ROBERT FRANCES GROUP
Business Advisors to IT Executives

June 30, 2021

# Summary Insights: Data Loss Prevention (DLP) Best Practices

**RFG Perspective:** Data loss can undermine customer trust, damage an enterprise's reputation – and ultimately cause a loss of revenue. Even though email remains the top attack vector for security events, followed by Web applications, many employees are either cavalier or careless in their daily use of corporate data. Therefore, IT needs not only to protect the company from breaches due to employee lapses, but must also put measures in place to ensure that the impact of any breach is minimized.

## INTRODUCTION

While data loss prevention (DLP) has been a major concern in top Financial Services firms and elsewhere, IT has not had sufficient visibility into data movement or employee behavior to minimize the security gaps. Business and IT executives have identified the primary causes of DLP, but most have yet to reduce the exposures that might impact their organization.

There are two main causes for security exposures: emails and Web applications. Employees and business partners fail to be concerned enough about security because it relates to the large number of emails sent or received. This situation results in confidential data and sensitive data being exchanged, without encryption, (internally and externally) without concern for the implications to their organizations. It also accounts for the number of successful phishing scams that continually occur, day in and day out.

From the standpoint of IT development, developers tend to focus on the line of business requirements that are clearly articulated. They usually do not devote the time needed to ensure that all of the security angles have been addressed when writing new code or modifications. This oversight opens the gate wide for attackers who are looking for any flaw to make an entry into the system.

It is common knowledge that there will always be flaws in the perimeter barriers that safeguard security; yet little effort is devoted to ensuring that the blast radius of penetration is minimized. It is clear that IT executives need to address this shortcoming – and the sooner they do so, the better.

### RFG 100 Panelists and Insights from Participating RFG 100 Executives

On May 5, 2021, Robert Frances Group facilitated an RFG 100 videoconference entitled: "*Data Loss Prevention: Gaining Competitive Advantage through New Advancements and Best Practices.*" This video conference examined what processes are required for the development and maintenance of a quality data loss prevention (DLP) program.

The panelists on the call were:

- Deal Daly, CIO, HammerSpace
- Evan Bauer, CEO, OpStack
- Christine Huang, Director, SAP
- Stephen Harris, Enterprise IT Architect (Security), Fiserv

## POLL RESULTS

On May 5, 2021, RFG 100 executives were polled about their biggest DLP concerns and consequences. We found the following results in the polling data for this group (n < 100):

- 80% of all respondents said that they view emails as the biggest threat vector for their organization's data security. 40% of respondents said that they are apprehensive about the exposure of confidential or PII (personally identifiable information) data being openly transmitted in internal and external emails. And 40% were concerned with incoming malicious emails.
- The remaining 20% of respondents said they were worried about weaknesses or vulnerabilities in Web or network code or configurations.
- 50% of all respondents said their organization can detect that an employee or contractor has downloaded, saved, or sent work-related document personal accounts. 40% of respondents felt that they could only determine that in an after-the-fact forensics analysis. The remaining 10% felt that they had no access to such information.
- Regarding the major consequences of a security breach, 40% of all respondents stated reputational damage was the biggest consequence; 30% felt it was loss of customer trust; and the remaining 30% said it was revenue loss.

## Visualization of the Data Estate

An effective DLP strategy must present a clear view of the organization's entire data estate. In other words: You cannot protect what you do not know exists. The best way to do that – and the ideal way to do that – is to have a single pane of glass for the entire data estate. Otherwise, fragmentation of the corporate data estate results in added complexity and potential gaps. However, most organizations do not have that "ideal" 360-degree view of their entire data state. Instead, they may consolidate multiple views – and integrate them into a few "views" that security personnel may view to get a broader picture of the "state" of their corporate data.

One of the most effective ways to address a data estate is to separate the metadata from the data itself. This separation and virtualization of the data estate enables the creation of a unified "data plane" – which helps to simplify the implementation of automated policies and objectives. Separating the metadata from the data itself reduces the number of data copies, allowing managers to choose the least expensive data stores and simplifying their move to the hybrid multi-cloud environment.

## DLP Processes

The panelists and participants provided their expert opinions on ways to improve an organization's DLP processes:

- **Identifying the privileged and/or sensitive information that exists within the organization**. This is the first step in getting the process of managing the data estate started. Data access rules must be applied consistently so that there are no gaps that could enable inappropriate imports or exports of privileged or sensitive information.

- **Determining policies for what is appropriate for infiltration and exfiltration of sensitive information.** Emails and data files will be transmitted outside of, and imported into, the organization through standard business activities. But rules must be in place to prevent data import and export failures. At a minimum, there should be a rule that all data must be encrypted for handling sensitive information – including corporate, PHI and PII information. Policies for data transparency, duplication, export controls, and redundancy must be established, as well as determining which data is suitable for each of the supported cloud environments.

- **Minimizing the risk surface.** Enterprises must limit the penetration of any attack by reducing the control surface for data security. Rather than focusing on scripting, managers must gain a clear understanding how the control surface impacts all of the data traffic (transfers) in the organization. The downside of this approach is that it requires more keys and more access credential restrictions to be issued. While this generates more security and governance work, it limits the size of the risk surfaces and it minimizes the risk exposures of any given attack.

- **Creating data pillars.** Executives can look at the various data challenges through the lens of data pillars. There should be a recognition that everything – all applications and all database accesses – are data-oriented, and that the creation of data pillars can simplify the DLP discussion. Some of the data pillars are the encryption pillar, the resiliency pillar, the anti-malware pillar, and the scanning pillar. Discussing the issues in terms of pillars, and taking action accordingly, makes it easier for non-IT executives and staff to understand the challenges and support the activities.

- **Simplifying, standardizing, automating, and monitoring.** End-to-end automation is a desirable goal, but in the interim, all organizations need to constantly verify that they are secure. This underscores the truth behind the motto: Trust but verify.

- **Managing security keys.** In a multi-cloud environment, enterprises should bring their own key (BYOK), because it is the only way for ensuring data consistency. Giving out a key to decrypt traffic creates a security gap. When protecting data on behalf of customers, decrypt the data (or use homomorphic encryption techniques) and then pass the data on to others. Furthermore, there needs to be a key rotation plan to improve the current state of data protection over time.

- **Building the scan rules and scan filters.** While scan filters are a must-have, a number of iterations are usually needed to properly deal with standard usage behaviors while posting a limited number of false positives. Also, the rules must be applied consistently across the enterprise. One participant said his organization found that it took 12 iterations before they had the filters reasonably tuned.

- **Moving toward having one management plane.** To achieve the goals of effective DLP, it is necessary to have the knowledge of the totality of the data and metadata. Consolidating inputs for data monitoring allows organizations to categorize and classify the data, and to translate the policies into DLP rules, attributes, and terms. Having a dictionary and policy templates will allow an organization to address company-classified and secret data, as well as highly sensitive personal data.

- **Frequently updating the data classification policy.** A very comprehensive data classification policy must be established and kept current. Businesses and regulations change constantly, as do the actions of end-users and suppliers (including IaaS, PaaS, and SaaS cloud providers and software vendors). IT groups cannot ignore the fact that businesses constantly expand into new sectors or make acquisitions (M&A activity) – and both forms of expansion have major impacts on the types and quality of the data being captured, stored, transmitted and used.

## Security Processes and Speed Bumps

The participants expressed a number of security concerns relating to the DLP process and its implementation. Here are some key points:

- **People forget to manage the security processes as relates to DLP.** The DLP and the data discovery process needs to be validated with compliance, legal and security teams. Participants said the single biggest challenge they face is the continuing use of data silos in the multi-cloud environment. These data silos cause inconsistencies in an organization's highly distributed data universe, due to inclusion of unstructured databases – many of them from cloud-native applications – and labeling inconsistencies by the hyperscalers. When cleaning up this process, one should start with the applications and datasets that account for the bulk of the business revenue, because they generally carry the greatest amount of risk for the company and its financial results.

- **Taking inventory of the data landscape.** There is a need to do a thorough discovery against all the data being managed across the organization. The lack of consistency across all types of data in hybrid multi-cloud environments is a major concern – and must become a focus for remediation. Organizations must use risk-based analysis to prioritize next steps in their plan to address data inconsistencies, because they lack the time and resources to immediately address all of the inconsistencies found in the discovery process.

- **Creating "speed bumps" to prevent DLP workarounds.** It is often the case that users in the organization will find a way to avoid, or to go around, the roadblocks put in place to safeguard data protection. That is why there is a need to create enough "speed bumps" to discourage or slow these activities down, especially for business units that are impatient to deploy new cloud-enabled systems. These steps are different than other security controls, which represent more formal ways to protect data, such as access control encryption.

## Beware of Data Drift

Once DLP plans move into production, continual monitoring and management are needed to ensure that the plans are properly executed and maintained. Following Day 1 DLP implementations, Day 2 brings its own set of challenges that must be constantly monitored. The participants identified the following key points to monitor:

- **Shifting borders.** Amazon Web Services (AWS) and other cloud service providers (CSPs) continually introduce changes, which can shift the "border" for security perimeters. Typically, CSPs do not inform you of the impact caused by their changes. For example, for AWS services, there could be Alexa triggers or Lambda function changes that breach the zero-trust parameters. When the AWS relational database service (RDS) is used, users are dependent on AWS for the controls. Therefore, all of your cloud assets need to have authentication and authorization for all service changes.

- **Human error.** Administrators often click the wrong button by mistake on their management console, which can introduce a gap in the data protection protocols. The key is to know what a good (known) configuration is, and then keep monitoring that configuration for data-drift.

- **Many CSPs do not provide two-step validation.** In this scenario, one person makes a change and it goes "live." Clearly, there needs to be a second person who approves and executes it. Separation of duties is critical to DLP success (and to key management) to ensure the configurations are not drifting from accepted parameters.

- **Broken applications expose security credentials.** Once an attacker can bypass all of those controls, they can freely get and exfiltrate the data. That is why sensitive and confidential data must be encrypted throughout the organization.

- **Application developers in a DevOps environment may be unaware of the differences in the multiple levels of encryption.** When a CSP provides encryption capabilities, it is usually at the server-side, block level or total database level – and not at the field level. This means that if the attacker gets access to the database, he or she has access to all the data. With field-level encryption, general access does not grant the hacker access to the individual records. Enterprises need to build and to manage their own security keys and to keep the key management in a separate environment.

- **Deleting old keys.** If you should decide to move out of a vendor, delete those keys and crypto-shred all your data. Otherwise, there could be remnants of it that other, non-certified users might be able to access.

## Identity, Access Management and Encryption

- **Frequently cleanse access management rights to applications and data.** When somebody joins a team, they get the access to the applications and data they need to perform their assignments. But when this person leaves the team, the access rights are usually not removed. That is why there needs to be an incentive or automated process to ensure that this is type of data-access failure does not occur.

- **Monitoring encrypted data traffic.** One of the trickiest bits for a multi-cloud multinational firm is setting up the ability to have the installed DLP software suite actually looking at encrypted traffic. This becomes a tricky challenge for key management unless one has an enterprise-wide distributed secrets management capability in place.

- **Using higher-level controls.** Utilizing application stack controls instead of trusting a bunch of network controls is an alternate or additional approach worth considering. Networking controls have not always worked as desired – and they may provide the gateway to higher-level applications, which have access to secure data. That is why access monitoring needs to move "up the stack" to ensure there are application controls in place.

## Summary

DLP is both a major business challenge and a technical one. It is so pervasive and important to the organization, that it needs to be understood by all levels of control throughout the enterprise. Many end-users remain blissfully ignorant of the roles they can play in protecting the data upon which their enterprise depends. Yet, most organizations do not devote much time and energy to getting everyone to understand the true impact of data loss. Clearly, awareness of the importance of DLP in safeguarding corporate-wide security must be top-of-mind for all the enterprises that are rapidly moving into the hybrid multi-cloud world.

**RFG POV:** The three most important assets of an enterprise are its data, employees, and intellectual property; yet, ironically, data is the least respected and protected asset. Loss of critical corporate data can damage or impair the other two assets (people and IP) – and put the entire enterprise in jeopardy when a security breach occurs. Business and IT executives must devote funding and resources to ensure that key corporate and PII data are properly protected at all times. To ensure that these important data practices are effectively established and maintained, a DLP program must be established, monitored, and reported to the Board.

*Additional relevant research and consulting services are available. Interested readers should contact Client Services to arrange further discussion or interview with Mr. Cal Braunstein, CEO and Executive Director of Research. Jean S. Bozman, President of Cloud Architects Advisors LLC, co-authored this report.*

–----------------------------------------------------------------------------------