Beyond data loss prevention

Prepared by:
William Burlend, Security Consultant
# Table of contents

1. Introduction 3

2. Risks to corporate data 4

3. Current solutions 7

4. Integrating with cloud-based solutions 12

5. The next frontier 14

6. Conclusion 16

7. References & further reading 17

8. About NCC Group 19
1. Introduction

Data Loss Prevention (DLP) is a security control aimed at detecting sensitive data on a corporate network and highlighting when it leaves the network, or is accessed without authorisation. DLP tools can take many forms including endpoint security, email, cloud-based solutions and mobile device management software. DLP implementations can be seen by some employees and departments as a blocker or as a delay to their everyday work functions and responsibilities. However, when configured correctly a DLP solution can be a great asset to a business and support a range of security goals and compliance, such as PCI and General Data Protection Regulation (GDPR).

The diagram below provides an overview of some of the different types of DLP solutions available and the devices/applications that they seek to protect from potential data loss.

This whitepaper aims to discuss the various benefits and pitfalls of DLP solutions currently available. It will also address how DLP can be integrated with cloud providers, given the ever-increasing demand to place data in the cloud.

![Diagram detailing the split of devices and general function of a DLP solution](image_url)

*Figure 1: Diagram detailing the split of devices and general function of a DLP solution [1]*
2. Risks to corporate data

As businesses evolve their working environments and support their IT infrastructure through the introduction of new and exciting technologies, the risk to data loss (both customer and business) increases.

In a typical office there are many ways in which confidential data might be leaked, including:

- Social media networks
- Corporate instant messaging
- Use of smart phones to take photos and record audio/video
- Misconfigured CCTV and network devices
- Printers
- Cloud storage
- Personal email
- Smart watches
- Internet of Things (IoT) devices

2.1 Examples

Seemingly innocuous devices, for example printers, can hold a wealth of information and therefore need to be taken into consideration when devising a DLP implementation on a corporate network. Some of the risks that can be taken include increased print jobs from employees leaving the organisation, indicating potential intellectual property theft. Also sensitive data could remain on the printer’s internal hard drive when the printer is either scrapped, sold or sent away for maintenance. If they are misconfigured and exposed to the Internet, they can become victim to a malicious attack which could result in an unauthorised person being able to copy the data being sent to the printer from outside of the network.
Data loss software can be installed on endpoints to prevent sensitive documents such as M&A, intellectual property or client data from being printed. It is also important for effective DLP solutions to have a policy which takes standard deviation into account. For example, if an employee hands in their notice and they suddenly increase their printing from approximately 20 to 80 documents a week, this could be flagged for investigation through the use of a DLP solution.

Removable media such as USB devices, CD-ROMs, mobile phones and smart watches can be used to take information from a network. Therefore, controls are required to mitigate these risks. There have been recent examples of data being taken from a network without authorisation, resulting in a serious data breach. One example is a Europol employee who copied a 700 page report on terrorist groups to a USB device. This device was then moved to a personal storage device connected to their home network. Unfortunately, this device was misconfigured and was actually exposed to the Internet without any form of authentication [3].

Not all of these types of risk can be mitigated through use of DLP tools alone. However, combined with user education, well-written corporate policies, a good network security posture and an awareness of laws and regulations governing the information stored, the majority of data leakages are preventable.
### 2.2 When DLP could have saved the day

There have been a number of instances where a DLP solution could have prevented a data leak and therefore, the subsequent fine from the Information Commissioner’s Office (ICO). One such example involves a County Council fined £120,000 after staff emailed files containing sensitive personal data to the wrong recipients [4]. If a DLP policy had been configured to block emails which included personal data, this incident could have been avoided and saved the County Council from a substantial fine and significant reputational damage.

The table below shows fines issued by the ICO relating to email errors since 2011. In all instances email errors or abuse were the conduit for data loss or exposure [5]:

<table>
<thead>
<tr>
<th>Data controller</th>
<th>Sector</th>
<th>Nature</th>
<th>Date of final notice</th>
<th>Fine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surrey County Council</td>
<td>Local government</td>
<td>Various incidents concerning the disclosure of personal data via email to incorrect recipients</td>
<td>06/06/2011</td>
<td>£120,000</td>
</tr>
<tr>
<td>North Somerset Council</td>
<td>Local government</td>
<td>Various incidents concerning the disclosure of personal data via email to incorrect recipients</td>
<td>09/11/2011</td>
<td>£60,000</td>
</tr>
<tr>
<td>Worcestershire County Council</td>
<td>Local government</td>
<td>Disclosure of personal data via email to incorrect recipients</td>
<td>17/11/2011</td>
<td>£80,000</td>
</tr>
<tr>
<td>Cheshire East Council</td>
<td>Local government</td>
<td>Disclosure of personal data via email to unintended recipients</td>
<td>08/02/2012</td>
<td>£80,000</td>
</tr>
<tr>
<td>Stoke on Trent City Council</td>
<td>Local government</td>
<td>Disclosure of personal data via email to incorrect recipient</td>
<td>23/10/2012</td>
<td>£120,000</td>
</tr>
<tr>
<td>Ministry of Justice</td>
<td>Central government</td>
<td>Spreadsheets showing prisoners' details emailed to members of the public</td>
<td>22/10/2013</td>
<td>£140,000</td>
</tr>
<tr>
<td>Bloomsbury Patient Network</td>
<td>Health</td>
<td>Inadvertently revealed the identities of HIV patients through an email error</td>
<td>11/12/2015</td>
<td>£250.00</td>
</tr>
<tr>
<td>Telegraph Media Group Ltd</td>
<td>Media</td>
<td>Sending hundreds of thousands of emails urging readers to vote Conservative in the general election</td>
<td>15/12/2015</td>
<td>£30,000</td>
</tr>
<tr>
<td>Chelsea &amp; Westminster Hospital NHS Foundation Trust</td>
<td>Health</td>
<td>A health trust revealed the email addresses of more than 700 users of a HIV service</td>
<td>04/05/2016</td>
<td>£180,000</td>
</tr>
<tr>
<td>Chief Constable of Dyfed Powys Police</td>
<td>Criminal justice</td>
<td>An email, containing information that could be used to identify eight sex offenders, was sent to a member of the public in error</td>
<td>02/06/2016</td>
<td>£150,000</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td><strong>£960,250</strong></td>
</tr>
</tbody>
</table>
3. Current solutions

Today, there are a number of DLP products available on the market which can cater for business’ unique network architectures. Some of the most widely used solutions can be seen on the image below and include Symantec DLP, Digital Guardian and Forcepoint. These have a host of options to refine policies and install on multiple device types. There are also some key advantages to choosing a market leader, mostly down to their budgets and higher staff headcount. Benefits often include being more feature rich and bug fixes being released in shorter time frames.

However, some organisations prefer a more personal approach and by going with a more niche company there might be more room for product customisation to better fit with the network it is to be installed on. It might also be that the company focuses specifically on only one aspect, such as Bring Your Own Device (BYOD) DLP and this is all that the client is looking to achieve.

Figure 3: Magic Quadrant for Enterprise Data Loss Prevention [6]
3.1 Advantages of DLP & its business impact

DLP software can be a great asset to a business in many ways. Depending on the sector of the organisation, DLP tools can be honed to help support Business as Usual (BAU) tasks. This can include monitoring network activity, preventing data from being leaked out of the network perimeter, supporting internal investigation teams or helping to obtain a regulatory or compliance certification.

An example of this could be monitoring insider threats. The information security team, in collaboration with various business stakeholders, could detail the number of customer records a member of the call centre staff typically accesses in a standard shift. Based on this information, a standard deviation rule could be configured that triggers an alert when a user accesses ten per cent above the agreed number of records. This would then detail the staff member, file paths accessed and the times they were viewed. Depending on the type of organisation, this may then be passed to the internal investigations team as a form of evidence and may also form part of an ongoing assessment into staff activity.

Data loss tools can also be used to support regulatory goals for the company. For example an organisation may be working towards PCI-DSS compliance. DLP can be used to passively scan various network segments to analyse the type of information being stored, such as credit card details, addresses, dates of births and other Personally Identifiable Information (PII). Armed with this knowledge, network teams can go about implementing effective network segmentation and ensuring sensitive information is removed to obtain compliance and regulatory sign-off.

Projects such as Identity Access Management (IAM), whereby user roles are defined along with their associated access, can utilise DLP data. This could include tracking the type of information a user accesses on a daily basis or online services which require some form of file upload process. The majority of DLP solutions include a dashboard which can be used to generate custom reports and this feature can be extremely useful when trying to uncover shadow IT within the business. By filtering on some of the most used email or cloud storage providers, it might be possible to uncover poor business processes which have been developed to speed up everyday tasks. This information can then be used, along with the affected business unit to come up with a more secure solution to the problems they face.

Employee education can also be greatly assisted through the use of a DLP solution as most support alerts which advise users that what they are attempting to do may breach company policy. For example, if you were to upload a document which contains details of a corporate patent to personal cloud storage, you could receive a warning pop-up advising that this action is not permitted with an explanation. You could then choose to either ignore the alert, mark as a false positive and include a comment or abide to the alert and prevent the document from leaving the network. Over time, this would also re-educate users to the types of data that is acceptable to send and upload.
Managing removable media on corporate networks can become a contentious issue. There has to be a trade-off between security and usability with a lot of solutions either using a whitelist approach, software blocking or physically disabling the ports.

Group policy can be used to perform a blanket block on all ports as a quick and effective fix to the risk they pose. However, there are many valid reasons why this would not be appropriate.

Another method is to use a whitelisting approach which requires devices to be approved by IT before being plugged into a machine. This works well with DLP software as the whitelisted device can then be monitored for the types of files that are uploaded. Physical devices exist which can be placed over ports to lock them down, preventing storage devices from being inserted into the machine. Another variant of physical port lockdown is filling the ports with an epoxy resin which when hardened, physically protects those ports from physical device connection. However, this is not very practical for a large organisation as the number of port locks/blocks required could run into thousands.

One of the major drawbacks of any DLP solution is that in order for it to be effective, it needs to be able to decrypt encrypted data in order to inspect its content and nature. While it isn’t possible to view the content of an encrypted file which is being copied off the network (or attached to an email), a policy can, however, be set to detect encrypted data and set it to either block or raise an alert. For policies that require scanning of all content, consideration will be required where encrypted streams are terminated. This may require installation of encryption (SSL) keys and certificates on DLP scanning devices so that the encrypted streams can be terminated and the data inspected. This will require careful key management and a robust configuration to ensure that stream encryption is not inadvertently weakened or exposed through the introduction of DLP solutions.

Thin clients can be a good compromise in order to reduce the costs associated with DLP technology. If an organisation has end-user terminals that do not store data, have input ports or any method to transfer data (other than via a web browser and email), this could be a good solution to secure a sensitive environment [8]. All data and applications reside on a server that is typically hardened and maintained in a secure area of the network. This option would certainly not be for every organisation as there are a multitude of reasons why accessing and transferring data via removable media is required. However, it might be appropriate for highly sensitive areas of the network, such as card data environments, or for teams that analyse protectively-marked documents.
3.2 Why solutions can fail

Understanding what the business classes as sensitive data is one of the most difficult aspects to get right when configuring DLP policies. There may be differences of opinion between the information that the security team deem to be important compared with the various business units. If this issue is not addressed then two scenarios may occur:

1. The DLP policies are too broad and may trigger a vast number of security incidents that need to be manually reviewed. A significant proportion of the incidents would be false positives.

2. Polices configured are too relaxed and sensitive information is not picked up. This may result in a data breach along with potential reputational damage or regulatory fines and investigations.

To overcome these issues, business engagement and a detailed breakdown of what each department wants to protect is required. Based on the outcome of these meetings, policies can be set up to focus on capturing specific file types or taking a hash of a document and checking files for matches.

Alternatively, if time and funding allow, a data classification project can be undertaken. This involves going through the data stored and generated by the business and modifying it to include metadata to help with tagging. This can then be integrated in DLP polices in order to search for data which matches the various tags. There are a number of benefits to undertaking this type of project, such as removing duplicate files and subsequently freeing up storage space on the network. Defining data owners can also help to ensure the process is rolled out systematically throughout the company while also facilitating GDPR compliance, due to having known and defined data controllers.

In terms of the information security team managing the DLP incidents, segregation of duties needs to be enforced. Failing to consider an insider threat could jeopardise the whole implementation of DLP within the organisation. One recent study stated that “26 per cent of employees admitted to uploading sensitive information to cloud apps with the specific intent to share that data outside the company” [9]. If an information security analyst has the ability to dismiss incidents that are generated by themselves this could bypass all DLP controls. The manager of the team should also receive automated weekly results of the top ten worst offenders (with regards to employees triggering alerts) and also destinations or methods of the incidents. This enables transparency within the team and should mitigate any such occurrences of misconduct.

Depending on the size of the organisation and the number of DLP policies in force, a significant number of DLP incidents can be generated. The time spent by information security staff remediating and categorising them correctly can be huge. Manually inspecting thousands of incidents to determine if they are false positives or require further investigation is not time well spent and could be put to better use. Another disadvantage of poor policies is the amount of storage required to quarantine incidents.

BYOD is another major issue that faces both DLP systems and organisations. If users are accessing corporate data on their personal devices it should fall under the scope for DLP. However, some users may have concerns around privacy as the company could end up viewing personal, non-work related content.

Other data transport mediums such as Bluetooth, infrared and Near Field Communication (NFC) are not currently natively supported by DLP solutions. This does present a small data risk, though the number of business cases that would require sensitive data to be transferred via these channels will
be relatively low. It could therefore be reviewed on a case by case basis. Simple lockdowns of these interfaces through group policy could minimise the potential risk of data loss through wireless networks, however, as endpoint devices become ever more reliant on wireless technology for networking, there is a need for DLP wireless solutions to facilitate assurance in this space.

3.2.1 A web of encryption

The Internet is increasingly using more secure methods to serve content and in recent years there has been a surge in sites utilising HTTPS over the insecure plain-text version HTTP. This is partly as a result of a number of high profile attacks against organisations as well as various government whistle-blowers. According to data supplied by Mozilla [10] and Google [11], HTTPS traffic now makes up more than half of the traffic seen online. Apart from this being a great step forward for online security and the communities that make use of the services, it does pose challenges for companies want to ensure their DLP systems can view the raw data and check for sensitive documents before allowing it to leave their estate. A number of tools and physical devices have been developed to help resolve this issue and are discussed further in the section 4 ‘Integrating with cloud-based solutions’.
4. Integrating with cloud-based solutions

In recent years, there has been a big shift in moving corporate data away from local hardware and onto Software-as-a-Service (SaaS) environments such as Box, Google Drive and Dropbox. A study of 1,060 professionals across a range of different organisations found that 89 per cent used some form of public cloud storage [12] as part of their company’s infrastructure. With this in mind, the standard on-premises DLP solutions can often no longer fulfil the business requirements, such as trying to keep track of access and retain compliance certifications.

Shadow IT is a significant problem that faces organisations as employees attempt to bypass corporate politics and red tape by using online tools to complete tasks faster. DLP can be used to retain control over who is accessing certain data types and log interactions, while ensuring that data is secured at all times. A number of DLP products natively integrate with cloud storage companies through custom APIs or by installing a physical device to the network. The device usually sits at the Internet gateway, intercepts all HTTP and HTTPS traffic and inspects it for sensitive data based on policy settings. One advantage of inspecting all data being uploaded into a cloud environment is to check that it has been encrypted before leaving the network.

SSL traffic inspection is a crucial tool, because if the traffic is not decrypted before leaving the network it will bypass policies which could enable accidental or malicious users to smuggle data outside of the company. One issue with this is that terminating the SSL connection at the gateway to view the traffic can impact user privacy. It can also have significant network overheads through the issuing of new certificates and checking each packet.

The general flow of SSL packet inspection from within the corporate network is as follows [13]:

1. Intercept the client request.
2. Establish a secure connection to the requested website and validate the site server’s certificate.
3. Create a new SSL certificate for the communication between the Security Gateway and the client. Send the client the new certificate and continue the SSL negotiation with it.
4. Using the two SSL connections:
   a. Decrypt the encrypted data from the client (from within the corporate network).
   b. Inspect the clear text content against the DLP policies set.
   c. Inspect the traffic coming from the website into the organisation network.
   d. Encrypt the data again to keep client privacy as the data travels to the destination web server resource.
Utilising cloud storage as part of an organisation's corporate infrastructure is simpler to manage as everything will be in one place. This allows features such as continuous monitoring to occur and feeds into regular audits detailing the amount, type, location and encryption status of the data held. Conventional corporate networks are large with legacy equipment attained over time and these networks and devices are generally geographically disparate. This adds a lot of complexity and requires a lot more time to be spent to try and get a clear picture of sensitive data across a network.

Organisations such as the ‘Cloud Security Alliance’ have created guidelines for companies to use, are publicly available and can help when making a decision on which cloud service provider to use. It covers security best practice in terms of infrastructure, legal considerations, IAM and data security [15].

There are also legal aspects to consider when using cloud storage providers. When signing a contract, it is important to make clear that the customer owns their data and the responsibilities of the provider are clearly outlined. Failure to do so could cause issues if a data leak were to occur, such as a compromised server. In this instance, having a fully integrated DLP solution would not protect the organisation responsible for holding customer data securely and reputational damage would fall on the company that collated or generated the data, not on the third party responsible for storing it.
5. The next frontier

As user interaction with data evolves and new applications are developed to handle information, DLP also has to mature and keep up with trends. Failing to do so will not future-proof a company against accidental or malicious data breaches, which could result in certifications being revoked or irreparable reputational damage.

**Social network analytics** can be used to get an understanding of user intention and third party collaboration. This can be of benefit to an internal investigation as they will be able to understand the reason why certain data was sent, or if they are networking with known malicious data brokers in exchange for cash incentives. Monitoring these channels of communication, however, brings up the question of data privacy and the users right to privacy. In an employment contract it is common to find that communication channels will be monitored on corporate networks and equipment, therefore, the onus is on the employee to be careful what they wish to disclose on social media while at work. The Data Protection Act and GDPR require that any sensitive personal data collected will be handled following best practice.

**Offline policy enforcement** allows the use of unauthorised devices to be flagged. When files are downloaded to authorised devices they will be decrypted and accessible. If a file is downloaded onto an unauthorised device it will remain encrypted and an alert will be sent to the information security team with details of the device and document being accessed.

**Context-based alerting** is when a document is scanned for sensitive data types, but when an alert is raised, instead of blocking the email or file upload for manual inspection the sensitive information is simply removed automatically. This allows the document to be sent or uploaded successfully without impacting business operations or clogging up consoles with unnecessary alerts. This process utilises text mining, deriving meaningful information from qualitative data through the use of patterns and natural language processing.

**Machine learning** is generally agent-based and monitors details such as the number of customer records accessed within a shift, email recipients the employee keeps in touch with on a regular basis or work patterns. The greatest benefit is when machine learning targets unstructured data, for example web pages, intranet sites, images and videos. When all of these sources are correlated together it can then be used to recieve a baseline for an individual which can then be used to form part of standard deviation alerting. This is when a significant statistical limit has been reached, such as when an employee typically views 45 customer records in a shift, however on one particular day they accessed 88. The result would be an alert to the DLP console, highlighting the anomaly so that it can then be investigated to decide if there was any malice intended.

**Artificial Intelligence** (AI) will begin to incorporate all of the above elements into DLP environments to enrich dynamic and evolving policies. These DLP deployments will be aware of their environments, current security trends, news and the user base it serves. This could then continuously monitor endpoints, gateways and users to ensure a complete cover of the corporate environment.
Wi-Fi networks are quite difficult to monitor, especially as there is currently no solution that focuses just on the data being transferred via wireless means. The best solution is to install DLP software onto individual devices that connect to these networks. An example would be that any companies embracing BYOD should implement some form of Mobile Device Management (MDM) solution, allowing for more granular controls to be refined. This could include checking the type of data being uploaded via a mobile phone or to disable the ability to access the Internet entirely and only allow access to internal content hosted on the intranet. Other solutions that are currently available include DLP agents being installed onto all connected devices. This acts in the same way a physically connected device would as the data will go to the Wi-Fi access point, passing through the DLP engine before hitting the corporate router to the internet.

NFC and Bluetooth will eventually be supported by DLP software as more IoT devices are developed and integrated with people’s everyday technologies, such as smartwatches and personal assistant devices like Amazon Echo and Google Home. These all present a risk to corporate data and make handling access to it significantly more difficult.

IoT devices are extremely affordable and provide quick solutions to numerous everyday challenges. Some recent high profile attacks have managed to cause widespread disruption due to a lack of consideration for security. One such example includes an IoT camera which had to be recalled due to a default password that was shipped with all devices. The cameras became part of a vast botnet and took down some high profile sites including Netflix, Twitter, Reddit, Spotify and the UK government website [16]. At present, there are no specific DLP products that can protect these devices from leaking data. However, there are some steps users can take to ensure the highest degree of security. These include:

- Change default passwords straight away.
- Complete a code review on the device and after being assessed as safe get the device code signed to protect against tampering.
- If data should not leave the corporate network, set up a specific policy on the companies firewall to block the device from calling out to the Internet.
- Ensure that firmware and patching levels are kept up-to-date on the device.
6. Conclusion

DLP solutions are only as effective as their configuration. The more time and effort that are spent in the early stages of deployment to define and refine the rules and polices, the better the ROI that will be seen by organisations. Getting company stakeholders involved early in the policy generation phase will ensure more value can be gained and less disruption will be faced when the system is rolled out.

DLP can be an invaluable safety net which could save organisations from significant fines or brand damage. However, DLP is not the only element needed to keep corporate data safe. It should be used as part of a ‘defence-in-depth’ approach with the most valuable tool being an organisation’s employees. User education and training can therefore be the most effective means of preventing in data loss.

DLP should be an integral part of all data consumed and generated by a business. Only then can a high degree of security be ensured. Checking data at rest can be used to locate sensitive files for either tagging or complying with PCI-DSS certifications. Data in use can ensure only authorised users are accessing the data and within appropriate volumetric limits. Data in transit is where the biggest risk lies as data is continually moving in and out of an organisation through network connectivity, or even onto removable media.

As new technologies are developed, DLP will have to evolve with it in order to prevent users either accidently losing or maliciously taking data without authorisation. The biggest threat, as far as DLP is concerned, is with new IoT devices which have multiple methods of data transport. This could potentially create a problem for IT support and security teams within organisations as they try and find a balance between outright blocking and supporting the company in embracing new technologies.
7. References & further reading


7.1 Figure table

Figure 1: Diagram detailing the split of devices and general function of a DLP solution [1].........................3
Figure 2: Security risks against printers and their data [2]........................................................................5
Figure 3: Magic Quadrant for Enterprise Data Loss Prevention [6].............................................................7
Figure 4: Email popup alert advising users [7]..........................................................................................9
Figure 5: Example of SSL traffic inspection [14].......................................................................................13
8. About NCC Group

NCC Group is a global expert in cyber security and risk mitigation, working with businesses to protect their brand, value and reputation against the ever-evolving threat landscape.

With our knowledge, experience and global footprint, we are best placed to help businesses identify, assess, mitigate & respond to the risks they face.

We are passionate about making the Internet safer and revolutionising the way in which organisations think about cyber security.