ISSN No. 0976-5697

Volume 8, No. 5, May-June 2017

International Journal of Advanced Research in Computer Science

REVIEW ARTICLE

Available Online at www.ijarcs.info

Data at rest and it's security solutions-A survey

Gazala Matloob
Dept. of computer science
Jamia Hamdard University
New Delhi, India

Dr. Farheen Siddiqui
Assistant Professor, Dept. of computer science
Jamia Hamdard University
New Delhi, India

Abstract: Security of the data is the major challenge for every organization and to secure the archived data or you can say the data at rest is critical because this may consist of some of the confidential information which can be used in future. So , in this paper I tried to discuss the data at rest with it's possible security solutions.

Keywords: Data at rest , data in transit , security, encryption , biometric, content address storage

INTRODUCTION

Data is defined as any raw facts and figures, it can be of any form the data present in the hard drives, pend drives, etc. Previously work which is done manually are all processed by the computer even a simple admission form can easily be submitted by just pressing a button due to which this gradually gave rise to data. Even this is explained by Gordon Moore in 1965 co- founder of Intel by given the Moore's law which states that the number of transistors on integrated circuit doubles approximately every two years and this will continue in future. In upcoming years the speed may get slow down but data density will get doubled on every 18 months. The research study is been conducted by the digital universe in united states according to which the digital data is expected to grow from 898 Exabyte's to 6.6 zeta bytes between 2012 and 2020 or more than 25% a year, which means it will double about every three years (see fig:1). Now the question arises what is the cause of the growth of this data.[1]

The reasons are:

i. Increase in the use of internet, smartphones etc.
ii. Falling cost of devices that create, capture and store information.
iii. Migration from analog to digital TV.
iv. Rise in e commerce.
v. Improvement in available internet services previously it is 2G but now 4G is available which speed our work and also attract people to do everything on net from the social networking to banking and the bill payments etc.

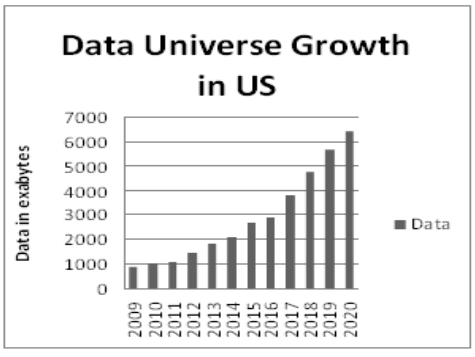


Fig: 1 Data Universe Growth in US[1]

The graph shows how the data increases gradually from 2012 to 2020. This increase in data is known as "Big Data" and the cause of this growth is increase in the use internet, popularity of E commerce which attracts the people to use them etc.

This processed data is store somewhere in the hard drives, in the form of backup files or in the server's storage drives, phone memory etc. Even according to the ILM strategy, which state that the value of information changes with time when created the value is high and is used frequently but with time value drops but it's still has a space in our storage device we move that less frequently used data to archives because can be used in future and that data which stored in the archives are known as "data at rest". So, Data at rest is defined as the data stored in storage system (e.g. databases, data warehouses, spreadsheets, archives, tapes, off-site backups, mobile devices etc.) that is inactive. From the term "inactive" we mean which is not in use for long period of time or processed by the central processing system. Data at rest is the data which is inactive and does not change frequently whereas Data In use is the data is the active one and changes or updated frequently. **[2]**

CONCERNS ABOUT DATA AT REST

The data at rest or the archived that is of great importance specially in the business enterprises because it may consist of the past records , some of the confidential information which should not be disclosed to the unauthorized person and everything is on cloud whether business confidential information, student database, banking information etc. so data is in the cloud and accessed by every ambiguous person .so, to provide security to that confidential data from unauthorized person by taking into account the confidentiality, integrity and availability of data is a major concern today.

So, how these CIA traits is been compromised is discussed below:[3]

➢ **Confidentiality:**

It refers to sensitive and protected data should be accessed by authorized user. The attacks from the insider or external threat the confidential information.

Insider threats: malicious insider from inside the organization.

External threats: The threats from external attackers are from the internet public cloud. Cloud delivery models are affected by external attackers, particularly in private clouds where user endpoints can be targeted. The cloud consist of customer all sensitive information ,credit card details, personal information, government sensitive information, intellectual property is being attack by the attacker . This can be any form hardware attack, social engineering and supply chain attacks.

Data leakage: This issue particularly occurs in public clouds, where an attacker through a virtual machine attack another cloud instance and compromise the sensitive or confidential information from the same physical layer.

➢ **Integrity:**

It refers to the stored data to be the one as transmitted by the user .How the Integrity of data is being affected is discussed further:

Data segregation: in cloud the data is shared among multiple customers the data of multiple customer is stored and processed on the same computer which do effect the integrity and it is difficult to ensure data segregation.

User access: the poor access control may threaten the store data because any ex-employees or any other employees of cloud provide organization may access to administer customer cloud services and can cause intentional damage to the system.

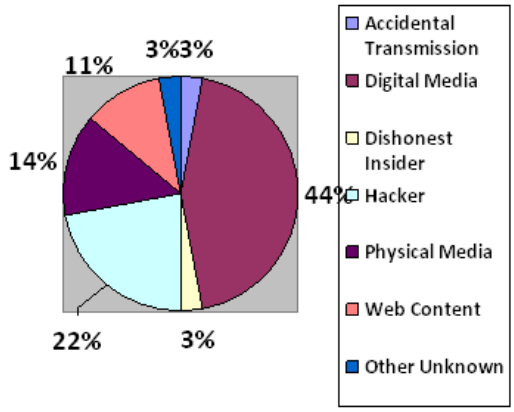
➢ **Availability:** the change in the cloud service model as the increase in the demand of the change management. These changes of hardware or software to existing cloud services threaten the stored data.

Denial of service : the attacks from the networks which is actually the external agents coming from the public cloud services introduce themselves as the internal agents gain access to the stored data and cause DoS.

Physical disruption: the physical damage failure to the data center threatens the stored data the damage may cause by the internal attacker or external attacker.[4]

Above all are the major concern of security of data at rest there can be many more security issues I have tried to discuss some of them. After discussing the security issues threatening data at rest here comes the thing in mind what can be the measures or the ways to addresses these security issues.

According to the study , since January 2005 more than 260 million records containing sensitive information are involved in security breaches in united states.



So, from the above analysis we find that two most frequent type of data breaches involving personal information occur from the loss of digital media (44%) and hackers (22%).this can be prevented by data at rest security solutions if the data secured before the incident.[5]

Preventive measures for securing data at rest

Almost inevitably, information is going to end up spread across multiple devices and networks with varying degrees of security and risk. Before you can take effective action to mitigate your risk, you need to have answers to the following questions:

1) What types of sensitive data does your organization store, use, or transmit?
2) Who has access to this data?
3) Where, when, and why are they using it?
4) How is data stored when it is not in use?
5) How is access to databases controlled?
6) What mechanisms are used to transport data?
7) What are the pertinent laws, regulations, and standards?

Once you have a solid grasp of the potential risks, work with data security experts to determine the next steps to implement a total information security strategy. But don't wait for the risks to make themselves clear; by that time it will almost certainly be too late to take effective action.[6]

There is some of defense mechanism through which this security concern can be resolve

Available security solutions

Many of the reputed IT companies like **IBM, EMC, Google**, etc. are already working on the security solutions of data at rest.

➢ **Vormetric** is the first data at rest security solution available in IBM cloud marketplace. The software uses the encryption and key management, specifies access policies and privileged user that allow only authorized user to access the information.[7]

➢ Encrypting and decrypting data with symmetric cipher keys, AES cryptographic algorithm is used for encrypting. AES uses one of the 3 key lengths: 128,192 and 256. The larger the key length the more computation requires the greater security it provides.

➢ Oracle advanced security transparent data encryption (TDE) by stopping the attacker to read sensitive information from the storage by encrypting the data at database layer.[8]

➢ RSA is the encrypting algorithm used by many organizations for encrypting the stored information.

➢ Google who is famous for its security services not only secure the stored database with encryption but also check security at various level of organization like people accessing the information their employees, customer etc. .they gathered the information from inside and outside the organization and protect the stored information from possible vulnerabilities.[9]

POSSIBLE SECURITY SOLUTIONS TO DATA AT REST

One of the best way to secure the information is to prevent the security attacks, it can be be taking the some of the preventive measures before any security breaches. The best way to secure or to restore the lost information by having

the backup copies of it .this can be achieved by replicating the stored data.

Apart from this some of the other security solutions to data at rest are:

Encryption one of much known way of securing information is encrypting the information. If encrypt the stored information to some coded form this will prevent it from being lost or damage from any unauthorized access.[10]

One can also filtered the information before being stored in the storage device. Here filtering means finding out the information which is not important and storing the important information into the storage devices and applying the security passwords which will allow only the authorized person to access the information.it is important to filter the data because it is not guaranteed that your data is actually secure and is merely impossible to secure huge amount of data ,that is the reason data scientist are working on the big data so that to analyze and filter the data and secure that important data and disposing the one which is not important ,this approach not only reduce the data so that storage space can be easily provided to the useful data but also ease our task to handle it and to provide all the security approaches to that filtered data.

Another approach can be biometric identification, though it's a new approach and expensive one better it is more secure than the traditional password and encryption technique.

Content address storage can be used instead of physical storage like physical drives etc., because the physical device consist of physical address which can be copied whereas content address storage store the information with a unique identity which is unique to the stored information and cannot be copied not only this CAS guarantee authenticity and integrity and fast data retrieval compare to physical drives.

CONCLUSION

Though backup of the store information is the preventive measure to secure the data at rest which may leads to recovery of information in the case of disaster but which again increases the data and need more storage space among all approaches of security solutions like encrypting etc. the best way to secure the data at rest is to filtered the least important data by distinguishing with the important one and dispose the least important data ,securing the important one with the best security approaches like biometric or by applying the hashing algorithm to secure the information this not only reduce the need of extra storage space.

REFRENCES

[1] John Gantz and David Reinsel:"The digital universe in 2020,Big data bigger digital shadows. And biggest growth in far east" (2012). .

[2] Identity finder," the data loss prevention:data–at-rest vs data-at motion",www.identityfinder.com

[3] EMC ,"A detailed review on the approaches for encryption of data ay rest in the enterprises".

[4] Gazala Matlloob "A research survey on cloud computing security issues and it's possible solutions"International Journal of Advance Research in Computer Science Volume 8,N0.2 March 2017 ISSN NO. 0976-5697.

[5] NateLord,https://digitalguardian.com/blog/history-data-breachesA blog on "History of data breaches ".

[6] Ken Beer,Ryan Holland,Amazon web services,"Encryption data at rest" (2014).

[7] "https://www.vormetric.com/company/ newsroom/press-releases/vormetric-is-the-first-data-at-rest-security-solution-available-in" ,April 2014.

[8] Oracle" Encryption and Redaction in Oracle Database 12c with Oracle Advanced Security" March 2017

[9] G Suite Encryption, Google cloud" " How Google Uses Encryption to Protect Your Data".

[10] Intel," Protect Your Company's Data with Efficient Encryption and Secure Remote Management .