DATA LOSS PREVENTION & DLP SOLUTIONS EXPLAINED



Corporate networks are continually hardened against the growing risks of data loss. Yet, sensitive business information is lost at unprecedented rates around the world.

These cybersecurity stats give some idea of how severe data loss is:

- 1 billion data records were exposed in the first six months of the year 2019. (RiskBased)
- A cyber-attack is performed every 39 seconds, 2,244 times a day. (University of Maryland)
- The average cost of data breach is $3.92 million. (Security Intelligence)

It takes a strategic approach and advanced cyberdefense capabilities to protect sensitive data against sophisticated cyber-attacks. That's where data loss prevention comes in.

What is data loss prevention?

Gartner coined the term data loss prevention (DLP). Here's how Gartner defines DLP:

> *"A set of technologies and inspection techniques used to classify information content contained within an object—such as a file, email, packet, application, or data store—while at rest (in storage), in use (during an operation) or in transit (across a network)".*

The DLP practice encompasses the strategy, process, and the technologies involved in securing:

- Business information
- Data workloads

Companies develop security policies according to industry compliance regulations, such as HIPAA or GDPR. Once you've developed your security policies, you can use DLP technologies like contextual analysis and data inspection to execute your security measures about preventing data leak.

How data leaks happen

Before discussing how DLP works, let's understand how data leaks occur. These are typical ways of data leaking from a storage or network source:

- **Insider threat.** Disgruntled employees misusing authorized ways to access data and transferring sensitive business information outside of the organization's domain.
- **Extrusion.** External threat vectors compromising systems or the human element to obtain unauthorized access to sensitive information.
- **Unintended data exposure.** Loss of confidential information to the public, usually through media or social media channels. This can also occur by employees, without malicious intent.

How DLP technology works

Now let's see how DLP handles these data leaks.

Companies store data in both on-site data centers and public cloud systems. Both users and software applications access data through private or public internet networks. This usage leaves an activity profile—known as patterns or traces—that you can compare against the expected or authorized approach of consuming the data.

DLP solutions perform two functions:

1. First, evaluate the context of the network traffic.
2. Then, apply the appropriate data security policies.

Context analysis deals with the broader activity around data access and transmission, accounting for the three risk factors described above—inside threats, extrusion, and unintended exposure. Context analysis works in connection with content awareness, which deals with the data that may be leaked or misused.

In order to counter these actions, you can use the Data Leak Protection solution as a proxy system that monitors and evaluates the context of data access and transmission activities. It's designed to look out for a range of policy violations, using several content analysis methods:

- **Rules-based expressions** analyze the data against specific policies. For example, a file containing 16-digit credit card numbers must be encrypted at all times.
- **Database fingerprinting** is useful for structured databases, as it applies rules to information with specific attributes and content. For example, collecting purchase information of customers in a specific geographic location while ignoring the rest.
- **Partial document matching** applies DLP policies to partial content of a file that may be protected. For example, the DLP will flag when a user of application copies a few lines from a large, sensitive document to email communication.
- **Exact file matching** generates a file hash and compares it against other file fingerprints. This is particularly useful for media or other content files.

- **Conceptual/lexicon** is useful against unstructured ideas and communications that violate the company policy. For example, you can train the DLP solution to identify communications that may indicate insider trading or running a business using a corporate email domain.
- **Statistical analysis,** including advanced machine learning or simple statistical methods such as linear regression, can be used based on the complexity of the content analysis problem.
- **Categories** applies rules to specific categories such as health data, credit card details, or other classified information.

Components of a DLP solution

Now that we understand how data loss prevention solutions work, we can sum up the key data security capabilities that typical DLP solutions offer:

- **Monitoring** provides visibility into data assets and network traffic and identifies anomalous behavior that may require security actions, based on further analysis and predefined data security policies.
- **Filtering** refines data sets to prevent disclosure of confidential or sensitive information to unauthorized users. You can design different patterns of data filtering to filter data at various levels of access privileges.
- **Reporting** provides findings on past performance and contextual data visualizations to help make well-informed decisions on information and network security.
- **Analysis** uses data processing and analytics powered by advanced AI/ML tools to extract hidden insights into network traffic and user access behavior.

Examples of DLP services

Here are some common data loss prevention services and solutions:

- McAfee Total Protection for DLP
- Digital Guardian Endpoint DLP
- Symantec Data Loss Prevention
- SecureTrust Data Loss Prevention
- Check Point Data Loss Prevention

Related reading

- BMC Business of IT Blog
- BMC Machine Learning & Big Data Blog
- Introduction to Enterprise Security
- Data Management vs Data Governance: Main differences
- What Is Data Gravity?