and secure in all means. To provide security to this, data security frameworks have to be implemented and security polices are to be applied and followed to maintain high security standards in data-centers this paper proposes a new security framework to secure data center from all kind of threats and breaches. This framework also provides ways to handle situations even after breaches. The framework basically covers all important parts from defining Data-Center, addressing all kind of risks and providing relevant solutions to them..

This paper is organized as follows: definition of data-center is given in in Section 2. A detailed description of security threats and their cause as well as tending solution to them are given in Section 3. Subsequently, the security policies and proposed framework is explained in detail in Section 4. In Section 5. We put forward cryptography as a security enhancement block in framework. This leads to Hashing and Hash Functions that is used for maintaining data accountability and authenticity and this is discussed in Section 6. Finally, conclusions are drawn in Section 7.

II. Data-Center

A data center is a building or dedicated part of a building used to hold computer systems and associated components, such as storage and networking systems[3]. In general, it consists of redundant or backup elements and infrastructure as power supply, networking connections, environment management components (e.g. temperature conditioning, fire control) and various security devices [4]. A large data center is an industrial-scale operations that may be using as much electricity as a small town[5]. In smaller organizations a data center can be a single computer or a hard drive. Wherever in larger organizations data center can be referred to network control center, server room or a server hub. On other hand very large scale organizations may consider internet connected datacenter needs and can have cloud based data center consisting storage and networking structures [6].

III. Data-Center Security and its Issues

Data center security can be defined as the set of policies, precautions and practices that are to be followed to prevent manipulation of a data center’s assets and to avoid unauthorized access to the data. The data center stores the enterprise data that is why providing that data proper security is so crucial. At present denial of service (DoS), confidential information breach, data mutation or modification and loss of data are some of the common security problems that associates with corporate data center environments [7]. Oracles corporation have explain security risks and solutions[8].

Table 1: Security Risks and Solutions

Problem	**Solution**	**Security Technology**
Unauthorized users	Know your users	Authentication
Unauthorized access to data	Limit access to data	Access control
	Encrypt data	Data Encryption
	Limit privileges	Privilege management
Eavesdropping on communications	Protect the network	Network encryption
Corruption of data	Protect the network	Data Integrity
Denial of service	Control access to resources	Availability
Complexity to user	Limit number of passwords	Single sign-on
Complexity to administrator	Centralize management	Enterprise user security
Overly broad access to data	Dynamic query modification	Fine-grained access control
Too many accounts	Centralize management	Directory services
Operating system break-in	Encrypt sensitive data	Stored data encryption

IV. SECURITY POLICY AND PROPOSED FRAMEWORK

For data-centers, the security policy refers to set of rules on functions and their flow among them, constraints on remote access and adversaries including application/programs and data access by users (both authenticated or non-authenticated).

As per proposed framework all systems should be classified into three or more categories as per requirement, as per basic taking three categories as:

- Servers
- Host to server system
- Operator Systems

Servers will be the main database systems that have all data.

Host to server systems will be the just next to server systems that connects to servers via network.

Operator Systems will be the lowest level systems that are to be used by data entry operators and data retrieving and information systems.

So all the systems that are present in data-center must be categorized in this classification and remote access policies and data access policies must be applied as per this classification.

After that comes security policies, we divide these policies insub categories as:

- Data classification policy
- Data Access rights policy
- Remote Access policy
- Data transfer policy
- Data backup policy
- Data wiping policy
- After incident policy

Data classification policy: In this policy we classify type of data that is going to be handled in Data-Center. All data whether at static (i.e., stored in databases, tables, email systems, flash drives, etc.) or in use (i.e., being: processed by software systems, transmitted over network, used in e-sheets, or manually manipulated, etc.) must be classified into one of the three data classification levels described in this policy by each unit.

- Level I – Confidential Information: Critical financial data, legal subjection, public distrust, or publicly harmful data.
- Level II – Sensitive Information: Mid level requirement for Confidentiality and/or medium level or low risk of financial loss, legal subjection, public distrust, or harm if this data is disclosed.

- Level III – Public Information: Low requirement for Concealing data [information is public] and/or low or negligible risk of financial loss, legal subjection, public distrust, or harm if this data is disclosed.[9]

*System and Data Access rights policy:*In this policy we define how and who has rights to access what data.

- Servers: Only DBA(Database Administrator) has access to this type of system, highly encrypted, secured with multiple type authentication like biometrics. Access rights available to this type of system is FULL_CONTROL (Includes database configuration and settings and all other type of access rights).
- Host to server systems: Only DBA(Database Administrator), Sub-ordinates has access to this type of system, highly encrypted, secured with multiple type authentication like biometrics. Access rights available to this type of system are DATA_MANIPULATION_AND_HANDLING. (Includes all type data modification access and lower level access rights).
- Operator: DBA(Database Administrator), Sub-ordinates, Operators has access to this type of system, Low encryption level, secured with basic authentication measures. Access rights available to this type of system is DATA_READ, INSENSITIVE_DATA_WRITE. [10]

Remote Access policy: In this policy we define which kind of system is remote accessible and who has the remote access rights to these.

- NO_REMOTE_ACCESS: Servers and have only physical access to this type of system only by DBA.(Database Administrator)
- MID_REMOTE_ACCESS: Host to server systems can have remote access by Subordinates and DBA(Database Administrator).
- FULL_REMOTE_ACCESS: Operator systems can be accessed remotely by DBA and Sub-ordinates with FULL_CONTROL access right.

Data transfer policy: In this policy we define how data is being transferred from one place to another either it is from database to backup databases or data sharing via secondary storage device.

As per security policy are concerned all data that is being transferred in any media must be in encrypted form.

All data transferred should be done after proper authentications and handshakes between sender and receiver to verify data is being sent to correct place by correct sender.

For ensuring security all data transfer must be through secure chanel and encryption.

Here for encryption we use hashing and in 2 key factor cryptography that consists of two keys to decrypt and encrypt Public Key (Transferred publicly by any means)

- Private key (Transferred through secure medium)

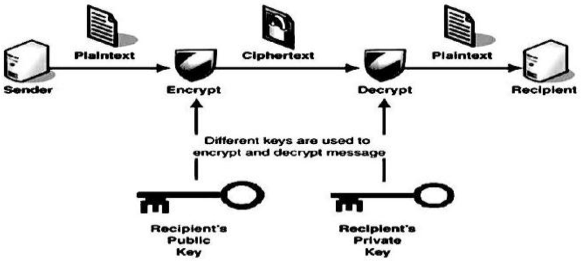

Fig. 1 Data Encryption

At the end of decryption these two keys are used in this data encryption and decryption. [11]

Data backup policy: As per this policy we decide how data is to backed up for ensuring protection of data during undesired situations that leads to data loss or other means of damage.

In this frame work there are two type of backups that holds up data:

- Short term Backup: This backup is a redundant copy of all changes made to database in past two days with what changes are made and timestamps of changes and other information like who changed data and from which IP address and from which interface etc. This backup is a real-timebackup (that means it stores all changes at every time a change is made) that has datalife time of 2 days and after two days of data backup data expires and free up the occupied space. This kind of backup will be used to handle after breach or after incident situations.
- Long Term Backup: This backup is also a redundant copy of all the data stored in database. The frequency of data backup for this system depends on number of database entries or updating done in a day, if number of entries or updating is very high it can be twice a day and if number of entries and updating is lesser than it can be once a day.

These databases are to be located at different physical locations for data safety from natural calamities or some kind of other threats.

All data stored in all databases (including backup and master) is immutable i.e. for each change a new entry is done with timestamps of new update and other info like what changed.

If one database crashes, then automatically server chooses another database from backup bases to act as a master database.

At time of backup from master to backup databases all data is transferred in form of hash to ensure authenticity and integrity of data being transferred/backed up. Forthis we use hash function [11].

All the data entries made in database are immutable data records i.e. once a data record is made it is not changeable even if it has to modified a new record has to be created with newer timestamp rather than changing previous record, this technique is called immutable Lazer technique [4].

For authenticity and integrity any data change is only possible in any backup database if other databases verify that change so it is a distributed authentication approach. If we have a look at the databases and backup timings and handling in graphical way, it would look something like in Fig.1 which indicates Server that decides master database (Indicated by ☐) from several databases present at different locations.

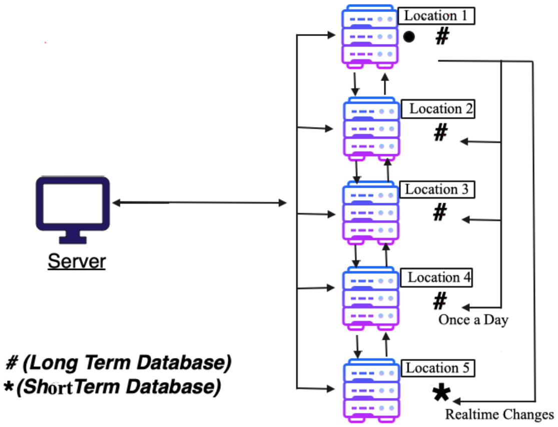

Fig.2 Databases and backup strategy

Data wiping policy: Inthis we define how data wiping and destruction is to be done to protect data leaks and handle other data threats. In this security framework we keep check on type of data that is to be wiped and type of hardwares used to store and also deal with drive and storage device health.

If data that is to be wiped falls under Level 1 Classification the suggested method to wipe will be to hard bit by bit wiping (replacing all data bits on drive with 0) that leaves no chance of data theft or leak after wiping. For data of Level 2 Classification the suggested method to wipe is data relocation and fragmented wiping with bit replacement mix this makes data recovery very hard i.e. nearly impossible. For any data that comes under Level 3 classification data wiping can be usual soft format that has a chance of data recover but that is not an issue at this data is not sensitive and can come handy if data may be needed later of deleted by mistake.

Dealing with storage devices is very crucial when working in data-centers as every storage device has some limited life span and they degrade over time affecting their efficiency and storage capabilities due to wear and tear during long operation time in different conditions, so here we classify all storage devices Health Status in three categories at the time of health checkup :

- Healthy (Performing up to mark)
- Replace (health is ok but not performing as per criteria)
- Discard (health is not ok)

When the status of any storage media is other than Healthy, either that drive is discarded or destroyed completely as per classification of data stored in it. When a drive is in Replace status that drive is wiped according to wiping policy and then reused with other low potential devices or systems. When a drive status is Discard that drive is to be destroyed completely.

After incident policy: In this policy we define how situations are to be handled after breaches like data theft, data manipulation, data loss etc. As per this security framework we take some precautions to prevent breaches and try to reduce or completely eliminate other security threats. The key points that are notable in this policy are:

- At first breach must be detected
- All breached data should be changed if possible
- If data is damaged or manipulated, then all the access rights must be revoked immediately for a time being in which data is secured again and security patches are done.
- The database that was breached must be reverted back to a state that was stable and secured, this will be done with help of other backup databases and short-term backup.
- A security audit must be done to find all vulnerabilities and creating patches for the vulnerabilities that were detected during audit.
- After patching and re-securing all databases and systems, breached should be identified and reported to be dealt according to legal procedures.

V. CRYPTOGRAPHY

Cryptography is the study of mathematical systems to solve two major security problems that are privacy and authentication. Where a privacy system prevents unauthorized extraction of information from a message being transferred via a public channel and an authentication system prevents the unauthorized injection of messages into a public channel, assuring the receiver of a message of the legitimacy of its sender [12].

So in cryptography we use different techniques to encrypt our data/information that is sent over any channel, but here all these techniques can be classified under two types: [13]

- Symmetric (Symmetric Key Cryptography)
- Asymmetric (Public Key Cryptography)

Symmetric: In this the one single key is used to convert the plain text into a cipher text and vice versa.

Asymmetric: In this the message sender and recipient do not share any common secrets and different keys are used to encrypt and decrypt message [14].

Encryption scheme is comprised of three algorithms: the first is for generating keys, second is the procedure for encrypting message, third is the procedure for decrypting. So at different stages these algorithms are used to encrypt and decrypt. And here for transferring any data over network from one database to other or system to database, that has to encrypted and to encrypt, this framework suggests using asymmetric cryptography. Here is a basic block diagram of how both type of cryptography works.

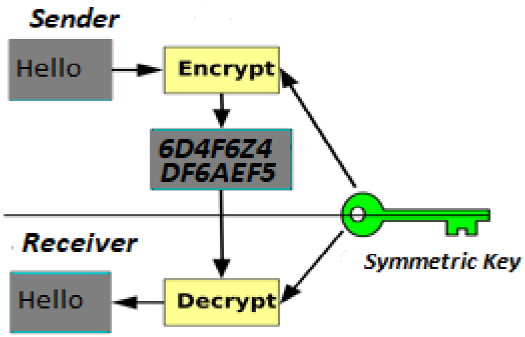

Fig.3 Symmetric Key cryptography

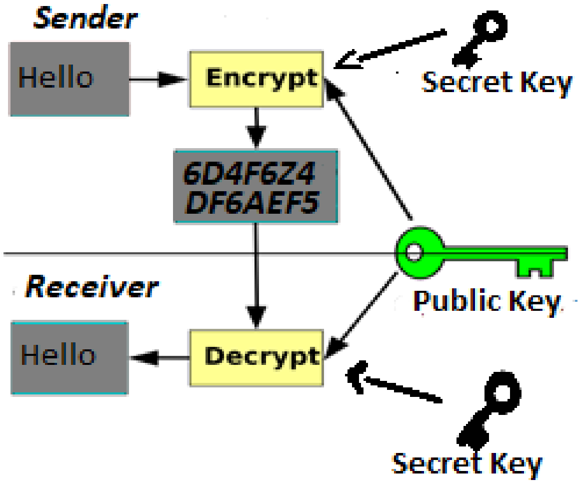

Fig.4 Public Key cryptography

VI. Hash Function

Hash function is a calculation applied to a key which may be a very large number or a very large string, to transform it into a relatively small number that corresponds to a position in a hash table, this index number is an effectively a memory address, for numeric keys in general, key K is divided by total number of available addresses N and take reminder to determine its position i.e. position=K mod N. And for alpha numeric keys, sum of all ASCII codes of in a key are divided by total number of available addresses N and take remind is taken to determine position [14] [15]. Here in this framework all the data that is being transferred uses hashing and hash function to ensure transfer authenticity and data security.

VII. Conclusion

Data-centers are hub of large data clusters and storages and securing these is a critical topic, to make these data-centers secure we have to develop such framework that have high potential policies that could reinforce data security and after considering all those factors, in this paper we propose a new security framework for data-centers that would ensure a secure environment to database. This paper suggests policies that covers key elements like Ensuring Data Accountability, Patch Management, Vulnerability Testing and System Audit, Account management and control and Response to incidents. This security framework covers all important elements that may effect security of a data-center and also suggest how data backup should be happen.

References

[1] Ben M. Chen, Tong H. Lee, Kemao Peng and VenkatakrishnanVenkataramanan, “*Hard Disk Drive Servo Systems*”, Springer, pp. 3-8, 2006

[2] Katherine Murray, “*Introduction to Personal Computers*”, Que Corporation,**India**, pp. 11-28, 1990.

[3] B.FakhimM.BehniaS.W.ArmfieldN.Srinarayana“*Cooling solutions in an operational data centre: A case study*” Elsevier Applied Thermal Engineering, Volume 31, Issues 14–15, p.p 2279-2291,2011.

[4] Al-Fares, M., Loukissas, A., & Vahdat, A. “*A scalable, commodity data center network architecture”* ACM SIGCOMM Computer Communication Review, 38(4),p.p 63-74,2008.

[5] S. Mittal, “*Power management techniques for data centers: A survey*,” CoRR, vol. abs/1404.6681, 2014.

[6] Knapp, K. J., Denney, G. D., & Barner, M. E. "*Key issues in data center security: An investigation of government audit reports*". Government Information Quarterly, 28(4), p.p 533–541,2011.

[7] Maurizio Portolani, Mauricio Arregoces. "*Data Center Fundamentals*". Publishers, Cisco Press, 800 East 96th Street Indianapolis, IN 46240 USA,2004.

[8] *“A Matrix of Security Risks and Solutions*"Oracle9i Security Overview, Release 2 (9.2)Oracle Docs.",2002.

[9] The University of Kansas Policy Library."*Data Classification and Handling Policy*", 2014.

[10] "*Data Center Access Policy*" Information Technology Services, West Virginia University, 2018.

[11] Laurens Van Houtven, "*Crypto 101*" Creative Comons,2013.

[12] Diffie, W., & Hellman, M. "*New directions in cryptography*". IEEE Transactions on Information Theory, 22(6),p.p 644–654,1976.

[13] Jonathan Katz Yehuda Lindell - "*Introduction to Modern Cryptography_ Principles and Protocols*",Chapman And Hall/CRC, p.p 241-245, 2007.

[14] Tadayoshi Kohno, Niels Ferguson, Bruce Schneier -"*Cryptography Engineering Design Principles and practical applications*" Wiley Publishing Inc.,2010.

[15] Goldreich O - "*Foundations of cryptography*", Now Publishers Inc. 2001.

Authors Profile

Narender Kumar pursed Ph.D. from Indian Institute of Technology, Roorkee. He is currently working as Assistant Professor in the Department of Computer Science and Engineering , H.N.B Garhwal University,Srinagar Garhwal since 2012. His main research work focuses on Cryptography Algorithms, Blockchain Technology, Data Mining,Machine Learning ,Genetic Algorithm, PSO and Digital Image Processing. He has 9 years of teaching experience and 6 years of Research Experience.

Avinash Karhana is currently pursuing Bechlor of Technology(CSE) in Computer Science and Engineering Department, H.N.B Garhwal University, Srinagar Garhwal. He has research interest in Data Security, Network Secuty , Digital Image Processing.