Cloud, IoT and Enterprise Security

BITS Pilani

Pilani Campus

Nishit Narang
WILPD-CSIS
(nishit.narang@pilani.bits-pilani.ac.in)

innovate achieve lead

BITS Pilani

Pilani Campus

<SSCSZG570 , Cloud, IoT and Enterprise Security>

Lecture No. 5: Enterprise Security – Securing Enterprise Data

Enterprise = Network + Systems + Data + Humans + …

Source Disclaimer: Content for many of the slides is from the course Textbook(s). Refer Course Handout for list of Textbooks.

What we shall cover?

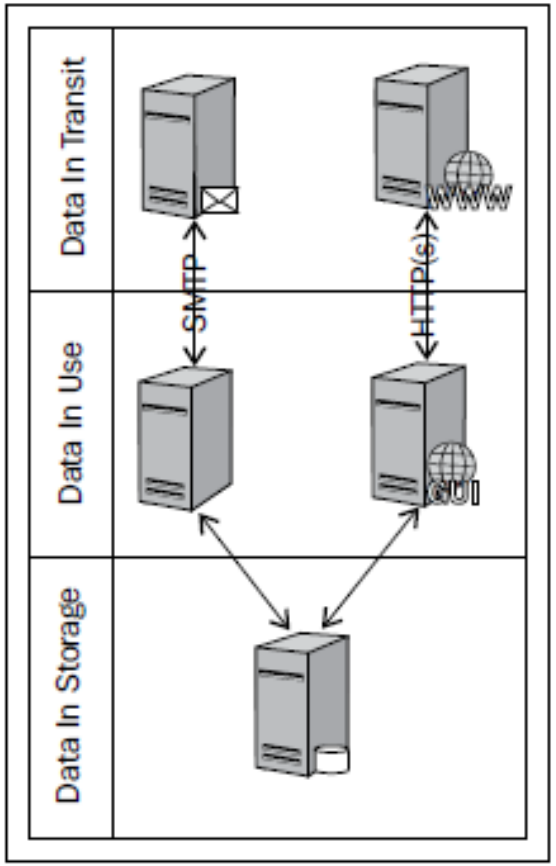
- Developing and enforcing a data classification model is a foundational component to securing enterprise data
- This lecture will focus on the steps required to develop functional data classification and how to protect high-value data in the enterprise
- We shall cover:
  - Data identification and classification
  - Data loss prevention methods and techniques
  - Data protection methods and techniques such as encryption, hashing, and access controls

Data Classification Process

- Involves two steps: identification and classification of enterprise data
  - specific handling methods defined for interacting with the classified data
  - data owners are assigned, enterprise criticality is scored
  - supporting processes are developed to ensure confidentiality, availability, and integrity
- Classification is done based on:
  - importance and
  - impact potential (i.e. impact of enterprise data compromise or loss)

Step 1: Data Identification

- What we have already said about this in past lectures?
  - There are many data types that exist in order for the business to operationally function
    - Example: Employee human resources data, Company private data (business plans, acquisition strategies, brands, and so on), Company confidential data, Company public data (product releases, press releases) etc
  - Data can be located in multiple places both internal and external to the enterprise network, including in employer-owned and employee-owned assets
    - Example: Network shares, Document repositories, File transfer systems, Business partner and third-party systems, Employer and employee laptops/desktops etc
  - Data can be at rest, in use or in transit
    - Each will have a unique set of challenges to provide the protection dictated by the classification model



Source: Aaron Woody, Enterprise Security: A Data-Centric Approach To Securing The Enterprise

Step 2: Data Classification Assignment

- The act of assigning a label to identified data types that indicate required protection mechanisms
  - driven by business risk and data value
- Example Data Classification:

	Restricted confidential (Level 1)	Confidential (Level 2)	Public (Level 3)
Data type	Customer: • CC# • PII Employee: • SSN# • PII Company: • Merger Plans • New product	Customer: • PII Employee: • PII Company: • Internal documents	• Anything not in the previous sections. • Items considered to be available in the public domain.
Data protection	Data encryption, hashing, or tokenization	Restricted access permissions	None

PII = Personally identifiable information
CC = Credit Card
SSN = Social Security Number

Source: Aaron Woody, Enterprise Security: A Data-Centric Approach To Securing The Enterprise

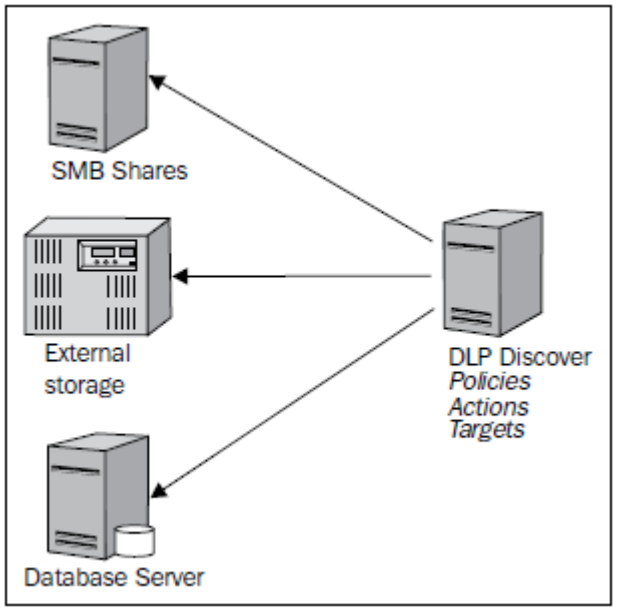
Data Loss Prevention

- **Data Loss Prevention** (**DLP**) is a tool that can enforce protection of data that has been classified
- The primary purpose of DLP is to protect against the unauthorized exfiltration of enterprise data
- In general, DLP solutions can:
  - Help find data in various locations within the enterprise
  - enforce encryption, in some cases
  - block insecure transmission, and
  - block unauthorized copying and storing of data, based upon data classification
- In next slides, we will cover the implementation of DLP for the common data locations in the enterprise

NEED FOR DLP: *No network monitoring device will detect if, for example, thousands of medical records are saved to a local machine and moved to a USB storage device, but Endpoint DLP can detect and prevent this action!!*

DLP: Data in Storage

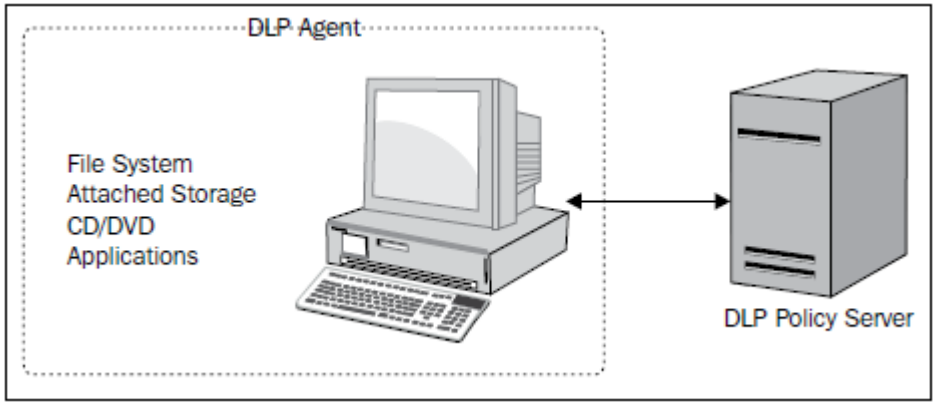
- Data can be stored in network shares, databases, document repositories, online storage, and portable storage devices
- Most DLP solutions have the ability to scan data stores and also provide an agent that can be deployed on end systems to monitor and prevent unauthorized actions for classified enterprise data
- Using DLP, a discovery scan can be initiated to identify data in locations
- Also, it can be used in an ongoing scheduled scan to continuously monitor the data stores for data that should or should not reside in the data location



Source: Aaron Woody, Enterprise Security: A Data-Centric Approach To Securing The Enterprise

DLP: Data in Use

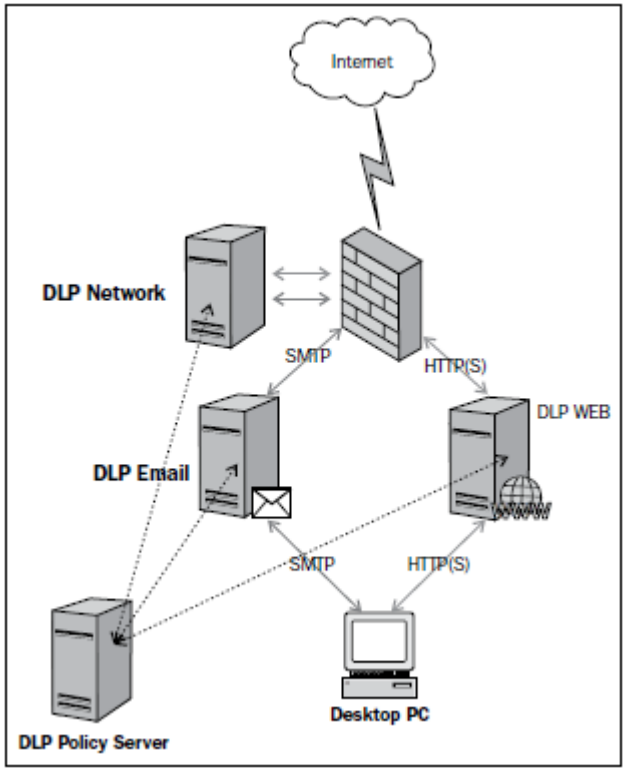
- Data in use is data that is actively processed within an application, process, memory or other location, temporarily for the duration of a function or transaction
  - i.e. enterprise data not stored long term, only long enough to perform a function or transaction
  - there is an application or function involved to read, add, remove, and modify data
- Data in use is the unique facet of DLP that is a little more complex than dealing with data in storage or data in transit
  - Data in use can be monitored by an agent installed on the end system to permit only certain uses of the data and deny actions such as storing the data locally or sending the data via e-mail or other communication method
  - implementation on employee-owned devices introduces privacy issues because any personal transactions such as online banking, medical record lookup, and so on may be detected and details of the transaction stored in the DLP database for review → must be carefully evaluated when considering a BYOD deployment!!



Source: Aaron Woody, Enterprise Security: A Data-Centric Approach To Securing The Enterprise

DLP: Data in Transit

- Data in transit is data that is being moved from one system to another, either locally or remotely, such as file transfer systems, e-mail, and web applications
  - focus of DLP for data in transit is specifically data leaving the enterprise through egress connections
  - Yet, it is recommended that all data including credentials be transmitted only using secure methods, even within the internal enterprise network
- Many enterprise communication applications may be invisible to network-based security solutions
  - Example, use of instant messaging to send files or data
- Various DLP solutions have accounted for this fact and provide solutions capable of intercepting and decrypting communications to look for classified data



Source: Aaron Woody, Enterprise Security: A Data-Centric Approach To Securing The Enterprise

*Careful choice needed from enterprise security admin if the **next generation firewall (NGFW)** can be better used or a DLP!!*

Implementation of DLP

- The challenge with the DLP toolset is deciding what methods to employ, in what phases, and how to digest the output from the tool
  - *=> challenge exists in operationalizing the solution and delivering value on the investment!!*
- The best method to implementing any solution in the enterprise is to first understand the problem to be solved, and then determine the course of action
- The following slides cover the DLP solutions, approaches to successfully implementing them, overcoming challenges, and getting value from a DLP implementation

DLP Network

- simplest solution to implement in an enterprise environment
- also the quickest method to determine what data is leaving the network in an insecure manner
- Implementation Considerations:
  - Volume of traffic to be inspected
  - Server size requirements to run DLP function (*else, over-flooding can lead to data being lost!!)*
  - Protocols to be inspected (to limit inspection volume)
  - Person or team to whom findings are to be reported

DLP Email and Web

- Email and Internet access are the most commonly used enterprise services
- DLP (Email and Web) goes beyond the basic network portion of DLP
  - Focus more on loss of enterprise confidential data via emails or web
- Implementation Considerations:
  - Placement of DLP (e.g. along-side existing Internet proxy servers and e-mail forwarders)
  - changes to the use of e-mail and web within the enterprise (e.g. encrypted emails)

DLP Discover

- Is a tool that can scan network shares, document repositories, databases, and other data at rest
- Requires an account with permissions to be configured, to allow the scans to open the data stores and inspect for policy matches
- Implementation Considerations:
  - advisable to run scans during off hours as the solution may increase the I/O on the system being scanned and impact performance
  - permission errors impede the success of the scan; testing by initiating a limited scan can help identify simple issues that will otherwise derail the scan
  - If there are file auditing controls in place, the DLP solution may trigger alerts based on file access operations → such false positive alerts should be possible to identify and ignore

DLP Endpoint

- DLP Endpoint is an agent-based technology that must be installed on every end point
  - closest to the end user where the human interaction is the highest and, in theory, where the greatest risk is introduced to enterprise data
- In a typical enterprise, there will be more end point systems than any other hardware combined
  - => requires a significant implementation of agents that have to be installed, managed, and the output operationalized for meaningful and actionable reporting
- Implementation Considerations:
  - Use of enterprise common software management tools to install agents remotely on end point systems
  - Verification of agent for a variety of platforms (OS etc)
  - incidents may be exponentially higher than from other DLP solutions
  - Employee personal data and operations privacy issue!!

Data Protection Methods

- Earlier slides discussed Data Loss Prevention methods and techniques
- Next few slides will discuss methods for Data Protection, using different methods
  - Encryption and Hashing
  - Tokenization
  - Data Masking
  - Authorization

Encryption & Hashing

- Both encryption and hashing are typically what is thought of when data protection is discussed whether in storage, transit, or in use by applications
  - Mostly for data in storage or in transit
- Encryption is the method of mathematically generating a cipher text version of clear text data to render it unrecognizable
  - There are two general types of encryption – symmetric and asymmetric
  - data encrypted using a symmetric key can also be decrypted with the same key
  - Asymmetric encryption is different than symmetric methods because the master key (private key) is never shared; data encrypted is done so using the server's public key
- Hashing is simpler, but only supports data integrity

Encryption: Data at rest

- encryption can happen at the location of storage, prior to storage, or during the process of storing
  - ensure the business processes and applications can support the method used
- Another aspect to encrypting data at rest is online versus offline encryption
  - online encryption is in effect while data is accessible
  - offline is when data is not directly accessible such as on backup tapes, turned off systems, etc
    - An example of offline encryption is a whole disk encryption, once the operating system is booted and the volume is decrypted for use. Post boot, data is no longer encrypted and can be accessed in an unauthorized manner

Data @ Rest Encryption

- Data stored in databases can be encrypted via two methods
- first method utilizes the built-in encryption capabilities of the database itself to protect the stored data
  - beneficial when attempting to make encryption invisible to the applications and processes accessing the data
  - Caveat: If not configured properly, the system administrators can circumvent the database encryption
- Second method uses encrypting at the application and process layer
  - All data is encrypted before it is stored in the database

Application Encryption

- the encryption of the data occurs in the application not the database
- data arrives as already encrypted in the database
- all applications and processes using this data need a method to decrypt and encrypt the data → typically a shared private key
- Benefits:
  - Database performance gains for not using encryption at the database tier
  - The data is always encrypted in the databases (no DB admin or SYS admin visibility)
  - Data encryption is implemented end to end

Selective Database Encryption

- refers to encrypting only portions of the database; typically selected columns that contain sensitive data
- Benefits
  - often employed to reduce the overall load on the database server for encryption
  - also to make it easier for the DB admins to ensure the data inserted into the database is correct
- Caveat
  - DB admin has full control over the database encryption, if the individual decides to see the data in an unauthorized method, by changing configuration
  - However, monitoring and detection of the unauthorized change can be the only real protection from this unauthorized access
- Alternate to selective DB encryption is the Complete DB encryption
  - The method implemented must make sense from data protection and risk analysis perspectives, due to its overhead costs

File Share Encryption

- As with databases, many operating systems offer native encryption
- There are technologies available that will encrypt data as it is being written to the file system
- Similar options exist:
  - Encryption within the application
  - Encryption outside the application
    - Require other methods for enforcing least privilege and ensuring only the necessary processes, applications, and users have permissions to access data

Data in Use Encryption

- Not many use cases
- An example could be fraud investigators leveraging stored credit card and transaction information for an investigation
  - In this scenario, access to the data is necessary but should not be visible to prying eyes on the network
- Requires commercial software offering secure communication and views that can be created to ensure that only the fields needed are viewable

Data in Transit Encryption

- Performed via use of secure transport methods to transfer data
  - E.g. SSL, SFTP, FTP-S, and SSH, in addition to proprietary solutions
- If the transport cannot be secured, then the data itself must be encrypted

Tokenization

- **Tokenization** is a method that assigns a value to a segment of data, so that the initial sensitive data value no longer exists
  - use in applications and storage in the database
  - processes, systems, and applications are able to process the token value as they would process the sensitive data
  - however, this method ensures that the token has no real value to anyone or anything outside of the process
- A database is used to map the original data to the token value
- A common use for tokenization is in the retail industry for the replacement of credit card data within the network and assets
  - allows retailers to escape the prescriptive security controls required (i.e. reduce PCI DSS scope)
  - is an option gaining momentum
- There is no real standard for tokens but one method to consider is format preserving to reduce complexity in rewriting applications for new formats

Data Masking

- This method is commonly used in processes where there is human interaction
  - example would be looking at your stored credit card information at an online retailer
  - Typically, your credit number will be masked (series of asterisks) except for the last four digits
- A similar method can be achieved in database views and specialized encryption solutions to enforce the least privilege and access only on a *need-to-know* basis
- Pros-and-Cons:
  - + relative ease of implementation
  - - Masking as used on a database implementation is simply a view presented with the original data intact and viewable by database administrators
  - - While the solution does provide some protection, it is not at the same level as tokenization, encryption, or hashing

Authorization

- Granting permissions based on who or what the authorized is
  - An important part of the enterprise data protection and security program
  - each of the previous approaches on data security relies on proper authorization to underlying operating systems, applications, and the data
- This facet of data security highlights the defense in depth mantra of information security
  - Regardless of the technologies implemented for encryption, tokenization, and masking, a developed process for authorization including access provisioning, account removal, level of access, and auditing will not only ensure that the data remains secure, but provides a defensible data security strategy that can aide in reducing risk and cost associated with external auditing engagements