



A Buyer's Guide to Data Loss Protection Solutions

Table of Contents

Policy Definition.....4

Detection4

Incident Response.....5

Role-Based Administration6

Network Monitoring6

Web Security7

Email Security7

Data Discovery.....8

Endpoint8

A Buyer's Guide to Data Loss Protection Solutions

Data loss prevention (DLP) solutions are designed to prevent the inadvertent loss or deliberate removal of sensitive and confidential information. Organizations commonly use them to secure communication channels, such as email, to ensure employees do not send sensitive information to unauthorized recipients. However with the increasing use of Web 2.0 applications and a growing mobile workforce conducting business on portable devices like laptops (with their potential for theft and loss), effective DLP solutions must be able to provide coverage for a wide range of communication channels.

To achieve this goal, a DLP solution must, at a minimum, include Web, email, and endpoints or laptops among the communication channels it can secure. If necessary, it must also be able to block transmission of data on these channels. Managing separate policies for each of these channels can quickly become cumbersome. A DLP solution should be able to provide policy management and reporting capabilities that administrators can easily extend to support several channels from a single policy.

One of the most common drivers for deployment of a DLP solution is the need to meet regulatory compliance. Organizations with this need will benefit greatly from a solution that comes with pre-built templates covering various regulations. The solution should also be customizable, so that organizations can tailor their built-in policy based on their specific regional or industry requirements.

This guide provides a list of recommended features and capabilities for buyers to consider when purchasing a DLP solution. Buyers can also use this guide to develop requests for proposals, as well as to help differentiate one vendor's products from another during demonstrations and proof-of-concepts.

Policy Definition

An effective DLP solution makes it easy for organizations to define their DLP policies, so they can more easily meet their compliance requirements.¹ Here are some capabilities to consider:

- Policy types that include not just keywords and regular expressions but also dictionaries and the ability to articulate context as well as content (e.g., when a name is found near a valid credit card number) for greater accuracy.
- Built-in policies for multiple industries and geographies that users can access, use, and apply simultaneously (e.g., health care and Insurance policies) to facilitate compliance.
- A single policy to scan data wherever it is stored, transmitted, or used, both on the network and on the endpoint, ensuring consistent coverage.
- A centralized interface for policy editing and policy management, across all components (across monitoring and prevention and across network and endpoint), which simplifies and streamlines administration.
- The ability to define policies based on any combination of the following: content, sender/recipient, file characteristics, communications protocol, and destination category, depending on an organization's specific needs, for greater visibility and control.
- Configurable scoring of incident severity based on content identifiers, such as file type, file size, and keywords for flexible incident management capability adaptable to individual needs.
- Inclusion and exclusion detection rules based on corporate directory data to enforce policy based on the senders and recipient/destination.
- Predefined detection policies to cover regulations and detection best practices, including pre-defined lexicons for commonly required regulations.
- Predefined content classifiers that users can combine to make new policies.

Detection

The importance of detection accuracy in a DLP solution cannot be overstated. If an organization does not have a high level of confidence regarding detection accuracy, its IT resources cannot focus on genuine incidents and will be easily overwhelmed by the large number of false positive and negative incidents. Capabilities to consider include:

- Identical detection capabilities across all threats covered (e.g., for both network and endpoint-based products, and for both data monitoring and prevention and data discovery and protection) to provide consistent policy enforcements.
- The ability to extract and inspect the text content of files and attachments for better visibility into your data.
- Detection capabilities that can support contents written in various languages and language types including Western European and Asian (Japanese, Chinese simplified, Chinese traditional, Cyrillic, and Korean).

¹ This Guide is not intended to provide legal guidance on regulatory compliance. If you have questions about the meaning of a particular provision of a regulation, you should consult your attorney. The Websense® data loss prevention product suite is a tool that people can use to help them comply with these requirements, but it is not a regulatory compliance product and will not guarantee that a given business practice is compliant with particular laws or regulations.

- The ability to recursively inspect the contents of compressed (e.g., a file that is ZIP-ed, TAR-ed, and then RAR-ed) archives and detect against fingerprinted content.
- A method for fingerprinting data such as customer records and validating the accuracy of a fingerprint at the time of its creation.
- The ability to detect fingerprinted data match on specific fields (e.g., only first name and last name from a customer database), without needing a pattern-based number (e.g., Social Security number, credit card number) to be present.
- A method of detecting fingerprinted documents that supports detection of the same text or portions of text in different file formats. For example, if a fingerprinted document is in Microsoft Word format, detects that same text that has been cut and pasted directly into an email in plain text.
- Content detection using fully customizable rules with keywords and key phrases, as well as detection for pattern matching combined with validations specific to the content being detected. For example, detects common credit card number patterns as well as doing the checksum validation to ensure a valid credit card number (the "Luhn" check).
- The ability to distinguish between different types of personally identifiable information (PII) or personal health information (PHI) numbers, such as distinguishing a customer's nine-digit Social Security number from a nine-digit phone number without the presence of a keyword (e.g., "SSN") or delimiters (e.g. , 985596761 or 985 59 6761 instead of 985-59-6761) for improved accuracy.
- Content matching of specific documents such as source code, specific paragraphs, design documents, marketing documents, or financials and support detection of derivative or cut-and-paste versions of content matching specific documents.

Incident Response

It is critical for compliance and audit-related requirements that an organization be able to quickly respond to a data loss incident. To do so, recommended capabilities include:

- The ability to view confidential data loss events via the Web in a format usable by non-IT business level users to reduce the strain on IT staff and empower other users to manage data loss incidents.
- A clear indication in the incident report of how the transmission or file violated policy (not just which policy was violated), including clear identification of which content motivated the match for greater accuracy in reporting and to improve processes that could prevent similar violations from occurring.
- The ability to view identity information on the sender (e.g., full name, manager name, business unit) and destination of the transmission (e.g., data sent to a blog, chat board, spyware site) to facilitate remediation.
- The ability to assign each user in the workflow for the remediation of a certain set of incidents so that the appropriate person is handling the incident.

- Automated notifications to designated incident manager(s) when they have new incidents to review.
- The ability to define and track a "case" or set of incidents users find to be related after an investigation for improved reporting and remediation.
- The ability to easily export a group of incidents from the system in a format that's readable by a person without system access (e.g., a PDF).
- The ability to add customized attributes to incidents to correlate with a unique remediation business process.

Role-Based Administration

Role-based administration in a DLP solution allows the management of various incidents without exposing sensitive information to unauthorized users, which can often exacerbate the problem. Only the designated administrator, based on policy, business unit, and other factors should be granted access to specific incident details. Capabilities to consider include:

- Control of incident access based on role and policy violated to ensure only an authorized administrator for the policy is managing the specific incident.
- Control of incident access based on business units or groups to ensure only the authorized administrator for the specific business unit or group is managing the incident.
- The ability to define a role to restrict viewing rights to identity-based information.
- The ability to define a role to restrict viewing rights to content of the message that violated policy.
- The ability to create separate roles for technical administration of servers, user administration, policy creation and editing, incident remediation, and incident viewing for data at rest, in motion, or at the endpoint.

Network Monitoring

Network monitoring capability for DLP solutions enables the inspection of traffic traversing the network providing much needed visibility into the types of network data. This visibility is crucial for analysis and reporting, as well as for creating accurate and relevant policies. Some capabilities to consider include:

- Notification of unprocessed traffic due to network bursts (e.g., dropped packets or sampling).
- The ability to monitor Web traffic, such as webmail, Web postings, and other protocols using HTTP and HTTPS including uploaded files.
- The ability to monitor and, when appropriate, prevent network printing of confidential information.
- Geographical and website detail to resolve/classify the destination of HTTP and HTTPS transmission beyond just an IP address.
- Native inspection of SSL communications.

- The ability to monitor both active and passive FTP traffic including fully correlating transferred file data with control information.
- The ability to detect multiple incidents over time. A slow leak over time can result in significant data loss.
- The ability to monitor network traffic on arbitrary ports or port ranges to deal with unclassified or rogue threats.
- The ability to resolve the identification of the offending user in real time (not just IP address and not just LDAP lookup after the event).
- The ability to operate without depending on a third-party proxy to enforce Web traffic including SSL traffic.

Web Security

More and more Web attacks are intended to steal sensitive and confidential data making tight integration of Web security and data loss prevention critical in current and future DLP solutions. Some of the capabilities to consider include:

- Support of content-aware blocking of network transmissions over HTTP natively and the ability to provide notification.
- Visibility into the type of site data is posted to and its geographical location.
- The ability to monitor and block network transmissions over FTP.
- Control of the latency the solution introduces to normal network communications.
- The ability to prevent data loss on Web 2.0 (AJAX based) sites that dynamically update content.
- The ability to monitor and control Web traffic including HTTPS without an external proxy.
- Support of ICAP for interoperability with third-party proxy as necessary for different deployment scenarios.

Email Security

Emails are one of the simplest methods of transmitting potentially sensitive and confidential information, which could occur in the body of the email as text or as attachments. Email DLP capabilities to consider include:

- The ability to block outbound emails that are in violation of company policy on confidential data.
- The ability to monitor and enforce for internal email traffic, including attachments.
- The utilization of either its own message transfer agent or another means of email prevention.
- The ability to quarantine emails that are in violation of company policy on confidential data.
- The ability to automatically encrypt emails based on company policy settings.
- The ability to take preventive actions without introducing another "hop" in the outbound message chain.

- The ability to release email from quarantine by end-users, their managers, or other designated users.
- The ability to ensure message delivery even in the event of a failure of your system.
- The ability to notify senders and security administrators of a blocked or quarantined email.

Data Discovery

Uncovering the location of sensitive data in the various servers, databases, endpoints (laptops), and other locations identifies whether data are stored in accordance with an organization's policies. Data discovery also continually monitors storage of sensitive data and if necessary, removes or encrypts them. Capabilities to consider include:

- The ability to scan Windows file servers, desktops, and laptops.
- The ability to scan locations such as custom repositories and support full reporting on policy violations found in those repositories.
- The ability to automatically move or remove files which violate policy.
- The ability to automatically quarantine and delete files which violate policy.
- The ability to scan to Inform file owners about quarantined files, including details of why the file was quarantined, such as which policy it violated.
- Integration with corporate directories to allow data-at-rest policy violations to be associated with a particular individual and business unit.
- Providing a single report covering data at rest (storage) throughout the global enterprise.
- Providing a single management interface for all scan configuration and control, enterprise-wide.
- The ability to scan remote locations with low network bandwidth.

Endpoint

Endpoint or laptops warrant special consideration for DLP solutions because laptops contain multiple methods of transferring data, and users can also store data in the internal hard drive, Removable media such as USB and CD drives as well as direct or network printing capabilities require special enforcement scenarios to prevent potential data loss. Capabilities to consider include:

- The ability to detect user attempts to copy confidential data to removable storage devices (e.g., USB drives, floppy, CD/DVD).
- The ability to monitor/prevent cut/copy, paste, print screen, file access, print to local printer, and print to network printer.
- The ability to perform the actions above based on a specific application (e.g., cut/copy is not permitted from Excel); content in use from a specific application (e.g., cut/copy is not permitted from Excel when displaying confidential data); application category (e.g., all spreadsheet applications).
- Automatic prompts of external security controls (e.g., file encryption).

- Enforcement of different policies when the endpoint is connected to the trusted corporate network and when it is connected to public network that cannot be trusted (e.g., airport, coffee shop, home network).
- Protection of confidential content regardless of file type or file location (e.g., distinguishes between an Excel document with confidential data that must be protected versus an Excel document without confidential data which should not be protected).
- Support of detection based on fingerprinting of content.
- The ability to perform detection locally, avoiding the need for network connection or to transmit potentially sensitive data over such connection.
- The ability to define policies once and applies them for both network (agent-less) and agent-based discovery in a centralized management interface.
- The ability to scale as additional endpoints are deployed.
- Support of geographically dispersed machines for global deployments of endpoint agents while maintaining a central management/reporting interface.
- The ability to display complete details about the incident including the file name, user information, policy match details, and a copy of the original file that violated policy.
- The ability to operate without requiring much system resources, including CPU, disk, and memory footprint or requiring third-party management tools.
- Native encryption support and key management for protected content copied to USB devices.