ECONSPEAK: A Journal of Advances in Management IT & Social Sciences,
Vol. 6, Issue 10, October 2016 Impact Factor 4.282 ISSN: (2231-4571)
www.skirec.org Email Id: skirec.org@gmail.com

A STUDY ON DATA LEAKAGE PREVENTION FOR SUSTAINING COMPETITIVE EDGE

Abdeali Patanwala
Student of B.E, SVITS, Indore

ABSTRACT: Organizations apply data leakage prevention solutions to monitor and control data access and usage in the field of information security. As the organization progresses into the more technological environment, the amount of the digitally stored data increases dramatically, but keeping the track on data used in any organization is no longer as easy as before. While doing business there is a need to maintain the sensitive and confidential data. If the confidential data is leaked from the organization then it may influence on the organization heath. Data leakage is the unauthorized transmission of data or information within an organization or from an organization to the external destination. This study has examined Redis broker as a test case and offers security layer to prevent information leakage through Redis broker. The goal of this work is to extend the capabilities of such information brokers by implementing security layer on top of data service to prevent information leakage.
Keywords: DLP (Data Leakage Prevention), Information Security, Redis Broker, Transmission.

INTRODUCTION

Data Leakage is an incident when the confidentiality of information has been compromised. It refers to an unauthorized transmission of data from within an organization to an external destination. The data that is leaked out can either be private in nature and are deemed confidential whereas Data Loss is loss of data due to deletion, system crash etc. Totally both the term can be referred as data breach, has been one of the biggest fears that organization face today.

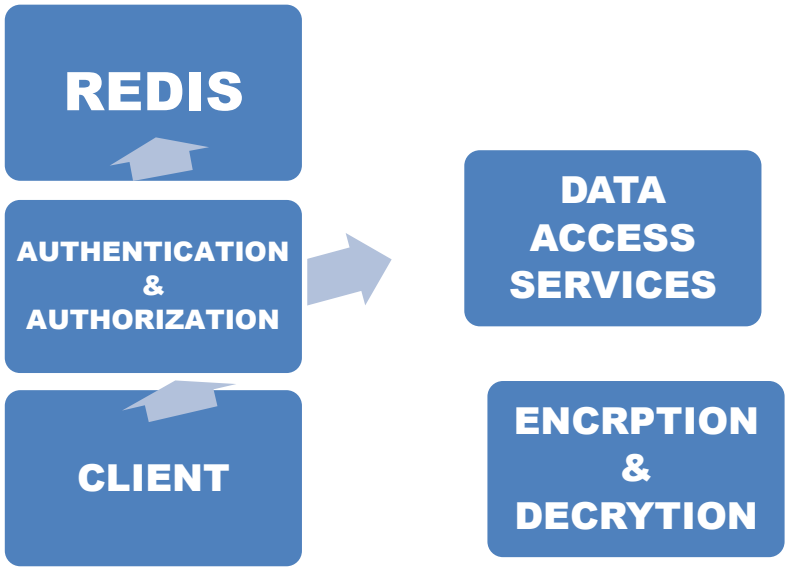
Data Leakage Prevention (DLP) is a computer security term which is used to identify, monitor, and protect data in use, data in motion, and data at rest by using deep content analysis to per inside files and with the use if network communications. DLP is mainly designed to protect information assets in minimal interference in business processes. It also enforces protective controls to prevent unwanted incidents. DLP can also be used to reduce risk, and to improve data management practices and even lower compliance cost.

There is a large security gap between the existing systems which are used to prevent the data leakage and the real life scenario. The gap analysis is undertaken as means of bridging that space. It is a technique for determining the steps that are need to be taken in moving form a current state to desired future state. Security gaps are nothing but the vulnerabilities or weakness in the organization which is a threat and can be exploited to make an attack.

There are two ways of attacks such as External and Internal. External Attacks are those attacks which are done by hackers and other people from the outside of an organization network. It is done by finding the vulnerability and exploiting that to make an attack. Internal Attack is performed from the internal perimeter of the organization by a disgruntled employee, contractors or vendors either for monetary benefits or to take away some confidential, sensitive data out of the organization. Software code, PCI DSS information, financial reports, NER report are few examples of inside attack which are performed from inside of the network.

Redis is a quite popular open source light weight information sharing system, a Key-Value pair based database which is predominantly broker-agent model based. Through a data agent to data broker (server) the data communication involves entities to store data of any data type. However such a system does not enforce any session or access control mechanism. Hence an intruder can easily get access to the data if he has the partial information about the data definition. Any intruder acquiring knowledge about the keys can get the values. Also as no authentication is supported in Redis system, anybody can acquire the knowledge of the data. In order to secure NSQL database and to provide highest data security, a Redis Client which provides Authentication, Authorization and Encryption services for both text and multimedia-binary data.

MODEL ON REDIS SYSTEM



REVIEWS

Information leakage can occur in the common ways corporations envision. Many organizations may invest heavily into firewalls, anti-virus software, encryptions and intrusion detection systems in order to protect lost of data to unauthorized parties such as external hackers (Mallery, 2009). Data leakages can also occur through comments left by a developer in the system script codes or HTML for future debugging or integration, providing unauthorized parties a view of how scripts work or even passwords and usernames used during the development phase (Information Leakage, 2005). Other times, organizations may collaborate with other organizations and information is shared between each other. A risk is posed through this as reliance in the other parties' system is required (Alawneh & Abbadi, 2008). However, the least expected form of data leakage can occur through insider parties of the organization. It appears data leak occur the most from insiders of the organization. This is supported as one study found that 87% of confidential information leaked out is from insiders (Baek, Kim, & Lee, 2008). Therefore, this is important for both management and auditors to take note of.

Many times employees may use simple portable devices (e.g. USBs, flash drives). These devices may look harmless as they become part of the normal work life and go unnoticed if employees use them to carry corporate data in and out of the organization. The main issue here is the ability to carry and store large amounts of files and data in these devices and import it to another computer easily (Mallery, 2009).

These devices also do not require technical expertise from the user when compared to hackers who need IT knowledge, and almost every computer used today contains a USB port (Mallery, 2009). Although many businesses may recognize the threat post by USB flash drives, a lot of the times portable "lifestyle" devices (iPods, MP3 players and digital cameras) are ignored (Mallery, 2009).

Employees may intentionally use these methods to transmit private information or use them as ways to vent their feelings as part of their normal lifestyles (e.g. diary of the work day through a blog). As these forms of communication are usually not monitored by the organization or logs kept for the information transmitted through these methods, confidential information (e.g. trade secrets) can easily be leaked out (Mallery, 2009). The use of e-mails to transmit data can be one of the worst leaks as found by a pharmaceutical firm during a data leak audit. Employees may send unencrypted confidential zip files, e-mails marked as confidential (which raises flags) or even unfinished research documents that are not encrypted (Gittlen, 2009). Through this, the firm is exposed to legal, regulatory and business partner risks if data leak occurs. (Gittlen, 2009).

THE CONCEPTUAL MODEL

This Model contains following major modules. In this subsection a brief overview of the modules has been presented which are:

a) Authentication and Authorization: We first create a universal key which can hold all username password pairs. Any user wanting to access the system needs to first register with the system.

b) Command Restriction: We restrict the commands to **Get, Set** commands and system does not support any other commands. This is done to show the capability of restricting command sets. Also we provide high level command for storing and retrieving images. The images are also secured.

c) Encryption Services: Every variable and their values are encrypted and packed in a key which is derived from the username. Therefore every user has a single data node or key in the Redis database that contains all other keys in the encrypted form. Symmetric key cryptography with AES is used for encryption service. In the symmetric key cryptography, decryption is performed using the same key as that of encryption. The keys are generated from the salt of password created at the time of login. Images are first converted into binary data and then are encrypted using AES using block cipher AES technique. Encrypted images are stored as binary data in the Redis database. Images of a particular user are packed in an image key corresponding to that user. Therefore the database contains an image node related to every user.

d) Network Security: Encryption is always performed at the client side. Therefore the data that is propagated in the network is secured data. Therefore chances of the data being acquired by intruder by network packet hacking is minimized.

e) Data Persistence and Persistent data security: Persistent data is one which is saved in the Redis database. This data should be available at any time when user comes back. This is done by packing the user variable which contains variable of current and previous sessions and their values being packed in the user data node in the encrypted form.

USAGE OF DLP TECHNOLOGY

There are various technologies being used in the organization to prevent data loss. Though these technologies are very powerful but can help majorly an outside attack on data, whereas the current DLP technology deployed is mainly focused on inside attacks. Below are the currently used technologies in the organization for preventing security breaches. Here we concentrate on how these technologies are addressing the security issues in comparison with DLP.

1) Anti-Malware:

Anti-malware is software used to protect malware attacks on computers, this software get into the operating system's core or kernel functions in the same way as malware, which attempt to operate from there. Each time the operating system does some job, the anti-malware software checks that the OS is doing an approved task.

Though this anti-malware software works in real time environment very effectively but it only looks for threats from outside, by scanning and signature validation it ensures that malware infection be removed. This anti-malware software helps in data loss prevention from external threats but for internal threats it doesn't have any mechanism.

2) Firewall:

Firewall is a software or hardware that helps in keeping network secure. Its objective is to control the incoming and outgoing traffic of networks by analysing the data packets and determining whether it should be allowed through or not. A network's firewall frames a brigade between an internal network (Secure), and an external network, i.e. Internet (Insecure). There are different types of Firewalls used in the organization, and it is one of the best security features to be implemented.

But the major problem here is Firewall works on Access Controlled List often know as ACLs. These ACLs either allow or deny completely. For example if a rule is set to deny any outgoing traffic with certain set of data, then it will block all such traffic and it will not even allow the legitimate traffic to flow.

3) IDS/IPS:

Intrusion detection system (IDS) is a device or software application that monitors network or system activities for malicious activities. It identifies a potential security breach, and logs the information and gives an alert by signalling. Intrusion prevention systems monitor network, system activities for malicious activity. It mainly identifies malicious activity, log information, attempt to block/stop activity, and report activity. Though both IDS/IPS and Firewall relate to network security.

4) SIEM:

Security Information Event Management (SIEM) is a tool used on enterprise data networks to centralize the storage of logs which was generated by the software running on the network. This has various features such as gathering information, analysing the information and also presenting the information from network and security devices; identity and access management applications; vulnerability management and policy compliance tools; database logs; application logs; external threat data and OS. It monitors and helps manage user and service privileges, and directory services; as well as providing log auditing and review and incident response. Though this technology can collect events or logs and store for certain period of time but it doesn't have the capabilities of preventing/protecting data loss . All the above technologies are used to prevent external attacks and act very minimal for preventing Insider attacks/threats. In contrast to the above technologies used for Loss Protection/Prevention, DLP provides a policy-based approach to secure data. It enables customers to classify their sensitive data, discover data across the enterprise, enforce controls, and generate reports to ensure compliance with established policies.

The following figure shows the exact scenario.

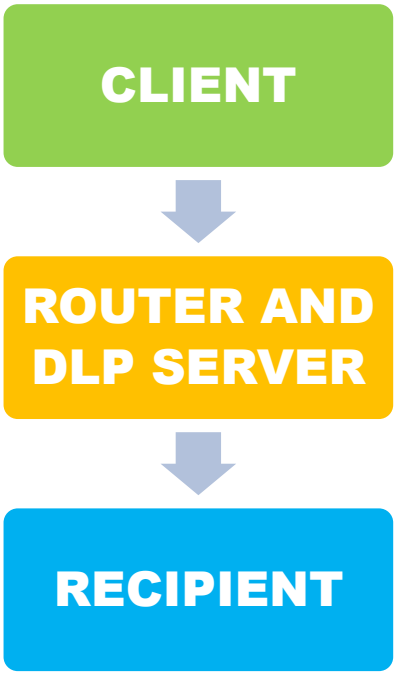


Figure 1: Transmission of an Email between client and DLP server

For the execution of the above DLP solutions it requires the following modules:

1. Email System
2. DLP Server
3. DLP Algorithms for pattern matching
4. Router Programming
5. File corruption system
6. Integration module to integrate all modules

CONCLUSION

If proper procedures, policies and controls are not taken Data leakages can be detrimental to an organizations. Management needs to understand what data leaks are, its effects, and proper measures should be taken as outlined in this study. However careful analysis and preparation is required to help monitor and prevent against data leaks through DLP solutions, especially if the communication involves open source protocols like NoSql or MQTT. In such cases, data is released in a shared buffer. Various schemes are suggested in order to prevent such undesirable information leakage. By using public key cryptography many solutions improve the inherent drawback. In public key cryptography, end entities exchange a key sacredly and the key pair is used to encrypt and decrypt data. But due to extra overhead in number of packets, it is not suitable for open stack information sharing like NoSQL. Therefore in this study Redis too has been applied which is a popular open source light weight network data structure storage and enhanced the security by offering access control, session management, authorization and authentication and symmetric key cryptography. The study also shows that the proposed system prevents the leakage of data. Results show that the system does not add communication or time overhead for the security enhancement. However symmetric key cryptography adds extra space complexity. The study has also extended the framework by offering an image storage in Redis database which does not support multimedia storage. By adding both image encryption and decryption, it has also extended the security to multimedia data.

Future Scope

Our framework can be utilized as a future scope to build digital based system where the information is entrenched in image and encrypted image is saved in the database instead of the plain image. Such a technique will improve the information security to greater extent.

REFERENCES

* Alawneh, M. A. (2008). Preventing Information Leakage Between Collaborationg Organizations. Proceedings of the 10th International Conference on Electronic Commerce (pp. 1-10).
* Aishwarya Potdar et, al., (2014) , Data Leakage Detection In Networks. KJCOEMR, Pune University, Pune, India.
* Gittlen, S. (2009). Inside a Data Leak Audit. Network World.
* Mallery, J. (2009). Overlooked Data Leaks. Security Technology Executive , 78-80.
* Sandip A. Kale & S.V.Kulkarni (2013) "Data Leakage Detection" Department Of CSE, MIT College of Engg, Aurangabad, Dr.B.A.M.University, Aurangabad (M.S), India.
* Rudragouda G Patil (2012) Development of Data leakage Detection Using Data Allocation Strategies. Dept of CSE, The Oxford College of Engg, Bangalore.
* Rekha Jadhav G.H.Raisoni (2014) "Data Leakage Detection" Institute of Engg. And Technology.