Technology Trends

Data Leak Prevention (DLP)

Enterprise Architecture, Chief Technology Officer Branch

Version 0.1
Date 2019-07-31

Shared Services Canada Services partagés Canada

Canada

Table of Contents

Business Brief

Data Leak Prevention (DLP), also known as “data loss prevention,” is a cybersecurity solution that includes a variety of strategies, processes, and tools whose purpose is to protect an organization‘s valuable data from being accessed by unauthorized users, released into an untrusted environment, or destroyed.

The term data leak or data breach refers to confidential information being released by an insider or an external threat for nefarious purposes. Examples of an organization’s valuable information can include financial data such as a credit card numbers, personal identifiable information (PII) such as the user’s identity, username, password and user activity, intellectual property like patents, trade secrets or source code, or classified documents.

Without implementing countermeasures, an organization risks the *Confidentiality, Integrity, and Availability,* known as the CIA or AIC Triad, of their data by leaving themselves vulnerable to cyberattacks. In the past, examples of these incidents have cost organizations millions of dollars in damages and loss of brand reputation.

DLP provides the tools to mitigate data leak incidents from occurring within an organization. DLP software usually includes the following functionalities:

- **Protection:** DLP tools implement safeguards such as encryptions, access controls and restrictions to mitigate possible vulnerabilities. An organization can regulate file access by classifying data according to their level of security and by defining a set of rules each user has to abide by.

- **Detection:** DLP can alert administrators by generating a real-time detailed report on policy violations such as an attacker attempting to access sensitive data. By creating a baseline behavioural profile of standard patterns, the software can detect abnormal or suspicious user activity. Some solutions accomplish this using machine learning.

- **Monitoring:** DLP monitors the behaviour of users on how the data is being accessed, used and moved through the IT infrastructure in order to detect irregular or dangerous user activity. If an event is triggered by a rule violation, the system will notify the security personnel. The system gains visibility in order to proactively secure data from leaving the organization on policy violations.

Technical Brief

Data Leak Prevention is the practice of detecting and protecting confidential information against data loss, data leakage and data breaches. Cyberattacks are caused by hackers, spies or even insiders, whose objectives include: to damage IT infrastructure, for financial or political gain, status or revenge.

In this ever-changing landscape, there are several factors that contribute to the increasing threats:

- **Data Value:** The monetization of data has created an environment that encourages the persistence of cybercrime.
- **Multitude of Access Points**: Many businesses embrace new technologies like social media and mobile devices, but thereby increase their exposure to internal threats by offering data escape paths.
- **Cheap IT Storage Units:** Modern storage units are light and cost less, making it easy for an employee to walk out the door with gigabytes of data.
- **Decentralized IT Systems:** This type of architecture provides many benefits like openness and information sharing, but makes it difficult for organizations to track and control their information due to lack of governance.

DLP technology is usually categorized into three different components related to each state of the data lifecycle: data at rest, data in motion, and data in use. In most DLP products, there is also a central management server acting as the control center of the DLP deployment. This is usually where DLP policies are managed, data is collected from sensors and endpoint agents, and backup and restore is handled. The components of a data leak prevention tool are, in general:

Storage DLP: "Data at Rest" refers to data stored on a "device," for example, on a server, database, workstations, laptops, mobile devices, portable storage or removable media. The term refers to data being inactive and not currently being transmitted across a network or being actively processed. A storage DLP protects this type of data by using several security tools:

- Data masking hides sensitive information like personal identifiable data.
- Access controls prevent unauthorized access.
- File encryption adds a layer of protection.
- Data classification uses a DLP agent to tag data according to their level of security. Combined with a set of rules, an organization can regulate user access to use, modify and delete information.

- A database-activity monitoring tool inspects databases, data warehouses (EDW) and mainframes and sends alerts on policy violations. In order to classify data, some mechanism uses conceptual definitions, keywords or regular expression matching.

Network DLP: “Data in Motion” is data that is actively traveling across a network such as email or a file transferred over File Transfer Protocol (FTP) or Secure Socket Shell (SSH). A Network DLP focuses on analyzing network traffic to detect sensitive data transfer in violation of security policies and providing tools to ensure the safety of data transfer. Examples of this include:

- An email monitoring tool can identify if an email contains sensitive information and block the action or encrypt the content.
- The Intrusion Detection System (IDS) monitors for any malicious activity occurring on the network and typically reports to an administrator or to the central management server using a Security Information and Event Management system (SIEM).
- Firewall and antivirus software are commonly available products included in a DLP strategy.

Endpoint DLP: “Data in Use” is the data currently being processed by an application. Data of this nature is in the process of being generated, updated, viewed, and erased on a local machine. Protecting this type of data is a challenging task because of the large number of systems and devices but it is usually done through an Endpoint DLP agent installed on the local machine. Some characteristics are:

- The tool provides strong user authentication, identity management and profile permissions to secure a system.
- It can monitor and flag unauthorized activities that users may intentionally or unintentionally perform, such as print/fax, copy/paste and screen capture.
- Some DLP agents may offer application control to determine which application can access protected data.
- There are advanced solutions that use machine learning and temporal reasoning algorithms to detect abnormal behavior on a local machine.

Industry Use

Implementing data breach and data leak countermeasures is a major concern for the industry. Over the years, a wide range of high-profile companies have been subjected to these incidents. The biggest security breach of all time happened to Yahoo in a series of breaches in 2013 and 2014, which resulted in all 3 billion user accounts being hacked and personal information being leaked. The company only first disclosed these events in 2016. At the time, the company was in the process of being sold to Verizon but these events had lowered the selling price of $350 million and it received 43 class action lawsuits as a result.

Due to the constant risk of possible breaches, such as in the example above, Data Loss Prevention technology is widely adopted amongst the tech industry to protect their data. When it comes to enterprise solutions, Gartner identifies four leading DLP vendors: Digital Guardian, Forcepoint, McAfee, and Symantec. Market worth around DLP is growing: in 2015, its estimated worth was around $0.96 million and is expected to grow to around $2.64 billion by next year at a Compound Annual Growth Rate (CAGR) of 22.3%. While data breaches and cyber-attacks have historically been the driver for demand, the growth of cloud storage will increase demand into the future. Furthermore, as things such as the use of digital services, social media, the Internet of Things (IoT) and e-commerce expand, the production of data, even big data, will grow with it as will the need for storage, whether on cloud or through other means. Thus, the desire and regulatory obligations to protect data, such as through DLP, will expand as well.

The DLP market used to have the same approach with respect to monitoring and protecting an organization’s data, but modern solutions differ significantly and have become more individualised. The traditional approach, sometimes called a project approach or a suite, involves a network gateway acting as a man-in-the-middle to monitor the traffic. It requires that the source, destination and type of sensitive information is known and well-defined. The newer method, sometimes referred to as the data visibility or individual approach, uses an agent installed locally on each system to monitor all user and system activity. This approach works well an organization is still in an age of discovery regarding its transmittal and sharing of data and most networks users would potentially have access to sensitive forms of data. The majority of organizations employ both DLP approaches to varying degrees.

Canadian Government Use

The Government of Canada (GC) has a responsibility to protect not only its data and IT assets but also that of its citizens and the data collected on or about them. Despite this, the GC itself is not free from experiencing data leaks. For example, the Canadian Revenue Agency (CRA) reported 3,763 data breaches in 2013, including incidents where taxpayer's information were lost, compromised, or accidentally released. In order to prevent such occurrences, as well as those on both smaller and larger scales, there are various DLP protocols in place throughout the GC. Currently, DLP operations are run independently in each department. However, this is in concurrence with federal supporting policies and procedures, some of which also extend to industry.

As of November 1, 2018, private Canadian businesses and industries, along with the health sector, which are subjected to The Personal Information Protection and Electronic Documents Act (PIPEDA), are required to report all data breaches involving personal information that may harm an individual, hold a record of all data breaches, and notify the affected individuals. The goal of this act is to assure citizens have their personal information protected by appropriate safeguards in accordance to their right to access their personal information. Similarly, the federal Privacy Act stipulates how GC departments can share and provide access to personal information on or about individual Canadian citizens and also mandates reporting of security breaches involving this data.

Since the GC relies extensively on IT to provide its services, the Operational Security Standard from Management of Information Technology Security (MITS) as well as the Operational Security Standard – Business Continuity (BCP) Program defines a baseline of security requirements which federal departments and agencies must fulfill to ensure the security of information are under their control. Those prevention safeguards include incorporating identification and authentication in all networks and systems, authorization and access control to restrict accessibility on a "need to know" basis, proper cryptographic and encryption protocols, and emanations security methods such as TEMPEST. In the event of a data breach, the Policy on Government Security (PGS) establishes a mechanism to coordinate the response and recovery. Since the data breaches are primarily caused by people, the Canadian Centre for Cyber Security offers up-to-date publications as part of an awareness campaign.

The Government of Canada's Cloud Adoption Strategy, as well as the Strategic Plan for Information Management and Information Technology 2017 to 2021 outlines a move towards increasing the use of cloud services for data storage and processing. Outsourcing to private clouds presents a certain level of risk if vendors are not vigilant against cyberattacks or if malicious themselves. The GC has developed various strategies, guidelines and best practices in order to mitigate the risks around cloud and Cloud Service Providers (CSPs). For example, the Direction on the Secure Use of Commercial Cloud Services: Security Policy Implementation Notice outlines measures such as third-party independent assurances, encryption and cryptographic algorithm, and vulnerability alerts, amongst others, as part of its attempt to minimize risk and heighten data loss prevention.

As with other nations, creating an open, collaborative, and accessible government is of prime importance to the Government of Canada. As described in the Digital Operations and Strategic Plan (DOSP), it holds that sharing data and information with Canadians and businesses with help to grow the economy and allow for more active participation in public life. Open portals and open information can present a more open possibility of breaches and attacks, however. Therefore, moves towards open government must involve DLP controls. Making data and information more open has inherent risks – it exposes networks, systems, devices and data, including personal information, to accidental or malicious breaches. As such, robust IT security protocols in the GC are of paramount importance. A layered security approach, such as the use of trusted access, protected assets, secure protocols by default and continuous monitoring are already in effect and will continue to be implemented in the GC.

Implications for Shared Services Canada (SSC)

Value Proposition

The value proposition of DLP relates directly to SSC's mandate to design and operate a secure IT infrastructure that protects GC data and technology assets. The primary business value in implementing a DLP strategy is the reduction of risks and impacts associated with data leaks. These incidents often affect an organization in the following aspects:

- **Operational:** A data breach often causes an interruption of services until the investigation process is concluded – this can take weeks or months, costing an organization business or other resources in the meantime. DLP ensures redundancies are put in place to counteract important data losses, thereby avoiding cost to operational resources to remediate lost data. In 2015, SSC implemented the Directive on the Use of USB and Other External Storage Devices to help manage these sorts of risks. All of SSC's electronic assets have a DLP software-based tool that monitors the use of unauthorized devices on the network. This prevents removal of data from the SCC system or prevent infecting the system with any malware, viruses or other malicious entities. A second phase of SSC's DLP program is in the planning stages and will monitor enterprise data in motion and at rest – this is already in place in terms of secret data, however.

- **Financial**: There are significant financial losses resulting from data breaches, including fines, audit fees and legal expenses. The Ponemon Institute has estimated that the average global cost of a data breach has risen to $3.9 million and $5 million in Canada specifically in a 2018 study. Contrast this to the average annual cost of a subscription based DLP solution of approximately $175,000, according to Forrester.

- **Reputational:** Data losses affect the reputation and damages the brand. Often, organizations will see a drop in their valuation, which results in potential loss of future revenue, their competitive advantage, and their market shares. Consequently, the consumer trust in the organization also experiences a decrease which can have large-scale ramifications in short and long-term revenues. Having a DLP solution helps disassociate the user's concern for safety and builds clients' trust.

Challenges

Integrating a DLP solution in the infrastructure is a complex undertaking, involving many components such as a database analyzer, an email system, a web proxy, etc. Adding to the complexity, data security and DLP initiatives face several difficulties as a result of the modern technological landscape. When it comes to integrating a DLP solution, there are several challenges and issues that are relevant to SSC:

- **Complex DLP integration**: Generally, enforcing DLP technologies is complex, varies depending on the organization's network architecture, and requires to work across many components such as security, networking, infrastructure, email, web, endpoint, storage, databases. Deploying, configuring, and managing these DLP systems is also complicated. In order to fully protect an IT infrastructure, it is important to employ a holistic approach, however organizations often do not have a clear strategy toward DLP and balancing new ways of working.

- **User Awareness and Engagement**: Organizations face several challenges of control over their employees' actions. It's common for employees to lack awareness, accountability and responsibility for their actions. Some training and awareness campaigns do not focus enough on protecting sensitive data and using security tools like file encryption. There is also a general sense that there is no risk involved in breaking the rules.

The following trends will continue to be a challenge for IT service providers in protecting data:

- **Emerging Consumerism**: The availability of computing devices and connectivity to the internet anywhere at any time has its benefits. Unfortunately, it facilitates the disclosure of personal or proprietary information by providing several exit points to the web. Policies like "Bring Your Own Device" (BYOD) are vulnerable to loss of physical assets such as laptops and end users may unintentionally spread confidential information through social media.

- **Business Continuity and Disaster Recovery:** The technological climate forces organizations to have 24/7/365 system availability. Outages interrupting the continuity of IT services could cause financial and reputational loss.

- **Persistence of Cybercrime:** Since data has real world value, cyberattacks are becoming more frequent and more sophisticated. While the majority of attacks are from external sources, The Verizon study estimates that 15% of the attacks involved insiders losing or stealing devices, transferring data to personal storage, etc.

Considerations

As with any program or tool, it is necessary to align policies with controls. The GC already has various policies in place pertaining to IM/IT infrastructure, including the security of these resources and information. However, if an organization has policies in place that prohibit or monitor certain activities but a control is not yet in place, or completely absent, then data leak still poses a large risk to the organization. Security policies exist but departmental compliance and control implementation remains an issue.

Although DLP protocols and controls have already been implemented into much of SSC's IT infrastructure, there are some areas in which improvements should be considered. With government-wide strategies around "Open Government" and "cloud computing," SSC will face increasing need to adapt DLP tools into these platforms as they evolve and expand.

Once aligned with policies, which may change and evolve as time goes on and technology advances, SSC must be prepared for its DLP controls to change with it. Leading experts in the area of DLP define DLP as a dynamic process, not an end-state. A robust DLP program is an opportunity to work with stakeholders and set the expectation that protocols should change and be adjusted over time. DLP must also be considered when the network architecture and tools change, SSC should evaluate how security checks are integrated into new projects.

Furthermore, while SSC will play a main role in procuring DLP tools for departments and delivery these services, the protection of data requires a team effort. Collaboration in terms of monitoring, surveillance, and the granting of access to local or departmental networks and resources will be needed. Also, engaging stakeholders helps to identify vulnerabilities that may otherwise be missed. A mindset of collective responsibility is a best practice for ensuring the most effectiveness of DLP.

One way of helping to achieve buy-in around DLP as an ongoing process, as well as creating a culture of collective responsibility, could be for SSC, along with its partner departments in the GC, to establish "Security Champions". The GC has introduced a national champion, Mr. David Jean, the GC's Champion of Security, to be the link between departmental security and national security interests, with respect to all forms of threats or safety issues, not only those related to cybersecurity. However, cyber-specific champions could also be introduced at a more local level and advance DLP "on the ground" as suggested in the Summer-Fall 2016 Consultations: Information technology Transformation Plan – What We Heard Final Report. Such employees can help promote the importance of security protocols and behaviours, and can be an important part of the DLP framework.

However, DLP tools and processes cannot work in isolation of systems and users. Without proper operationalization, DLP runs the risk of offering a false sense of security and merely becoming a risk generator. The SSC Departmental Plan of the Cyber and IT Security program identifies the following five risks with respect to cybersecurity, of which DLP is a part:

- **Resource Capacity**: SSC may not have the adequate financial and human resources to improve services and to introduce the latest technologies to counteract cyber threats.
- **Aging IT Systems**: Current IT infrastructure is at risk of failing due to its end of life.
- **Cyber and IT Security**: SSC is at risk of not being able to respond efficiently to IT security and cyber security threats, which would result in proprietary information being compromised and disaster recovery activities being impeded.
- **Service Delivery and Management**: SSC's enterprise tools and processes are at risk of not being able to improve the delivery of services to partner organizations.
- **Availability and Quality of Information:** Lack of availability and integrity of information will impede effective planning and decision-making.

References

Arellano, N. E. (2014, March 31). *Data breaches in federal departments soar in 10 months*. Retrieved from IT World Canada: https://www.itworldcanada.com/post/revenue-agency-bumps-up-government-data-breach-numbers

Brooks, R. (2018, November 29). *What to Know about a Data Breach: Definition, Types, Risk Factors and Prevention Measures*. Retrieved from Netwrix: https://blog.netwrix.com/2018/11/29/what-to-know-about-a-data-breach-definition-types-risk-factors-and-prevention-measures/

Canadian Centre for Cyber Security. (2019, May 15). *Five practical ways to make yourself cybersafe*. Retrieved from cyber.gc: https://cyber.gc.ca/en/guidance/five-practical-ways-make-yourself-cybersafe

Digital Guardian Guest Contributor. (2018, February 5). *Getting Successful with DLP: Two Approaches for Quick DLP Wins.* Retrieved from Digital Guardian: https://digitalguardian.com/blog/getting-successful-dlp-two-approaches-quick-dlp-wins

DLPexperts. (2019, may 17). *DATA LOSS PREVENTION BUYER'S GUIDE & VENDOR COMPARISON*. Retrieved from DLPexperts: https://dlpexperts.com/data-loss-prevention-buyers-guide-and-vendor-comparison/

Ernst & Young. (2011, October). *Data loss prevention*. Retrieved from EY: https://www.ey.com/Publication/vwLUAssets/EY_Data_Loss_Prevention/$FILE/EY_Data_Loss_Prevention.pdf

Governement of Canada. (2004, May 31). *Operational Security Standard: Management of Information Technology Security (MITS)*. Retrieved from Governement of Canada: https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=12328

Government of Canada. (2018, December 13). *The Privacy Act.* Retrieved from Government of Canada: https://laws-lois.justice.gc.ca/eng/acts/P-21/

Hughes, C. (2014, September 3). *The Three States of Digital Data*. Retrieved from ASPG: http://aspg.com/three-states-digital-data/#.XN7E0aBKi71

Imperva. (2019, May 17). *Insider Threats*. Retrieved from Imperva: https://www.imperva.com/learn/application-security/insider-threats/

Imperva. (2019, May 17). *Security information and event management (SIEM)*. Retrieved from Imperva: https://www.imperva.com/learn/application-security/siem/

Imperva. (2019, May 17). *What is a Data Breach | Tips for Data Leak Prevention | Imperva*. Retrieved from impperva: https://www.imperva.com/learn/data-security/data-breach/

Imperva. (2019, May 17). *What is Data Loss Prevention (DLP) | Data Leakage Mitigation | Imperva*. Retrieved from imperva: https://www.imperva.com/learn/data-security/data-loss-prevention-dlp/

Janacek, B. (2015, December 1). *Best Practices: Securing Data at Rest, in Use, and in Motion*. Retrieved from DataMotion: https://www.datamotion.com/2015/12/best-practices-securing-data-at-rest-in-use-and-in-motion/

Larson, S. (2017, October 4). *Every Single Yahoo Account was Hacked - 3 billion in all.* Retrieved from CNN Business: https://money.cnn.com/2017/10/03/technology/business/yahoo-breach-3-billion-accounts/index.html

Markets and Markets. (2015, September). *Data Loss Prevention Market by Solution Type (Network DLP, Storage DLP, Endpoint DLP), by Deployment Type (On-Premise, Cloud), by Applications, by Service, by Organization Size, by Vertical, and by Regions - Global Forecast to 2020.* Retrieved from Markets and Markets: https://www.marketsandmarkets.com/Market-Reports/data-loss-prevention-advanced-technologies-market-531.html

Meizlik, D. (2008, February 5). *The ROI of Data Loss Prevention*. Retrieved from Websense, Inc. : http://img2.insight.com/graphics/uk/media/pdf/whitepaper_roiofdlp_en.pdf

Office of the Privacy Commissioner of Canada. (2018, January). *PIPEDA in brief*. Retrieved from priv.gc: https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/pipeda_brief/

Osakwe, M. (2018, July 19). *Data Breaches vs. Data Leaks: What's the Difference?* Retrieved from NextAdvisor: https://www.nextadvisor.com/blog/data-breaches-vs-data-leaks-whats-the-difference/

McCarthy, Niall. (2018, July 13). *The Average Cost of a Data Breach is Highest in the U.S.* Retrieved from Forbes: https://www.forbes.com/sites/niallmccarthy/2018/07/13/the-average-cost-of-a-data-breach-is-highest-in-the-u-s-infographic/#58c9dcd32f37

Shared Services Canada. (2018, April 24). *2017–18 Cyber and Information Technology Security Branch Business Plan*. Retrieved from Shared Services Canada:

http://myssc-monspc.ssc-spc.gc.ca/en/worktools-processes/integrated-business-planning/CITS

Shared Services Canada. (2018, February 2). *Data Loss Prevention and the Use of Portable Storage Devices.* Retrieved from Shared Service Canada: http://myssc-monspc.ssc-spc.gc.ca/en/employee-centre/security/it-security/data-loss

Shared Services Canada. (2019, April 11). *SSC business planning*. Retrieved from Shared Services Canada: http://myssc-monspc.ssc-spc.gc.ca/en/worktools-processes/integrated-business-planning

Treasury Board of Canada Secretariat. (2018). *Digital Operations Strategic Plan: 2018-2022.* Retrieved from Treasury Board of Canada Secretariat: https://www.canada.ca/en/government/system/digital-government/digital-operations-strategic-plan-2018-2022.html#ToC8

Treasury Board of Canada Secretariat. (2017, November 1). *Direction on the Secure Use of Commercial Cloud Services: Security Policy Implementation Notice (SPIN).* Retrieved from Treasury Board of Canada Secretariat: https://www.canada.ca/en/government/system/digital-government/modern-emerging-technologies/direction-secure-use-commercial-cloud-services-spin.html

SiteUptime. (2017, June 8). *Data Leakage Vs Data Loss: What's The Difference?* Retrieved from SiteUptime: https://www.siteuptime.com/blog/2017/06/08/data-leakage-vs-data-loss-whats-the-difference/

Verizon Enterprise Solutions. (2019, May 17). *2019 Data Breach Investigations Report*. Retrieved from Verizon Enterprise Solutions: https://enterprise.verizon.com/resources/reports/dbir/

Wikipedia. (2019, May 10). *Data Breach*. Retrieved from Wikipedia: https://en.wikipedia.org/wiki/Data_breach

Wikipedia. (2019, May 5). *Information Security*. Retrieved from Wikipedia: https://en.wikipedia.org/wiki/Information_security

Zhang, Ellen**.** 2019, January 3). *What is Data Loss Prevention (DLP): a Definition of Data Loss Prevention.* Retrieved from Digital Gaurdian: https://digitalguardian.com/blog/what-data-loss-prevention-dlp-definition-data-loss-prevention