Data loss prevention software

Data loss prevention (DLP) software detects potential data breaches/data ex-filtration transmissions and prevents them by monitoring,[1] detecting and blocking sensitive data while *in use* (endpoint actions), *in motion* (network traffic), and *at rest* (data storage).[2]

The terms "data loss" and "data leak" are related and are often used interchangeably.[3] Data loss incidents turn into data leak incidents in cases where media containing sensitive information is lost and subsequently acquired by an unauthorized party. However, a data leak is possible without losing the data on the originating side. Other terms associated with data leakage prevention are information leak detection and prevention (ILDP), information leak prevention (ILP), content monitoring and filtering (CMF), information protection and control (IPC) and extrusion prevention system (EPS), as opposed to intrusion prevention system.

Categories

The technological means employed for dealing with data leakage incidents can be divided into categories: standard security measures, advanced/intelligent security measures, access control and encryption and designated DLP systems, although only the latter category are currently thought of as DLP today.[4]

Standard measures

Standard security measures, such as firewalls, intrusion detection systems (IDSs) and antivirus software, are commonly available products that guard computers against outsider and insider attacks. [5] The use of a firewall, for example, prevents the access of outsiders to the internal network and an intrusion detection system detects intrusion attempts by outsiders. Inside attacks can be averted through antivirus scans that detect Trojan horses that send confidential information, and by the use of thin clients that operate in a client-server architecture with no personal or sensitive data stored on a client device.

Advanced measures

Advanced security measures employ machine learning and temporal reasoning algorithms to detect abnormal access to data (e.g., databases or information retrieval systems) or abnormal email exchange, honeypots for detecting authorized personnel with malicious intentions and activity-based verification (e.g., recognition of keystroke dynamics) and user activity monitoring for detecting abnormal data access.

Designated DLP systems

Designated systems detect and prevent unauthorized attempts to copy or send sensitive data, intentionally or unintentionally, mainly by personnel who are authorized to access the sensitive information. In order to classify certain information as sensitive, these use mechanisms, such as exact data matching, structured data fingerprinting, statistical methods, rule and regular expression matching, published lexicons, conceptual definitions, keywords and contextual information such as the source of the data.[6]

Types

Network

Network (data in motion) technology is typically installed at network egress points near the perimeter. It analyzes network traffic to detect sensitive data that is being sent in violation of information security policies. Multiple security control points may report activity to be analyzed by a central management server.[3]

Endpoint

Endpoint (data in use) systems run on internal end-user workstations or servers. Like network-based systems, endpoint-based technology can address internal as well as external communications. It can therefore be used to control information flow between groups or types of users (e.g. 'Chinese walls'). They can also control email and Instant Messaging communications before they reach the corporate archive, such that a blocked communication (i.e., one that was never sent, and therefore not subject to retention rules) will not be identified in a subsequent legal discovery situation. Endpoint systems have the advantage that they can monitor and control access to physical devices (such as mobile devices with data storage capabilities) and in some cases can access information before it is encrypted. Endpoint systems also have access to the information needed to provide contextual classification; for example the source or author generating content. Some endpoint-based systems provide application controls to block attempted transmissions of confidential information and provide immediate user feedback. They must be installed on every workstation in the network (typically via a DLP Agent), cannot be used on mobile devices (e.g., cell phones and PDAs) or where they cannot be practically installed (for example on a workstation in an Internet café).[7]

Data identification

DLP includes techniques for identifying confidential or sensitive information. Sometimes confused with discovery, data identification is a process by which organizations use a DLP technology to determine what to look for.

Data is classified as either structured or unstructured. Structured data resides in fixed fields within a file such as a spreadsheet, while unstructured data refers to free-form text or media in text documents, PDF files and video.[8] An estimated 80% of all data is unstructured and 20% structured.[9]

Data loss protection (DLP)

Sometimes a data distributor inadvertently or advertently gives sensitive data to one or more third parties, or uses it themselves in an authorized fashion. Sometime later, some of the data is found in an unauthorized place (e.g., on the web or on a user's laptop). The distributor must then investigate the source of the loss.

Data at rest

"Data at rest" specifically refers to information that is not moving, i.e. that exists in a database or a file share. This information is of great concern to businesses and government institutions simply because the longer data is left unused in storage, the more likely it might be retrieved by unauthorized individuals. Protecting such data involves methods such as access control, data encryption and data retention policies.[3]

Data in use

"Data in use" refers to data that the user is currently interacting with. DLP systems that protect data in-use may monitor and flag unauthorized activities.[3] These activities include screen-capture, copy/paste, print and fax operations involving sensitive data. It can be intentional or unintentional attempts to transmit sensitive data over communication channels.

Data in motion

"Data in motion" is data that is traversing through a network to an endpoint. Networks can be internal or external. DLP systems that protect data in-motion monitor sensitive data traveling across a network through various communication channels.[3]

See also

- List of backup software
- Metadata removal tool
- Endpoint detection and response
- Endpoint security

References

1. *Hayes, Read (2007), "Data Analysis", Retail Security and Loss Prevention, Palgrave Macmillan UK, pp. 137–143, doi:10.1057/9780230598546_9 (https://doi.org/10.1057%2F9780230598546_9) , ISBN 978-1-349-28260-9*
2. *"What is Data Loss Prevention (DLP)? A Definition of Data Loss Prevention" (https://digitalguardian.com/blog/what-data-loss-prevention-dlp-definition-data-loss-prevention) . Digital Guardian. 2020-10-01. Retrieved 2020-12-05.*

3. *Asaf Shabtai, Yuval Elovici, Lior Rokach, A Survey of Data Leakage Detection and Prevention Solutions (https://www.springer.com/computer/security+and+cryptology/book/978-1-4614-2052-1) , Springer-Verlag New York Incorporated, 2012*
4. *Phua, C., Protecting organisations from personal data breaches (http://www.sciencedirect.com/science/article/pii/S1361372309700119) , Computer Fraud and Security, 1:13-18, 2009*
5. *BlogPoster (2021-05-13). "Standard vs Advanced Data Loss Prevention (DLP) Measures: What's the Difference" (https://logixconsulting.com/2021/05/13/standard-vs-advanced-data-loss-prevention-dlp-measures-whats-the-difference/) . Logix Consulting Managed IT Support Services Seattle. Retrieved 2022-08-28.*
6. *Ouellet, E., Magic Quadrant for Content-Aware Data Loss Prevention, Technical Report, RA4 06242010, Gartner RAS Core Research, 2012*
7. *"Group Test: DLP" (https://info.digitalguardian.com/rs/768-OQW-145/images/SC-Labs-DLP-GROUP-TEST-AND-DG-REVIEW.pdf?field_resource_type_value=analyst-reports) (PDF). SC Magazine. March 2020. Retrieved September 7, 2021.*
8. *"unstructured data Definition from PC Magazine Encyclopedia" (https://www.pcmag.com/encyclopedia/term/53486/unstructured-data) .*
9. *Brian E. Burke, "Information Protection and Control survey: Data Loss Prevention and Encryption trends," IDC, May 2008*

Last edited 23 days ago by Aoidh

WIKIPEDIA