BigData Analytics

Unit 6:
Big Data Security

Big Data Security

6.1. Overview of Big Data Security

6.2. Understanding Security Challenges in Big Data

6.3. Big Data Security Technologies

6.4. Big Data Authentication and Authorization

6.5. Big Data Security in Cloud Environment

6.6. Big Data Security in Hadoop

6.7. Big Data Security in AWS

6.8. Big Data Privacy and Security – Case Studies

CE: 6.1.
Overview of Big Data Security

CE: 6.1. Overview of Big Data Security

- Big data security can be termed as the ***tool*** and ***measures*** which are used to guard both ***data*** and ***analytics*** processes.

- The main purpose of Big data security is ***to provide protection against attacks, thefts, and other malicious activities*** that could harm valuable data.

  - Big data security challenges are multi-faced for the companies that operate on the cloud.

  - This challenging threat includes ***the theft of information stored online***, ***ransomware***, or ***DDoS attacks*** that could crash a server.

  - These threats can cause serious financial consequences such as losses, trial costs, and fines or authorizations of an organization.

CE: 6.1. Overview of Big Data Security (Conti…)

- Compassions around ***Big data security and privacy*** are a hurdle that needs to be overcome.

- ***Intelligent analytics*** has been introduced to enhance security with the help of the proposed security intelligence model.

CE: 6.1. Overview of Big Data Security (Conti…)

How you can implement Big Data Security in Organizations?

- ***Encryption is one of the most common security tools.***
  - Encrypted data is hard to decode for hackers.
  - Encrypted data is generally done for the *incoming data* as well as for *outgoing data*.

- ***If we look at the other Big data security tools, then the best one is Firewall.***
  - Firewalls are usually used to filter the traffic that enters and leaves servers. Firewall creates strong filters that prevent attacks from malicious activities.

- ***BI tools and analytics platform is another key to protect vital information of the organization.***
  - BI tools used to create an access system that can reduce the possibility of an attack to a great extent.

CE: 6.1. Overview of Big Data Security (Conti…)

❖ **Some of the reasons why it is important to follow the best practices for Big Data security are mentioned below :**

- It boosts the security of non-relational data scores.
- It helps to implement endpoint security.
- Ensures the safety of transactions and data storage logs.
- Rely on Big Data Cryptography.
- Use Customized solutions.
- Practice real-time security monitoring and compliance.
- Enhances communication and availability of information.
- Allows for convenient resource sharing.
- Increases systems efficiency and robustness.
- Avoiding unauthorized access thus protects and enhances the performance and security of the organization.

CE: 6.1. Overview of Big Data Security (Conti…)

- Thus, **Big data security's mission is clear enough:**
  - keep out on unauthorized users and intrusions with firewalls,
  - strong user authentication, end-user training, and intrusion protection systems (IPS) and intrusion detection systems (IDS).

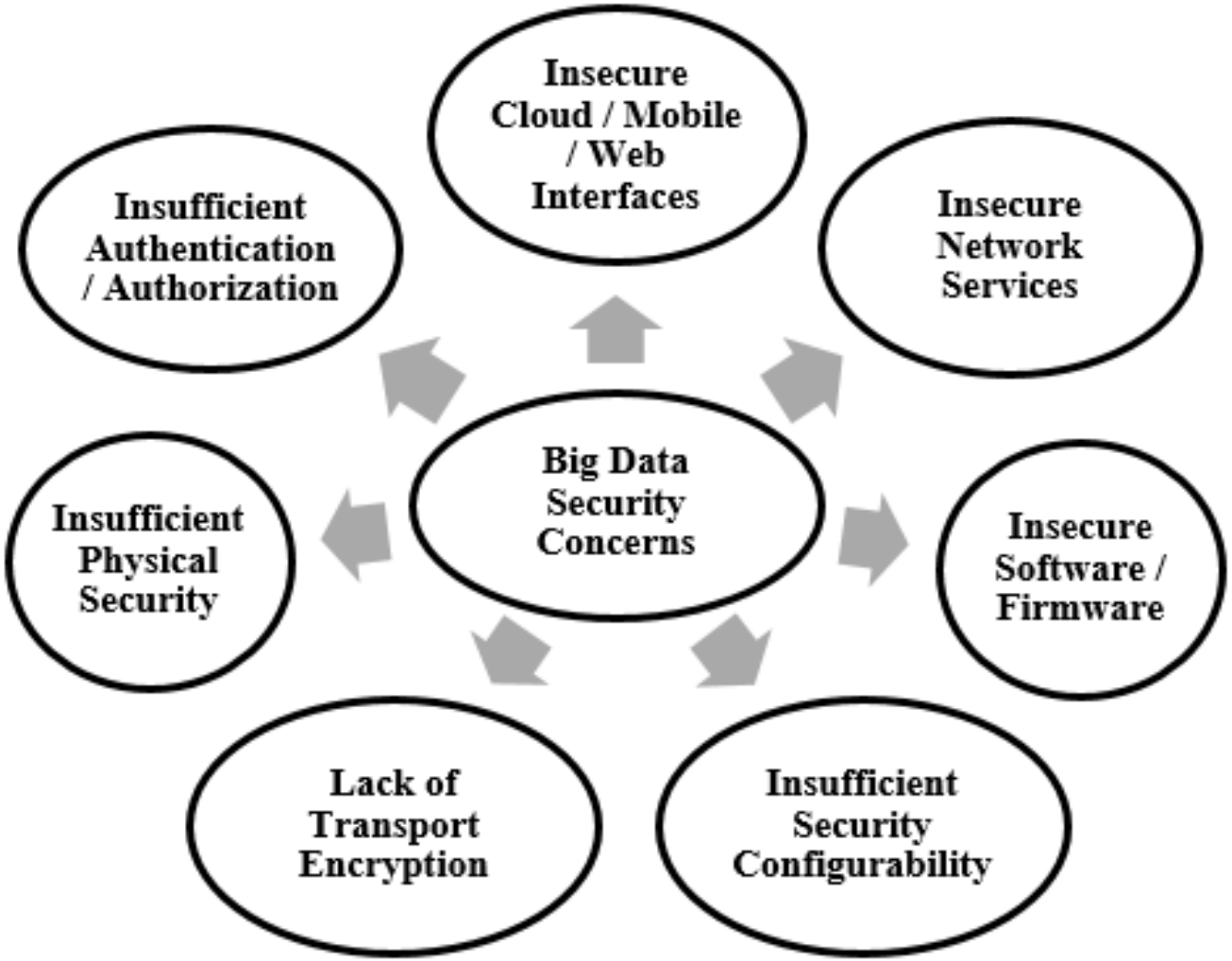
- In case someone does gain access, encrypt your data in transit and at rest.

- This sounds like any network security strategy. However, big data environments add another level of security because security tools must operate during three data stages that are not all present in the network.

- These are…
  1) data ingress (what's coming in),
  2) stored data (what's stored), and
  3) data output (what's going out to applications and reports).

CE: 6.2.
Understanding Security Challenges in Big Data

CE: 6.2. Understanding Security Challenges in Big Data



https://hevodata.com/learn/big-data-security/#challenges

CE: 6.3.
Big Data Security Technologies

CE: 6.3. Big Data Security Technologies

- **Big Data Security Technologies**
  1. Encryption
  2. User Access Control
  3. Physical Security
  4. Centralized Key Management

- Below are few of the representatives of Big data security companies:

  1. Cloudwick
  2. IBM
  3. Logtrust
  4. Gemalto

https://techvidvan.com/tutorials/big-data-security/

CE: 6.4.
Big Data Authentication and Authorization

CE: 6.4. Big Data Authentication and Authorization

What is Authentication???

- ***Authentication*** is used by a server when the server needs to know exactly who is accessing their information or site.
- ***Authentication*** is used by a client when the client needs to know that the server is system it claims to be.
- In authentication, the user or computer has to prove its identity to the server or client.
- Usually, authentication by a server involves the use of a *user name and password*.
- Other ways to authenticate can be through *cards, retina scans, voice recognition, and fingerprints*.

CE: 6.4. Big Data Authentication and Authorization

What is Authentication??? (Conti…)

- Authentication does not determine what tasks the individual can do or what files the individual can see.
- Authentication merely identifies and verifies who the person or system is.

CE: 6.4. Big Data Authentication and Authorization

What is Authorization???

- ***Authorization*** is a process by which a server determines, if the client has permission to use a resource or access a file.
- ***Authorization*** is usually coupled with ***authentication*** so that the server has some concept of who the client is that is requesting access.
- The type of authentication required for authorization may vary; passwords may be required in some cases but not in others.
- In some cases, there is no authorization; any user may be use a resource or access a file simply by asking for it.
- Most of the web pages on the Internet require no authentication or authorization.

CE: 6.4. Big Data Authentication and Authorization

https://docs.informatica.com/data-engineering/data-engineering-integration/10-2/big-data-management-administrator-guide/authentication-and-authorization/authentication-and-authorization-overview.html