MANAGED DATA LOSS PREVENTION SECURITY SERVICE IN CLOUD

Deepak H. Sharma*, Chandrashekhar A. Dhote †, Manish M. Potey

**K J Somaiya College of Engineering, Mumbai, India, deepaksharma@somaiya.edu.*
†P.R.M.I.T & R, Amravati, India.

Keywords: Cloud computing, Data Loss Prevention, Managed Security Service.

Abstract

Managed Security model emphasizes on security to be managed by cloud service provider; in managed security model the security is provided as cloud service instead of security solutions being provided on-premises. Data Loss Prevention (DLP) focuses on location, preventing, reporting and acting on violation of security policies related to Data. The DLP controls are used to ensure that the protected data is used by authorized users only. The DLP controls aim to protect Data against security risks for its entire life cycle from Data-at-rest, Data-in-motion to Data-in-use. The managed DLP security service implemented as the cloud service benefits the user with all the advantages offered by managed Security service model. The objective also is to make the process of Data loss prevention transparent to the user application. This paper discusses implementation of a Proof-of-Concept (POC) framework of managed DLP security. This POC has also been evaluated. The novelty in the approach is that the POC is integrated with other managed Security service options to provide a portal through which various security services can be provided. The users need only browser to use this service. The relevant standards and technologies are also discussed for preventing Data Loss or Leakage.

1 Introduction

Cloud computing is a computing model which is very rapidly evolving. The researchers all around the world are adding new aspects and capabilities to cloud computing domain. Cloud computing has its roots in large-scale distributed computing technology. It is in fact an extension model from grid computing, distributed computing, and parallel computing [1]. Nicholas Carr equates the rise of cloud computing in the information age to electrification in the industrial age. Carr argues that in the emerging future organizations will simply plug in to cloud (computing grid) for the computing resources they need [2].

Managed Security Service model focuses on managing security in the cloud by cloud service provider. In this model the security is provided as cloud service instead of security solutions being provided on-premises. The managed security model can increase the functionality of existing on-premise implementations by working with them in a hybrid manner. Data Loss Prevention (DLP) focuses on location, preventing, reporting and acting on violations of security policy. The DLP controls are used to ensure that the organization's protected data is used by authorized users only. The DLP must protect Data against security risks for its entire life cycle from Data-at-rest, Data-in-motion to Data-in-use. A holistic approach to end-to-end data protection must address the following characteristics [3]:

- Origination verification
- Integrity
- Confidentiality and access control
- Accountability

This paper discusses an implementation of a Proof-of-concept (POC) of Managed Data Loss Prevention security service framework. In particular this managed DLP security service is an on-demand portable, and available as pay-per-use cost model. The paper addresses various issues regarding DLP security delivered as cloud service. This paper addresses the following issues in separate sections. Section two discusses the related work. Section three describes scope and framework of the POC managed DLP service implementation. Section four evaluates the POC implementation and Section five concludes the paper and discusses the future work.

2 Related work

In CSA SECaaS defined category of service [4], the core functionalities of Data Loss Prevention are defined as Data labelling and Classification, Identification of Sensitive Data, Structured Data Matching, Regular Expression Detection, Automated Incident Response, Cryptographic Data Protection and policy management for protection of data throughout its life cycle.

DLP in the cloud as discussed in [3] presents certain risk factors that should be evaluated and addressed before the full extent of the cloud DLP solution's potential can be realized. Each phase of the data lifecycle presents risk factors that, from a DLP perspective, may be consolidated into three broad categories: Data in Motion, Data in Use and Data at Rest. Data Loss Prevention helps understand and address the following:

- Who is sender of data?
- What data is being sent?

- Who is receiver of data?
- Backup of data
- Appropriate use

The various policy considerations for implementing DLP as suggested by [3]

- Information classification (confidential, private, public)
- The nature of the information (different types of data, legal or trade secrets)
- Who is allowed access to this information
- Where the information is allowed to be used/sent/stored
- The severity of exposure
- Notification and alerts
- Actions to be take in response to detection of policy violations

In article [5] the authors have discussed the importance of DLP and elaborated on prevention of Data leakage. The various types of data losses their consequences have been discussed. The need for DLP has been established w.r.t various Government and Regulatory bodies. The various controls for DLP for various loss modes have been suggested. The various DLP solution capabilities required are manage, discover, monitor and protect from data losses. The several controls suggested are automatic encryption, quarantine, restricting printing, saving, copying, accessing, moving and downloading sensitive data. For overall effectiveness a DLP solution must operate without decreasing system performance.

In paper [6] the authors have surveyed and studied various DLP systems for comparison with other security and data protection approaches. The approaches like IDS/ IPS, Anti-Malware, Firewalls use Deep Packet Inspection methods for Data Loss Prevention. The survey discusses various DLP solutions for different phases of data throughout its lifecycle: data-at-rest, data-in-transit, and data-in-use. Security measures like data encryption and access control are used for data-at-rest. An agent is used to monitor data while it is being transported from endpoint to peripheral devices. DLP solutions are used to detect and monitor data while being sent across communication channels. Different DLP methods like content matching and Learning method have been proposed for effective and efficient DLP solution. The authors have compared Deep Content Inspection methods versus existing Deep Packet Inspection methods to prove effectiveness for former for DLP solutions. The disadvantage of the DLP system is that it cannot read encrypted data and the data hidden within images, audio and video content.

3 POC Implementation of Managed DLP Security Service

The DLP solutions can be implemented in different ways depending on Cloud Deployment model. The primary need is to handle data loss at various levels, Network (Data-in-transit), Storage (Data-at-rest), end point (Data-in-use – portable devices), and file level (protection of sensitive files / objects).

The architecture issues to be addressed are protection of data stored in the cloud, its location, users allowed to store the data in the cloud, data that is allowed to leave the cloud, encryption of data on cloud, identifying the users who are sending / receiving sensitive data, and methods of data access.

- A DLP solution should avoid any false positives.
- Encryption helps in confidentiality and integrity for the data from point-to-point.
- Encryption of data also provides a persistent layer data protection.
- DLP service should provide policies for the access control of data at rest (when in the cloud), as well as in transit.
- DLP service must have the ability of notifying and alerting the administrators and/or users in case of any violations of security policies.

The POC for managed DLP security service system architecture is shown in Figure-1, the main components of the system are DLP core – the core DLP functions are implemented in this module. These core functions have been implemented to handle data loss at various levels viz. Network, Storage, End point, File level. It would mean implementing at multiple locations in the set up. In DLP manager – the managerial functions like policy enforcement, security settings are implemented. The user application is on one side and on the other side of managed DLP service is the protected Data. The access to the users is provided through a web browser. As shown in Figure 1 when the data is entering managed DLP service engine it is in doubtful state represented by (?) mark but once it crosses the DLP security engine all the policies are enforced and before it reaches users or before it is updated it is verified represented by (√) sign. Thus preventing, detecting, and monitoring any possible Data losses.

The implementation has been done in form of a web servers running under windows OS Virtual machine in a public cloud setup. The DLP core and DLP manager have been implemented in separate Virtual machines. The user uploads a file which is to be inspected. In normal implementation this process will be transparent to user application. For example before attaching to email or saving on cloud server etc. The DLP core module applies encryption to data being stored on the cloud. This will help protect the data loss for data-at-rest. For outgoing data/ file the DLP core uses the method of content matching which is suitable for structured as well as unstructured data, with the help of keywords, pattern matching, regular expression, different file types, and other information to detect data loss incidents. The warnings are generated and user is alerted if any confidential data is leaking from the system. This will help protect data-in-transit. The novelty in the approach is that the POC is integrated with other managed security service options to provide a portal through which various security services can be provided.

In the POC currently only text files have been used but it can be extended to allow any types of files and other resources too. The objective is to do the process transparently to the user's application. The sample set of keywords has been used based on survey from various web sources for DLP. Figure 2 and Figure 3 shows some snapshots of the working prototype

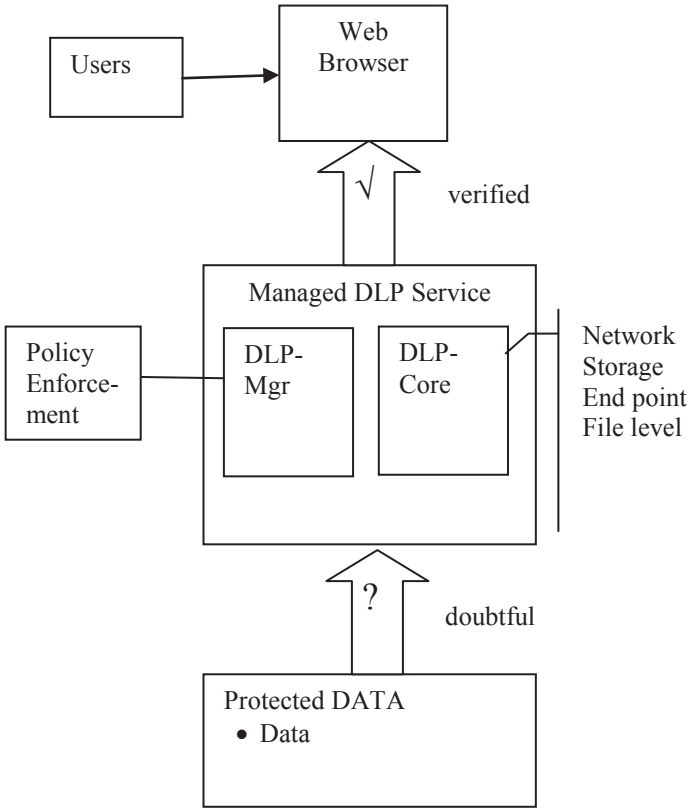

Figure 1: Managed DLP Security Service

Figure 2: Snapshots of POC of Managed DLP service

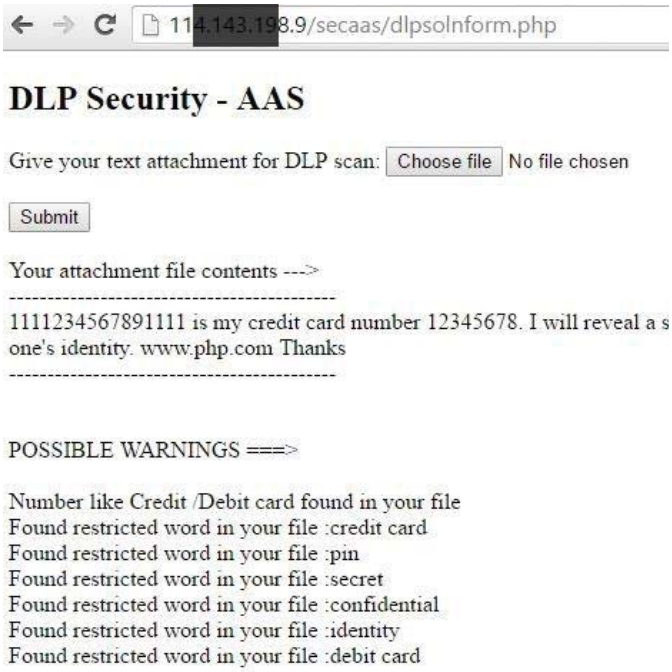

Figure 3: Snapshots of POC of Managed DLP service

4. Evaluation of POC Managed DLP Service

The evaluation of the POC is done with respect to our paper [7, 8] and white paper [9] in which the following criteria have been considered for the evaluation purpose:

- Reliability: the service can be provided in form of multiple web servers running in the cloud environment. The redundancy of servers will lead to high reliability and high availability to the clients. The POC was tested in form of two web servers to provide uninterrupted DLP services to the clients.
- Effectiveness: to make the service more effective, encryption has been provided to protect data-at-rest as well.
- Performance: the performance was tested by comparing the average execution time of managed DLP service provided as cloud service w.r.t. standard legacy DLP solution for a normal web server. The overall overhead also depends on the traffic in public cloud, but it does not increase by more than around 20-25% (refer Figure 4), which is fairly good given the advantages it offers over legacy systems.

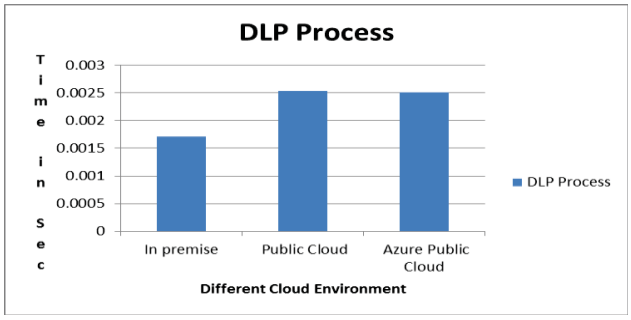

Figure 4: Comparison of performance

- Flexibility: the solution can work with existing legacy systems as well. Since the POC implementation is in the form of PHP programs it can easily work with legacy DLP systems. It can provide more flexibility to customers to choose varying levels of security as per their need.
- Control: the customer uses a web browser to access the service and it can be accessed from various devices viz. desktops and handheld mobile devices etc.
- Privacy and Security: the protected data of the customer is inside a private cloud, and access is provided only after successful screening. This ensures the privacy and security of protected data.
- Cost of ownership: the cost of ownership is borne by the cloud security service provider. The client does not invest in anything in on-premise solution. The client will have to pay only on the basis of pay per use model. Since DLP is available as cloud service it is only charged to customer in form of Operational Expenses (OPEX) model..

5. Conclusion and future work

In this paper, Managed DLP security service is discussed, which is implemented in the form of a framework that enables the cloud service provider to provide DLP as a cloud service. Managed DLP service is compatible with prominent cloud features including portability, elasticity, and pay-per-use service. The approach has been implemented as a collection of VMs in cloud architecture to comply with the cloud model. The protection is transparent to the user's application. This solution can work well with any existing on-premise implemented solutions in a hybrid manner to enhance their security capabilities. With managed DLP security service, users can define virtual private area with the cloud space for securing their protected resources.

The future work will involve enhancing the prototype for the more DLP functionalities. The POC has to be evaluated based on various attacking scenarios, component overhead, performance and effectiveness. The limitation of this DLP system is that it is unable to read data in encrypted form, hidden data within images, audio and video content, these issues need to be addressed in future evolutions of our POC.

References

[1] Shuai Zhang, Shufen Zhang, Xuebin Chen, Xiuzhen Huo, "Cloud Computing Research and Development Trend", 2010 Second International Conference on Future Networks, IEEE 2010

[2] Nicholas Carr,"The Big Switch Rewiring the world from Edison to Google", W.W. Norton & Co. January 2008

[3] CSA SecaaS Implementation guidance : Data Loss Protection September 2012

[4] Cloud Security Alliance, SecaaS Defined categories of service 2011

[5] Liu, Simon; Kuhn, R., "Data Loss Prevention," in IT Professional , vol.12, no.2, pp.10-13, March-April 2010

[6] Tahboub, R.; Saleh, Y., "Data Leakage/Loss Prevention Systems (DLP)," in Computer Applications and Information Systems (WCCAIS), 2014 World Congress on , vol., no., pp.1-6, 17-19 Jan. 2014

[7] Deepak Sharma, Dr. C A. Dhote, Manish Potey, 'Security-as-a-Service from clouds : A survey' IIJC Vol 1 Issue 4, October 2011

[8] Deepak Sharma, Dr. C A. Dhote, Manish Potey, 'Security-as-a-Service from Clouds: A comprehensive Analysis', IJCA Volume 67-Number 3, April 2013

[9] Websense white paper, seven criteria for evaluating Security-as-a-service solutions 2010.