International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Ensuring Data Integrity and Privacy: A Guide for Database Administrators

Balakrishna Boddu

Sr. Database Administrator
balakrishnasvkbs@gmail.com

ABSTRACT:
In today's digital age, information has become an invaluable asset for organizations across all sectors. From social and governmental bodies to educational institutions, nearly every organization has automated its operations, relying heavily on databases to store and manage crucial information. This reliance on databases has elevated the importance of database security to a critical concern.
Security is a critical challenge globally, including in the digital realm. This paper examines database sec urity, highlighting its growing importance for businesses. Databases hold vital information, making their protection essential. The survey covers issues and threats in database security, outlines necessary securit y measures, and explains how encryption is used at various levels to enhance protection.

Keywords: DBA, Threats, encryption, audit, access controls, policy, encryption, media, protection, safety, monitoring.

1. Introduction:

Database managers (DBAs) are important for keeping sensitive information secure. This journal provides DBAs with how to make sure the information in the database is accurate, consistent, and complete. There are ways to check the information for errors, organize it well, and make sure different parts of the database line up with each other. It's also important to be able to back up the information and restore it if something goes wrong. Another big concern is privacy, especially with laws like GDPR and CCPA. These laws have rules about how DBAs need to handle personal information. There are techniques to hide some personal details while still being able to use the information, and there are also ways to limit who can access the database and to scramble the information so it can't be read without a special key. By following this guide, DBAs can keep information safe and reliable for their organizations.
Database security refers to the protection of sensitive data stored within a repository. It encompasses measures to safeguard databases from unauthorized access, illegal activities, and various threats at all levels. Ensuring database security involves regulating user actions and granting or denying permissions to access the database and its contained objects.
Successful organizations recognize the imperative of maintaining database confidentiality. They strictly prohibit unauthorized access to their data, understanding the potential consequences of such breaches. Additionally, these organizations demand assurance that their data is shielded from malicious or accidental modifications. Data protection and confidentiality are fundamental security concerns that must be addressed with utmost priority.

The core properties of database security are confidentiality, integrity, and availability. Confidentiality ensures that only authorized individuals can access sensitive information, preventing unauthorized disclosure. Integrity guarantees the accuracy and reliability of data, safeguarding it from unauthorized modifications or alterations. Availability ensures that data is accessible when needed, preventing disruptions to critical operations. By upholding these three principles, organizations can effectively protect their valuable data and maintain the integrity of their operations.

2. Research Work:

The core properties of database security are confidentiality, integrity, and availability. Confidentiality ensures that only authorized individuals can access sensitive information, preventing unauthorized disclosure. Integrity guarantees the accuracy and reliability of data, safeguarding it from unauthorized modifications or alterations. Availability ensures that data is accessible when needed, preventing disruptions to critical operations. By upholding these three principles, organizations can effectively protect their valuable data and maintain the integrity of their operations.

Database encryption: is proposed in which database encryption can be provided as a service to applications with unified access to encrypted databases. Using such an encrypted data management model, applications can concentrate on their core businesses and protect data privacy against both malicious outsiders and untrusted database service users without need-to-know encryption details.

Data masking: is a technique used to protect sensitive data by replacing it with non-sensitive substitutes. This helps to maintain data privacy and security while still allowing for data analysis, testing, and development activities.

Feature	Encryption	Data Masking
Purpose	Protecting data in transit and at rest	Protecting data in use
Process	Cryptographic algorithms and keys	Tokenization, substitution, or generalization
Reversibility	Reversible	Often irreversible
Common Use Cases	Data transmission, storage	Development, testing, data sharing

Diagram: Encryption vs Data Masking

There are multiple ways of Security layer can be built to safeguard Databases and the below Diagram will help to prevent unwanted hits from malware and other threats.

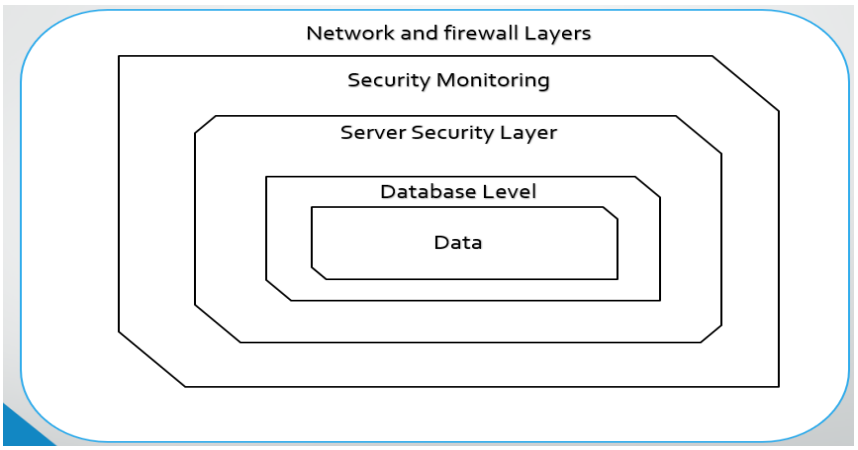

Diagram: Generic Layer of An Organization Security

3. Methodology:

some of the core areas need to be implemented to gain Data integrity and the Database Administrator can follow the below criteria

Assessment: Evaluate current data integrity and privacy measures. Identify gaps and vulnerabilities.

Policy Development: Establish clear data governance policies, addressing data handling, storage, and sh aring protocols.

Technology Implementation: Integrate encryption, access controls, and monitoring tools to safeguard d ata.

Training and Awareness: Conduct regular training sessions for DBAs on best practices and emerging th reats.

Continuous Monitoring: Implement continuous monitoring systems to detect and respond to potential breaches or anomalies promptly.

Review and Audit: Regularly review and audit data management practices to ensure compliance with e volving regulations and standards. To eliminate the security threats every organization must define a security policy. And that security policy should be strictly enforced. A strong security policy must contain well-defined security features. Figure 2 shows some critical areas that need to be considered.

Access Control: Access control ensures all communications with the databases and other system objects are according to the policies and controls defined. This makes sure that no interference occurs by any attacker neither internally nor externally and thus, protects the databases from potential errors that can make an impact as big as stopping the firm's operations. Access control also helps in minimizing the risks that may directly impact the security of the database on the main servers. For example, if any table is accidentally deleted or access is modified the results can be rollbacked or for certain files, access control can restrict their deletion.

Inference Policy: Inference policy is required to protect the data at a certain level. It occurs when the interpretations from certain data in the form of analysis or facts are required to be protected at a certain higher security level. It also determines how to protect the information from being disclosed.

User Identification/Authentication: User identification and authentication are necessities to ensure security since the identification method defines a set of people that are allowed to access data and provides a complete mechanism of accessibility. To ensure security, the identity is authenticated and it keeps the sensitive data safe and from being modified by any ordinary user.

Accountability and auditing: Accountability and audit checks are required to ensure the physical integrity of the data which requires defined access to the databases and that is managed through auditing and record keeping. It also helps in the analysis of information held on servers for authentication, accounting, and access of a user.

Encryption: Encryption is the process of concealing or transforming information using a cipher or a code so that it becomes unreadable to all other people except those who hold a key to the information. The resulting encoded information is called encrypted information.

Diagram: Encryption Process

4. Analysis:

Companies store their most important information, like financial records and customer data, in big computer systems. These systems are like treasure chests for the company, but if they're not built or set up properly, they can be easily broken into and stolen from.

It seems like every week there's a new story about hackers stealing information from poorly protected cloud databases or storage systems. This stolen information can include things like:

- Personal information on almost all the people in Panama (90%)
- Personal information on millions of people in Ecuador (20.8 million)
- Social media posts from 48 million people
- 142GB of documents

The yearly average data breach cost increased the most between the years 2020 and 2021 - a spike likely influenced by the COVID-19 pandemic.

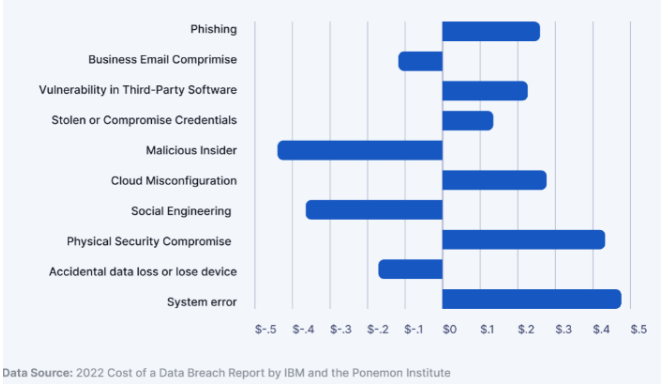

Diagram: Total data breach cost in different entities

To carry out the study, the different articles were analyzed to list the possible threats that a database may suffer, according to the results of the analysis it was possible to determine the following.

1) List of Attacks

Computer attacks are established as malicious acts by a group of people seeking the vulnerability of the system to cause damage to an infrastructure.

2) Global Attack Report

Data on total losses for the past 5 years were collected, and the total loss was determined to be $10.2 billion; the formula determined that the average reported threats are 341,523.6 per year; the data collected is global and dollar loss can be calculated based on threats found and reported.

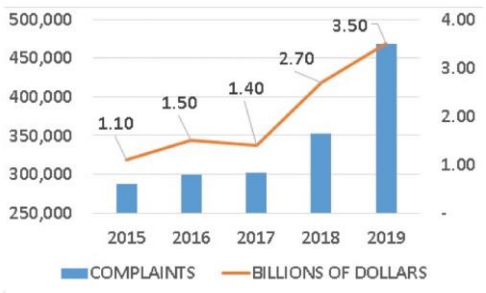

Diagram: Revenue Impact by Data Breach

4. Risks and Implementation:

The initiative database organization is subject to a prodigious variety of threats. Some serious threats are envisioned in this document.

High Privilege: When people are permitted to do things at work that they're not supposed to do, it can show that they have bad intentions. This can lead to them using their extra privileges to do harmful things.

Inference Policy: Inference policy is required to protect the data at a certain level. It occurs when the interpretations from certain data in the form of analysis or facts are required to be protected at a certain higher security level. It also determines how to protect the information from being disclosed.

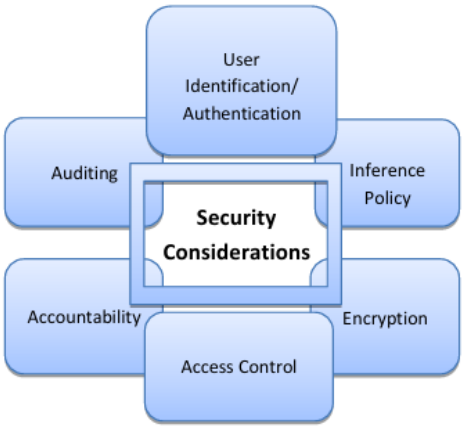

Diagram: Critical Areas under Consideration

User Identification/Authentication: user identification and authentication are necessities to ensure security since the identification method defines a set of people that are allowed to access data and provides a complete mechanism of accessibility. To ensure security, the identity is authenticated and it keeps the sensitive data safe and from being modified by any ordinary user.

Accountability and auditing: Accountability and audit checks are required to ensure the physical integrity of the data which requires defined access to the databases and that is managed through auditing and record keeping. It also helps in the analysis of information held on servers for authentication, accounting, and access of a user.

Empirical Analysis:

This empirical study is done by keen observation of the literature and then results are drawn. The frequency of benchmarks in different papers that were under consideration is shown below in a table.

Frequency: Frequency is the number of occurrences of a repeating commonness. The frequency is calculated in such a way that the paper which has an issue not common in some other paper is evaluated as having frequency "1" whereas the papers which have the common issues have been given frequency equal to the number of papers having that issue.

Criticality: To find the measure of the frequency of occurrence of an issue the Criticality factor is divided into four parts i.e. Medium, Moderate, High, and Very High. The percentage range for criticality is defined below:

Percentage	Criticality
10-20 %	Medium
20%-50%	Moderate
51%-80%	High
81%-100%	Very High

Conclusion:

Data to any organization is the most valuable property. Security of sensitive data is always a big challenge for an organization at any level. In today's technological world, the database is International Journal of Computer Applications (0975 – 888) Volume 47– No.12, June 2012 vulnerable to hosts of attacks. In this study, major security issues faced by databases are identified and some encryption methods are discussed that can help to reduce the risks of the attack and protect sensitive data. It has been concluded that encryption provides confidentiality but gives no assurance of integrity unless we use some digital signature or Hash function. Using strong encryption algorithms reduces the performance. Future work could be carried out to make encryption more effective and efficient.

References:

1. **"Database Administration: The Complete Guide to DBA Practices and Procedures" by Craig S. Mullins** - This comprehensive guide covers a wide range of DBA practices, including data integrity and privacy(https://books.google.com/books/about/Database_Administration.html?id=JWoKCHJheSUC)
2. **"Data Privacy Best Practices in Database Administration"** - This blog post on Data Sleek discusses best practices for ensuring data privacy in database administration.(https://data-sleek.com/data-privacy-best-practices-in-database-administration/)
3. **"Governance and Data Security for Database Administrators"** - This blog post on PostgreSQL explores governance and data security, essential for maintaining data integrity and privacy(https://www.postgresql.fastware.com/blog/governance-and-data-security-for-database-administrators)
4. **"Data Privacy and Ethical Considerations in Database Management"** - This article from MDPI discusses data privacy and ethical considerations in database management. (https://www.mdpi.com/2624-800X/4/3/24)
5. **" Ensuring Data Integrity in Databases with the Universal Basis of Relations"** - This article from Applied Sciences explores methods to ensure data integrity in databases.(https://www.mdpi.com/2076-3417/11/18/8781)