SecaaS Implementation Guidance

Category 2 //

Data Loss Prevention

September 2012

Contents

Foreword

Cloud Computing represents one of the most significant shifts in information technology many of us are likely to see in our lifetimes. We are reaching the point where computing functions as a utility, promising innovations yet unimagined. The major roadblock to full adoption of Cloud Computing has been concern regarding the security and privacy of information.

Much work has been done regarding the security of the cloud and data within it, but until now, there have been no best practices to follow when developing or assessing security services in an elastic cloud model—a model that scales as client requirements change.

One mission of the Cloud Security Alliance is to provide education on the uses of Cloud Computing to help secure all other forms of computing. To aid both cloud customers and cloud providers, the CSA SecaaS Working Group is providing Implementation Guidance for each category of Security as a Service, as delineated in the CSA's SecaaS Defined Categories of Service. Security as a Service was added, as Domain 14, to version 3 of the CSA Guidance.

Cloud Security Alliance SecaaS Implementation Guidance documents are available at https://cloudsecurityalliance.org/research/working-groups/security-as-a-service/.

We encourage you to download and review all of our flagship research at http://www.cloudsecurityalliance.org.

Best regards,

Jerry Archer
Alan Boehme
Dave Cullinane
Nils Puhlmann
Paul Kurtz
Jim Reavis

The Cloud Security Alliance Board of Directors

Letter from the Co-Chairs

Security as a Service is a specialized area categorized two years ago as growing rapidly and in unbound patterns. Vendors were struggling. Consumers were struggling. Each offering had its own path. We felt it was urgent to address the needs and concerns common to the implementation of Security as a Service in its many forms.

The Defined Categories of Service helped clarify the functionalities expected from each Category. In this series, we hope to better define best practices in the design, development, assessment and implementation of today's offerings.

We want to thank all of the many contributors worldwide who have worked so hard to produce these papers providing guidance for best practices in Cloud Computing Security. Many have been with the Security as a Service Working Group since the beginning; many others joined in this effort. Each has spent countless hours considering, clarifying, writing and/or editing these papers. We hope they help move forward toward those unimagined innovations.

Sincerely,

Kevin Fielder and Cameron Smith
SecaaS Working Group Co-Chairs

Acknowledgments

Co-Chairs

Wendy Cohen , GTB Technologies Inc.
Atul Shah, Microsoft

Contributors

Muhammad Asim
Jenn LeMond, Microsoft
Jim Peterson, PKWare
Roshan Sequiera, ISIT

Peer Reviewers

Abhik Chaudhuri
Wong Onn Chee, Infotect Security Pte Ltd
Ivan Macalintal, Trend Micro
Sandeep Mittal
Jean Pawluk
Jim Brigham, EMC/RSA
John Linn, EMC/RSA
Robert Polansky, EMC/RSA
Said Tabet, EMC/RSA

CSA Global Staff

Aaron Alva, Research Intern
Vicki Hahn, Technical Writer/Editor
Luciano JR Santos, Research Director
Kendall Scoboria, Graphic Designer
Evan Scoboria, Webmaster
John Yeoh, Research Analyst

1.0 Introduction

Data Loss Prevention and Data Loss Protection (also known as Data Leakage Prevention/Protection) are terms often used interchangeably to describe the controls put in place by an organization to ensure that data (structured and unstructured) of value remains under authorized use and care. However, prevention and protection are two distinct characteristics which call for different solutions using different controls. The thrust of this paper is on Data Loss Prevention (DLP) as it resides, moves and departs from a cloud service provider offering a security as a service (SecaaS) solution.

DLP must be considered an essential element for achieving an effective information security strategy for protecting data as it moves to, resides in and departs from the cloud. DLP has two facets: one as viewed from the owner's perspective and one as viewed from the custodian's perspective.

DLP is one of the many layers of an information protection and control strategy. What could be an effective control for one organization might not be adequate for another, as every organization must rate the criticality, importance and value of their information. This becomes even more pronounced when data is stored, transits or is handled at a third-party facility which caters to organizations having different classification levels , different lines of business, and which are governed by different laws and legal jurisdictions, all of which must be considered in determining a DLP strategy for the cloud.

1.1 Intended Audience

This paper is an attempt to highlight DLP cloud customer issues and guide organizations—both owners and custodians—to adopt appropriate controls for preventing unauthorized use, access and transfer of data. Section 2 is tailored toward executive level personnel who need to understand the importance of and requirements for DLP security in a cloud environment. Section 3 outlines some of the considerations and concerns that should be part of any discussion surrounding the use of a SecaaS DLP strategy in the cloud. Section 4 provides a detailed description of the architectural components of DLP in the cloud, and provides best practice recommendations for the implementation of Data Loss Prevention as a security service in the cloud.

1.2 Scope

Data Loss Prevention (DLP) is an important element of a broader security strategy surrounding data protection. A holistic approach to end-to-end data protection must address the following characteristics:

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

A policy-based persistent data protection strategy conceptually covers data loss prevention as well as data loss protection. However, this is a broad strategic focus that requires more work and development of the concepts before broader solutions can be commercially available. Point solutions are available in the marketplace today and are used to solve specific problems related to data as it resides in, moves through and departs from a cloud service provider. This document is focused on Data Loss Prevention (DLP) as part of a broader security strategy of policy-based persistent data protection.

This paper focuses on the DLP problem space from a customer's viewpoint. The cloud provider can then extrapolate what is needed to meet the customer's requirements. This guide will address the following:

- What DLP can/should provide
- Assessing risk factors
- Data lifecycle considerations
- Data sovereignty
- Fitting DLP within the broader security strategy

2.0 Requirements Addressed

Effective DLP in a cloud Security as a Service (SecaaS) model must automatically discover confidential ("sensitive") data, monitor its use, and enable administrators to create and implement content enforcement policies. Protection (enforcement) must apply to the content, inside files and packets, depending on policy requirements. Accordingly, the system should be able to identify secure content on any file format even with severe modifications to the original content, to the extent of granularity required by applicable policies.

If implemented properly, DLP can help address the residual risks left behind by inbound security systems such as those covered by intrusion management guidance. As theft of sensitive data is a common characteristic of targeted attacks by determined adversaries (known as advanced persistent threats, APT), DLP is an important layer of defense against such malicious data theft.

DLP in the cloud 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 sending the data? Insiders, intruders, spyware, malware, viruses
- What data is being sent? PHI, PII, source code, IP—including video files, etc.
- Who is receiving the data? IP address, email destination, geographic location
- Backup
- Appropriate use
- Forensics and legal obligation questions

2.1 Fitting DLP within the Broader Security Strategy

The increasing complexity of computing environments in the cloud requires effective data loss prevention to ensure and enforce the appropriate use of vital business data both in motion and at rest. Protecting this information must not disrupt or prevent business operations that rely on this information.

Data Loss Prevention measures should not be considered as the single solution for ensuring data protection in the cloud, and DLP should be used in support of a broader information security strategy which encompasses clearly defined roles, risk assessment, classification, policy, awareness and training. The role of DLP is to locate, report, and optionally act on violations of security policy.

The Data Lifecycle presents various risks at each stage: Data in Motion, Data in Use, and Data at Rest. A DLP system must protect against security risks at all stages of the data lifecycle.

2.1.1 Data Sovereignty

Some cloud service providers (CSP) provide data backup facilities by copying and storing multiple versions of user data across geographical locations. While this creates high availability and fault tolerance, among other benefits, there are growing concerns of the legal complexities and political considerations when cloud data storage spans across sovereign states. These non-technical issues need to be addressed before the true potential of cloud-based data protection solutions can be realized.

2.2 Setting Policy

A policy framework must be established to express security requirements and appropriate use of data. Policy considerations should include:

- 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

2.2.1 Enforcement

Detection of information in violation of policy most often triggers an enforcement response. Typical enforcement considerations for use with DLP within the cloud are similar to those for traditional IT infrastructures, and include the following:

- **Alert & Log** – DLP responses may not always require actions to data found in violation of policy. Notification may be sufficient to report to a user, the data owner, or an administrative/supervisory resource that a file or other data was detected through a policy detection action. This action will not disrupt workflows.
- **Reroute and Pass, or Block/Quarantine** – Policy may require that data not be allowed to move to an intended destination, triggering an enforcement action that blocks transit by either stopping, or quarantining a file. This type of action may disrupt legitimate workflows and is highly dependent upon the detection engine; therefore testing of the provider’s DLP detection capabilities is essential.
- **Delete** – Policy may require that a file be deleted. Legal considerations must be taken into account prior to any data/file deletion.
- **Encrypt** – Unprotected information may trigger an enforcement action to encrypt the data. Encryption may occur in a number of ways, and DLP systems should be capable of integrating with a suitable range of encryption methods. Encryption can avoid disruption of workflows for information by providing recipients with encrypted copies of a file. However, encryption must align with encryption policy. Once information is encrypted, it may not appear again in violation of policy, but users of that now-encrypted

data must be able to access it and use it according to the defined policy. A number of guidelines should be considered for enforcement through encryption:

- Interoperable encryption should be used to ensure that encrypted data could be decrypted across different platforms and applications through standard APIs and interfaces.
- Encryption keys should be held by the data owners or possibly delegated to trusted third parties (preferably not the cloud provider who stores the data).
- Data should be encrypted before it moves to the cloud and attached with appropriate usage permissions. However, prior to encryption, the customers' DLP system may inspect the data based on policy, and then allow the data to pass to the cloud.
- The data owner should be able to specify the policy governing access and use of the data (e.g., which identities/roles can use which data objects for which purposes).

2.3 Cloud DLP Provider Legal/Forensics Requirements

For DLP over the cloud to be feasible, cloud forensic aspects should be considered. Cloud DLP providers should have forensics requirements built into their solutions. Customers should consider what the cloud provider would allow them to do in terms of forensics and investigations post-facto. Legal discovery scenarios may dictate the need to search the customer’s data, and the cloud provider should provide mechanisms to enable such legal compliance. Likewise, the cloud provider should provide mechanisms for differentiating and protecting a customer’s data from another customer’s legal discovery requirements.

3.0 Considerations and Concerns

This section provides discussion on the implementation considerations and concerns related to the policy specification and enforcement in the context of a DLP engine.

3.1 Considerations

3.1.1 Regulatory

There is an extensive set of regulations around the globe that defines the security and privacy requirements for collecting, creating, maintaining, using, disclosing and disposing of individually identifiable information. Regulations from the United States and Europe include, but are not limited to:

- European Commission Directive 95/46 and corresponding national laws of each member state
- EU Data Retention Directive 2006/24/EC
- EU E-Privacy Directive 2002/22/EC
- Privacy Act of 1974
- Right to Financial Privacy Act (1978)
- Privacy Protection Act of 1980
- HIPAA Security and Privacy Regulation (45 CFR § 160 and 164 Part C and 45 CFR § 160 and 164 Part E)

The **Cloud Security Alliance Cloud Control Matrix** provides a useful reference that maps relationships of various industry-accepted regulations to cloud service security concerns.

Security and privacy policies should take relevant regulations into account.

3.1.2 Geographic

An effective information security strategy tries to protect data within and outside of the enterprise boundary. As Cloud Computing becomes more pervasive and is increasingly used for enterprise IT operations, the boundary between the cloud and local storage has become blurred. Data no longer resides in storage facilities under the tight control of the enterprise network and storage administrators. Data may reside on servers and systems running different operating environments, whose whereabouts are no longer fixed, and in some cases may not even be fully known.

Consider the enforcement of DLP techniques on data that is stored in the cloud. Many organizations are averse to data being stored in a different geographic locale that is governed by different rules and regulations. In such cases, the DLP provider could offer solutions to ensure that the data is accessed and used according to the same access and usage policies used in-house.

3.1.3 Policy/Source Considerations

There may be different types of policies based on the entities who specify them. Corporate policies are defined by the corporations which will be using cloud for their organizational IT related functions or for providing services to customers. Policies capture requirements necessary for the protection of information (as specified by business needs, criticality and legislation), while at the same time enabling the workflows necessary for business operations.

User defined policies are needed when corporations use the cloud to provide services to customers. They enable (often mandated by regulations) customers to specify their consent /privacy preferences.

A policy that is input to the DLP engine may be a combination of corporate and user defined policies. A typical policy may include:

- Information about the subjects who are allowed to use the information. The subjects may be indicated by their identities or their attributes/claims.
- Information about the objects that need to be protected by the DLP service. Examples of objects may include the following: customer information, individually identifiable health information (IIHI) or third party owned content.
- Actions (quarantined, released) that a user may perform on data that would be allowed by the DLP service.
- The duration of access rights applicable to a subject and/or object.
- Compliance and legal requirements (e.g., auditing).

3.1.4 Architectural Considerations

Customers should ensure that their business needs can be met by the cloud provider's solutions. Considerations for discussion include:

- What data is permitted to be stored in the cloud, and what is the location of stored Data? This could be important with regard to compliance and regulatory policies.
- Who is allowed to store the data in the cloud?
- What data is allowed to leave the cloud? (If data is encrypted, can you inspect and enforce actions prior to it leaving?)
- Can encryption (if needed) be automated when leaving the cloud?
- If sensitive data left the cloud, are you able to identify who sent it - regardless of protocol?
- Who is allowed to access the Data in the Cloud - including culling of metadata?
- How will data be accessed – types of devices, channels, protocols?
- Assurance levels for the different devices which would be storing the data (for example, sandboxing features if this is a shared resource , hardware capabilities to ensure this resource will be highly available for legitimate use and can withstand current known attacks)?

3.1.5 Enforcement Considerations

The enforcement capability of a DLP service enables the prevention/protection of data according to the policy specified by a data owner. From the cloud perspective, the fundamental question is whether a DLP engine in the cloud provides the same range of enforcement capabilities as an on-premise DLP engine. Consider the following when assessing the enforcement capabilities of a DLP engine in the cloud:

- In DLP enforcement mode, even a small amount of false positive detection can wreak havoc in an organization. Therefore, a DLP solution should employ detection technology to avoid false positives. A high degree of false positives may impact the enforcement of policies. Any degree of false positives also brings into question the detection engine's false negative rate (secured content leaves without being detected by the detection engines).
- TLS/SSL offers point-to-point confidentiality and integrity for the data that is received or sent outside by a DLP engine.
- Solutions for the content centered encryption of data should be reviewed, as they provide persistent protection of data.
- Determine if the encryption functionality provided by DLP service is dependent on a specific file format, and if so, if that format is acceptable. Many standards-based tools (e.g., CMS for XML Encryption standard) are agnostic as to format; these standards consider the data as a whole, and encrypt it. The customer should carefully examine the compatibility of the encryption functionality offered.
- Ensure that the DLP service attaches usage permissions to the data once it leaves its scope.
- Ensure that DLP service provides support for the enforcement of access control policies for information at rest (when in the cloud), as well as in motion (leaving the DLP service).
- Consider whether a DLP service also can integrate and interoperate with other solutions such as rights management. In some cases, the issue of interoperability becomes critically important.
- Ensure that the DLP service has the ability to notify and alert administrators and/or users in case of an incident (e.g., through email).

3.1.6 Encryption Considerations

Encryption of the content results in transformation of the data so it is unreadable without an appropriate key. A DLP service may encrypt the content as required by policy. The following should be addressed when considering the encryption capability of a DLP service:

- The encryption algorithm should meet appropriate regulatory and industry cryptography requirements (e.g., FIPS 140-2).
- Before applying encryption, cross-platform and interoperability needs of the data must be defined to ensure the data will be accessible.
- Determine what the encryption is applied to: data (or object) or storage or transport channel.
- Consider the performance impact of encryption. If necessary, implement selective encryption for performance reasons.

- Data encryption introduces the need for encryption keys and key management. Consider who controls the encryption keys.

3.2 Concerns

3.2.1 Service Level Agreements

Every cloud service provider is subject to local laws and regulations based on its geographic location, making it imperative that they comply with various standards like SAS 70 , ISO27001:2005 , PCI-DSS , GLBA , HIPAA etc. While compliance with these standards provides a level of assurance, regulatory standards may not meet the minimum the requirements of clients who might be spread across different geographies and might not be subject to the same laws and regulations.

Users are advised to discuss and review SLAs with their corporate counsel prior to the execution of any cloud provider's agreement. The majority of standard vendor SLAs will not indemnify or assume liability if a breach occurs.

SLAs should be written by customers to ensure policy compliance and satisfy auditing and regulatory obligations in the areas listed below:

- Communications and networks
- Systems and applications monitoring
- Information security incident handling
- User management
- Protection against malicious software
- Backup management
- Vulnerability management and security audit
- Systems and applications security monitoring
- Mobile computing and teleworking
- Appropriate operating procedures to protect data from unauthorized disclosure, modification, removal, and destruction
- Deployment and availability of required and necessary skills, resources, etc., to assure the quality of the service

3.2.2 Legal Discovery

Shared resources in a Cloud/SAAS model bring the possibility of metadata and/or customer data being culled by another cloud user (different organization) during a legal “discovery” scan for ESI (Electronically Stored Information).

3.2.3 Policy Specification

The goal of the policy specification should not be to overload the consumer of the service with a very complex task. The high level policies should be first translated into the detailed policies and later into the machine readable policies which then serve as input to the DLP engine.

Consider the level of support provided for administrative roles. Can access and administrative rights be controlled selectively for both internal and external administrator functions by role? Can policy administrator access be separated from system administrator access?

There may be conflicts between the policies of a corporation and a user, hence rules needs to be specified for the conflict resolution and user may need to be notified in case his/her privacy preferences are being overridden.

Different corporations and users may apply security controls at different levels of data granularity (e.g., user or category of users, data instance or type of data, etc.). A DLP engine should classify the data according to the policies defined by the data owner. The data classified as very sensitive related to the company/user strategy may require encryption in order to meet regulatory needs.

Identity management is required for the identification of users. Policies should be correlated with the user and objects, and the operations users can perform on the objects.

3.2.4 Detecting Policy Violations

DLP provides the means for detecting policy violations by inspecting information and applying policy to determine whether it conforms to appropriate or authorized use. Inspection may occur in real-time as information moves within network channels, including to and from cloud services, or it may operate on stored data.

3.2.5 Enforcement Concerns

The enforcement capability of a DLP engine enables the prevention/protection of data according to the policies specified by the data owner. From the cloud perspective, the fundamental question is whether a DLP engine in the cloud provides the same range of enforcement capabilities as by an on-premise DLP engine. The following concerns should be addressed when assessing the enforcement capabilities of a DLP engine in the cloud:

- A DLP engine has the ability, based on policy, to block, quarantine or delete information in case of an incident, virtually making it impossible for the entities to get access to the information. How will such a response influence the intended business operation? If information is deleted, is support provided for secure (remote) wiping? Does it sufficiently wipe the content?

- Legal ramifications should be examined as engine response mechanisms are set. Regulatory responsibilities must be accounted for prior to any deletion of information by the DLP engine in the case of an incident.
- Does the DLP engine sufficiently protect all critical channels and ports, both positive and negative? Positive channels and ports are necessary for the smooth management of business workflows. Negative channels and ports are those that can be used to evade detection such as through port-hopping.
- Plans must be made for supporting an increasingly mobile workforce. An organization usually needs the ability to control access to data from entities (devices) that meet the organization's security policy. A DLP engine should have such capability.
- Consider whether a DLP engine has the ability to monitor and protect the data in real time and the ability to allow the use of data only according to the specified policy.

4.0 Implementation

4.1 Architecture Overview

The cloud can be categorized and deployed in numerous ways, with the most common being Private, Public and Hybrid Clouds. Similarly, DLP solutions have many different designs. Some are designed as a reverse firewall, requiring only one server per egress point (monitors and blocks on all protocols), while others require more than one server (DLP Network Prevent Server, DLP monitor Server, 3rd party proxy server, etc.) per egress point. Some providers deploy DLP client agents on customers' endpoints to enforce controls. Customers themselves may even deploy client agents on endpoints (workstations and servers) as part of their overall DLP solution. Choose a provider whose options best meet the needs of the business.

Several common architectures exist for implementing DLP. Providers may offer some or all of these options. Often, centralized administration is available for defining policies and remediation actions that can then be applied consistently to network, storage and endpoint DLP. Understanding these at a high level can ensure appropriate matching of an offering to the need.

4.1.1 Network DLP

The most common type of DPL engine is Network DLP. It enforces applicable policies and encrypts data (with appropriate permissions attached) if it leaves the safe boundary. Network DLP also tracks the transfer of data from the cloud to the end points by monitoring all possible ports.

- **Real-time Inspection & Detection of Content** – The unique advantage of real-time inspection is the ability to prevent data loss over the network on any protocol and in any format. Many current solutions can prevent data loss only on non-real-time protocols: SMTP via MTA, HTTP/S or FTP via a proxy and/or ICAP server. Real-time detection is important because it allows content aware blocking for very large files.
- **Detecting Policy Violations in Real-Time** – Detection applies policies through inspection of the data within the channels monitored by DLP. While the most common scenario is email, DLP policies should be applied to all transfer channels and protocols through which data may pass, including VNC filters. Common detection mechanisms include:
    - **File Detection** – The typical use of DLP consists of data inspection.
    - **Partial File Detection** – Partial file detection may include data modifications such as excerpting, inserting, file type conversion, formatting, ASCII to/from UNICODE conversion, and UNIX to/from Windows conversion, etc.
    - **Decryption** – Information already protected by earlier DLP enforcement or by users may require detection to be preceded by a decryption step to allow scanning for disallowed content that users may seek to hide using file encryption.
    - **Rogue encryption** – Encryption should be aligned with encryption policy. Encryption found that is not compliant with encryption policy must be identified by DLP.

- **Detection Algorithms** – Each DLP provider has proprietary methods to describe and detect sensitive data. However since the key requirement is that the detection results are of high reliability and accuracy, thorough testing is necessary.
    - **Precise methods** are, by definition, those that involve Content Registration and **trigger almost zero false positive incidents**. These methods should be used for Data Leak Prevention (enforcement/blocking). When an organization has a large number of sensitive files, the best approach to securing these files is to use fingerprinting. Information in corporate DBs can be protected by using fingerprinting. This is similar to fingerprinting files, except that each column of each row in the DB is fingerprinted. For example, a security event can be triggered when the fingerprints of a Name AND Credit Card number are detected.
    - **Imprecise methods** are all other methods. They include keywords, lexicons, regular expressions, extended regular expressions, meta tags, machine learning, etc. Those imprecise methods that do not require Content Registration yield high false positive and false negative rates and are useful only if combined with Severity Thresholds. It is unlikely that imprecise methods would be used for Data Loss Prevention (enforcement/blocking); rather, they will be used only for monitoring. However, ultimately the customer's internal risk assessment determines whether to use precise methods only, or accept less than precise methods in some cases.

4.1.2 Storage DLP

Storage detection is very important because it allows discovery of data at rest that may otherwise go undetected and is usually not the focus of network DLP solutions. Storage DLP solutions typically are implemented using servers configured to inspect storage repositories and are often used for DLP engines applied to data at rest. Support is commonly available for file storage such as file servers and SAN/NAS. Increasingly specialized inspection support is available for use with structured data repositories and databases.

Though storage refers to data at rest, the access, use and processing of this information should be subject to the same rules and conditions as the policy that governs access and use of the data in-house. This calls for an understanding of:

- Who accessed the information
- When it was accessed
- How the information was transferred (encrypted or unencrypted form)
- Whether the data was encrypted using point-to-point encryption (TLS) or content centered (or object level) encryption techniques
- Where the information was transferred
- If the access and usage policy was attached when the information was transferred
- Trigger alerts and enforce actions as desired by the policies

To this end, Storage DLP needs to be tightly integrated with network and endpoint DLP rules, in addition to its own set of rules.

4.1.2.1 Architectural Considerations for Storage DLP Solutions

A breach of privilege by an administrator for the physical machine may render all virtual machines vulnerable to attacks. File shadowing features offered by some vendors may need to be checked against this backdrop, as new technologies allow complete virtual machines to be deployed across multiple physical machines.

Where a storage DLP server is placed depends on where the storage repository resides. If the storage repository were hosted in the cloud, then it would be best served by a storage DLP server also in the cloud. Likewise, storage repositories that are on-premise are best served by storage DLP server on-premise as well.

4.1.2.2 Detecting Policy Violations in Storage

Detection applies policies through inspection of the data within the repositories monitored by DLP, as well as access to those repositories by authorized legitimate users using local and remote access. Detection policies should just not be restricted to the "data" aspect, but should be able to analyze and correlate information from other rules which could include, but not be limited to the following: who accessed the information, how the information was accessed (local or remote), day and time the information was accessed, different channels used to move the information (screen captures, copy, etc.), and trigger alerts and enforce actions.

With the use of virtual systems, all information to the virtual systems passes through a shared physical infrastructure. This throws up many challenges both in terms of detection and effective prevention. Policies should include amongst others "time of the day" and " day of the week" access considerations and alerts , local access policy violations and movement of data between virtual systems using screen captures and "$ share" copying .

4.1.3 Endpoint DLP

Endpoint DLP systems may run on the endpoint devices (PDAs, notebooks, iPads) that use and consume data, or on an organization's server. Endpoint systems also may include servers that do not reside in the cloud; for example, file servers, application servers and database servers. Endpoint DLP systems usually have the ability to protect the data at rest (portable devices), in use or in motion. Endpoint DLP solutions monitor data continuously and may provide immediate feedback to the user if actions are found in conflict with the defined access control and security policy.

Endpoint DLP systems include multiple levels of functionality:

- **Basic** – Content discovery scanning and unauthorized file transfer prevention
- **Medium** – Providing consistent enforcement even when the endpoint is off the network (often due to mobility)
- **High** – Focuses on how a user interacts with the system, integration with the advanced protection tools such as digital rights management technologies, and may be used to enforce more advanced access and usage control policies

Customers may want to consider having their own DLP solution to protect data en route to the cloud as well.

4.1.4 File-Level DLP

File-level DLP resembles Storage DLP with the distinction that its main focus is on the identification and protection of sensitive files/objects. Whenever sensitive data is to be sent outside of the cloud, it is first inspected to be sure that said data is permitted to leave. File-level DLP then encrypts the file or object and associates the security policy with the file so that it travels with the data outside the cloud. If the Endpoint DLP engine receives such a file with the associated policy, it then can control and track access and usage according to the associated security policy.

4.2 Guidance and Implementation Steps

4.2.1 Policy

Once a cloud provider has configured the DLP solution for an account, rules (policy) setup begins.
When defining rules, the objects describe a subset of network traffic. Then rules are applied to those objects.

For a quick validation test, set up two object-rules: one using the SMTP protocol and the other using the HTTP protocol. Before defining these rules, use the pattern editor to add a definition to match the text “password.” As a test, define an Object-Rule to send an alert email to the appropriate security respondent in your organization, when an email is sent to a specific email address and contains the word “password” in the body of the email. Prepare an email destined for a restricted email address and place the word “Password” in the body of the email. Send the email to the restricted email address. The Security Respondent should receive the alert notification of the event. Test the other protocol in the same way.

DLP providers often include a range of policy templates that provide packaged policy configurations. Templates typically address policy needs by geography, industry, or regulatory compliance (e.g., FIPS, HIPAA, GLBA and PCI-DSS). Use of templates can often reduce policy setup time.

4.2.2 DLP Implementation

Based on the input provided by business owners, along with the company policies and procedures, define what data needs to be protected, and which users need to be given what rights, depending on defined data access parameters. The success of this effort depends on the availability of a company data classification and control policy. The template derived will be used a ready reference point for the steps to follow.

Some solutions require the use of metadata tags and classification in order to enforce data protection policies. *If the solution does not require metadata tag classification for enforcement, or if fingerprinting detection will be used, go directly to configuration.* Otherwise, specify the policy to identify, detect and apply actions (e.g., encrypt) to the data, based on information deduced from the metadata; or write Object-Rule sets for patterns, keywords detection.

Configure the DLP system for further complex tasks which would make the solution more robust and foolproof (e.g., fingerprinting of the information). Start with basic information—protect employee Social Security numbers, customer credit card numbers, or government issued identity cards. The organization's DBA should know what data resides in the master database.

4.2.3 Fingerprinting

Unstructured data fingerprinting, and where to start that fingerprinting, should be driven by corporate policy with input from stakeholders.

1. Fingerprint and build your Policy (Object-Rule) sets. Define which protocol to monitor and enforce on (block), and the workflow of incident handling (who is to be alerted, workflow per incident, etc.)
2. Run the DLP service in monitor mode for an appropriate amount of time (likely a few days) in order to determine the detection and policy accuracy, data collection and log parsing activities.
3. When the system is functioning as needed, move to enforcement mode.

The steps above should be repeated based on the organization's policies, compliance regulations, and what IP needs to be protected.

4.2.4 Enforcement

After establishing policies for the detection of data in violation of acceptable use criteria, enforcement actions should be configured to respond appropriately to incidents of data found in violation of those policies. Enforcement actions may range from simple actions such as notifying the user, to stronger actions such as encrypting the data/object.

4.2.4.1 Implementing Enforcement

Specify the policy and configure enforcement action(s) as a response to a violation incident. An enforcement action may require either a form of manual intervention, or, where supported by a DLP solution, it may be an automated action reducing operator or user interaction to resolve. The enforcement actions could be:

- **Alerts** – Assign an alert when a notification must be sent in response to a detected incident. Alerts should be configured as email reports or logged events. Alerts of one type or another should often be used in combination with other available enforcement actions. Alerts should go to one or more parties responsible for remediating an incident. Those alerted may include:
    - **DLP Administrator, console operator or security officer** – This resource typically monitors and responds to incidents.
    - **User** – This resource typically is the person that caused the incident to be triggered by performing an action on data that violates policy. A user may attempt to place a file into the cloud when either user is not authorized for that action, or the data is not allowed to move to the cloud. Alerts to users responsible for a violation can provide strong reinforcement to the

user of appropriate use policy. This can significantly help to reduce repeat incidents. Additionally, custom messages can be provided with the alert that inform the user of the violation and include instruction to resolve the incident. This may include steps on how the user should first encrypt the information using the approved encryption software before trying to send the information again.

  - **Data owner** – This resource is the person that is responsible for the security of the information. This person is not necessarily the person that caused the violation, but may be informed of the activity on their data.
  - **Managers** – This resource is a manager of either a user causing the incident, or of the data owner. This alert provider notification is for incidents where supervisory action is required for an incident, such as approval of the release of blocked or encrypted data.

- **No alert** – Policy may require notifications are not sent in some instances, for example, during investigative actions.
- **Block** – Assign a block action when policy requires that data not move to the requested location.
- **Quarantine** – Assign a quarantine action when policy requires that data be placed into a holding location for subsequent review or release approval. Data placed into a quarantine location may also be encrypted while it resides within the quarantine location.
- **Re-route** – Assign a re-route action when data should be routed to another service for additional processing. One example is to route data to a server that applies the standard encryption format to data.
- **Encrypt** – Assign an encryption action when sensitive data is authorized for a requested action but only if appropriately protected. Content centered encryption should be configured to protect data using interoperable and portable encryption methods. Interoperable encryption such as is available through specifications such as the Cryptographic Message Syntax (CMS)[4] standard can be used to ensure that encrypted data can be decrypted across different platforms and applications through standard APIs and interfaces. Encryption may be automated or manual per DLP policy and operational guidance. Portable encryption container formats can be used to ensure that encrypted file data can move to and be reused on any platform, and are accessible for decryption anywhere using universally available programs. These formats commonly include compression and archiving to efficiently store and transfer encrypted files
  - Encryption may be applied to data manually by a user, data owner or administrator when informed of the requirement through a DLP alert. Appropriate instruction provided in an alert can help users to apply the right encryption format before repeating the action to move or store a file.
  - Encryption actions may also be automated with DLP. To automate an encryption action, configure the encryption operation as an automated response. Before configuring an automated encryption action, make sure encryption keys are available through standard sources such as LDAP. To ensure encrypted data can be recovered, include the use of a master or contingency key whenever an encryption operation is performed.
- **Delete** – Assign a delete action when data is not allowed to reach its intended location and should be removed through deletion.

- **Pass** – Assign a pass action when sensitive information is of a low-risk nature that it can be allowed to move to the cloud, but may warrant a notification to the sender that caution should be used for this data as it moves to the cloud. This action allows the data to reach its intended destination.

5.0 References and Useful Links

5.1 References

Callas, J., Donnerhacke, L., Finney, H., Thayer, R. (2007, November). *OpenPGP Message Format*. Retrieved from http://tools.ietf.org/html/rfc4880

Cloud Security Alliance. (2011). *CSA Cloud Controls Matrix*, v1.2 2011. Retrieved from https://cloudsecurityalliance.org/research/ccm/#_version1_2

Cloud Security Alliance. (2011). Security Guidance for Critical Areas of Cloud Computing – Version 3.0. Retrieved from https://cloudsecurityalliance.org/research/security-guidance/

GTB Technologies. *Core Technology of DLP*. Retrieved from http://gtbtechnologies.com/en/company/about/core-technology

Houseley, R. (2004, July). *Cryptographic Message Syntax*. Retrieved from http://tools.ietf.org/html/rfc3852

PKWARE. (2007). *APPNOTE - .ZIP File Format Specification*, APPNOTE.TXT. Retrieved from http://www.pkware.com/documents/casestudies/APPNOTE.TXT

5.2 Useful Links

AES Encryption Information: Encryption Specification AE1 and AE2 http://www.winzip.com/aes_info.htm

Federal Information Processing Standards Publication 197 Advanced Encryption Standard (AES) http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf