Ford's Notes: Data Loss Prevention (DLP) Solutions

Data Loss Prevention (DLP) solutions are used to detect and prevent sensitive information from being stored on unauthorized systems or from being transmitted over unauthorized networks.

There are three components in most DLP solutions:

- a policy server,
- an endpoint agent, and
- a network agent.

The policy server is used to configure the solution and develop rule sets that are used to detect data loss based on data classification, confidentiality, or privacy level. The policy server is where logging of system behavior and detections is configured.

DLP rules can be based on detections such as:

- document classification,
- files containing certain dictionary words,
- policy templates based file attributes such as permissions, dates, or location in the file system,
- exact data match (EDM) to character patterns,
- and other techniques.

A DLP endpoint agent is used to enforce the rule sets and policies on a specific client computer, even if those computers are no longer connected to your corporate network.

A DLP network agent is a specifically configured network appliance that's placed at the network boundary, and it's going to be used to scan different web, email, and messaging protocols as the messages attempt to leave the network.

Using a DLP system, an organization is going to be able to block any data that doesn't conform to a predetermined and acceptable policy. When a DLP solution applies rules which when successful allow data to pass that process is known as whitelisting. An example of a whitelist rule would be allowing any Microsoft Word file that has the word 'Public' in the footer to pass. When a DLP solution is configured with a blacklist when the rules block data that matches a predetermined sets of conditions. DLP solutions may also be configured to block when data fails to match any of the configured rule conditions.

DLP rules can be configured with various actions. These rule actions are carried out by either a DLP endpoint or network agent. Common DLP rule behaviours are:

- Alert
- Block
- Quarantine
- Tombstone

If a DLP rule is set to alert only, the agent is going to allow the data to still transmit and go on its way to its destination but the action will be logged and an administrator possibly alerted when that happens.

Ford's Notes: Data Loss Prevention (DLP) Solutions

A DLP rule can be configured so that the agent blocks data when it finds a match in the rule set. This would actually stop the user from being able to copy that file from a sharedrive to the external hard drive. In this case, though, the user could still open the file and read it from the sharedrive. Since they're not allowed to copy it, they could actually see it on the screen and maybe pull out their cell phone and take a picture of it.

Using a quarantine rule, the DLP system will block the user from copying that file and then take away the user's access to even read or open that file. This is because the system is under the assumption that if you tried to copy that file and failed, you might try to do something else, like open the file and print it, or open it and copy and paste it into an email or some other mechanism to move the contents of that protected file off of the system and out of the network. To quarantine the file, many DLPs will simply encrypt the file and turn it into ciphertext that the end user can't read or comprehend.

The fourth action a DLP system can take is to tombstone the file. When executing this action the original file is going to be moved or encrypted on the sharedrive and it's going to get replaced by a new file that simply contains a message that states there was a policy violation that has occurred.

Many DLP solutions have evolved and have features that address ways of moving data other than over a network. If removable or external media blocking is enabled, the DLP system is going to prevent the user from being able to read or write to an external device, such as a CD, a DVD, a USB, or other external storage device like a flash drive or a hard drive.

If print blocking is enabled, the DLP system will block the ability for a user to print to a networked or USB connected printer.

If RDP blocking is enabled, the DLP system will prevent the user from conducting a copy and paste between the remote client PC that they're connected to over the Remote Desktop Protocol and their normal host machine.

A popular DLP solution is Proofpoint DLP (https://www.proofpoint.com/us/products/information-protection/enterprise-dlp). Proofpoint calls configured rule policies 'Detectors'.

Proofpoint Smart Identifiers (Smart IDs) combine regular expressions (regex) with additional code and are used by Detectors to locate potentially sensitive data in files. Smart IDs reduce the number of false positive identifications by identifying not only the number of characters in an expression, but also patterns with logic and connections between the characters using checksums and other algorithms.

Proofpoint DLP Dictionaries contain lists of terms used by Detectors to locate potentially sensitive data in files. When a file is scanned, a Detector compares all words and phrases in the file against all Dictionary terms in the enabled Dictionaries.

Proofpoint Indexed Document Matching (IDM) improves detection capabilities to protect sensitive data. IDM indexes unstructured text in files and DLP Detectors detect when an end user tries to exfiltrate part or all of the text or the entire file. The match is based on the percentage defined within the DLP detector.