Security in Database Systems

Research paper by Dimitar Petrov

Abstract:

The focus of this paper is about database systems, which are commonly used methods by firms and products as keeping track of various data and securing one efficiently. A database is, by definition, an organized collection of structured data stored in a computer or server, protected by a database management system (DBMS). As this storage is entirely digital, it is protected from physical theft, but on the other hand, without the proper safety procedures, is at the risk of being stolen via hacking. Naturally, it is of upmost importance to know the many ways data can be compromised and make a database system as secure as possible.

I. Introduction

Protection of data is not something trivial. Suitable methods and tools are necessary to satisfy the following three requirements that should be met in any feasible database system:

1. ***Identification.*** Any system must be able to identify its users and confirm their identity.
2. ***Authorization and Authentication***. Any system must have different permissions for different users.
3. ***Access Control***. Access controls maintain a separation between the users, as well as the various data and computing resources and protect all internal resources from unauthorized or improper modification.

The paper will focus on all 3 points separately and their importance on a DBMS. Note this paper will not go into full descriptive detail about every solution to each of the requirements, but will instead demonstrate the most efficient, state-of-the-art resolutions for each of these problems.

II. Risks of an Unsafe DBMS

Before analyzing the methods of securing the database, it is beneficial to first understand what we are attempting to avoid.

We consider database security for the following situations:

- Theft and fraudulent activity.
- Loss of confidentiality or secrecy.
- Loss of data privacy and integrity.
- Loss of availability of data.

But lists like these never truly accurately portray the potential catastrophic loss when one company's database is breached, so we will continue with examples of the biggest losses as of just the previous year:

- During March and April of 2019, Facebook has exposed users' credentials on numerous occasions due to poor password storage management, going as far as storing Instagram users' passwords in plaintext format with no encryption whatsoever. The numbers of affected clients have exceeded 540 million. (Ng, 2019)
- In April 2019, another major breach in a database managed by an Indian government healthcare agency was the cause for the leaking of 12.5 million medical records of pregnant women. Records go as far back as five years, to 2014, and include detailed medical information for women who underwent an ultrasound scan, amniocentesis, or other genetic testing of their unborn child.
- On July 15th, 2019, the NRA (National Revenue Agency) of Bulgaria has taken a large blow from a hacker attack with the intent of taking down its corrupt government, using the data as a hostage of sorts. The hacker has shared the news to large media outlets with parts of the stolen data, claiming that he has collected around 5 million Bulgarians' records overall. On the following day, the NRA has confirmed the hacker's threats and further elaborated that the total leakage amounts to 3% of their database and the attack has been done via an SQL injection. While the stolen data was never released by the hacker and he was never caught, one media outlet showed the email sent by the hacker without censoring a torrent link to a small part of the leaked records, which was linked by the hacker to show proof of his threats to the media outlets, jeopardizing about 250,000 people's records as any person was able to download the data from the torrent for several hours.

Again, may I put focus on the fact that these major attacks happened within a span of just 4 months. Hopefully, this suffices as a reason enough to go through the procedures of ensuring a database.

III. The Aspects of DBMS Security

One of the first and key lessons when studying anything connected to storing data digitally is identification, authentication and I would like to add on top of those two – authorization. Even when following a simple website tutorial, most go into detail about the topic of user organization and there is a reason for that. This is the most basic method of protecting a database system but a paramount one.

Identification – A good password is among the most basic precautions to take to secure a database, but also among the most important. Hackers have extensive tools for cracking passwords, so it is highly recommended to make one complex. For reference, see below a table of how much time it takes for a password cracking tool to crack a password depending on several criteria:

number of Characters	Numbers only	Upper or lower case letters	upper or lower case letters mixed	numbers, upper and lower case letters	numbers, upper and lower case letters, symbols
3	Instantly	Instantly	Instantly	Instantly	Instantly
4	Instantly	Instantly	Instantly	Instantly	Instantly
5	Instantly	Instantly	Instantly	3 secs	10 secs
6	Instantly	Instantly	8 secs	3 mins	13 mins
7	Instantly	Instantly	5 mins	3 hours	17 hours
8	Instantly	13 mins	3 hours	10 days	57 days
9	4 secs	6 hours	4 days	1 year	12 years
10	40 secs	6 days	169 days	106 years	928 years
11	6 mins	169 days	16 years	6k years	71k years
12	1 hour	12 years	600 years	108k years	5m years
13	11 hours	314 years	21k years	25m years	423m years
14	4 days	8k years	778k years	1bn years	5bn years
15	46 days	212k years	28m years	97bn years	2tn years
16	1 year	512m years	1bn years	6tn years	193tn years
17	12 years	143m years	36bn years	374tn years	14qd years
18	126 years	3bn years	1tn years	23qd years	1qt years

Authentication/Authorization/Access Controls – **"**Access controls maintain a separation between the users on one hand and the various data and computing resources on the other and protect all internal resources from unauthorized or improper modification. – Bertino E. – Database Security: Research and Practice

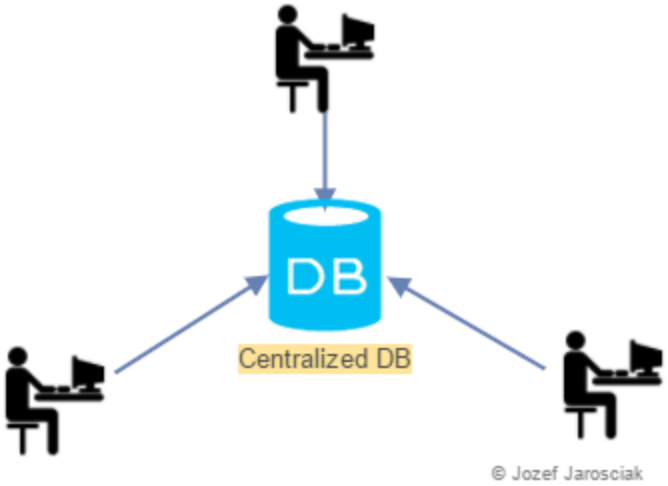
This is less of a safety measure against hackers but for internal tampering from employees that have access to the database, be it intentional or not. The more structured a database is and organized for users, the safer and foolproof it will be.

Encryption – Encryption is among the most important precautions you can take for a secure DBMS. It does not only serve to protect your data in the case that a successful hacker attack occurs, but it also improves a database's defense from hacker attacks – as an encryption can render an SQL Injection – one of the most common methods of hacking a database – completely useless. For encryption techniques, I personally recommend this article[1] that gives extensive explanations of eight common encryption algorithms and extra recourses on them.

IV. Local or Distributed

Developing a database requires this question to be answered beforehand – should it be based locally or online? While this decision benefits or in some cases hinders the business standpoint more, it should be noted that it has an effect on security quality as well.

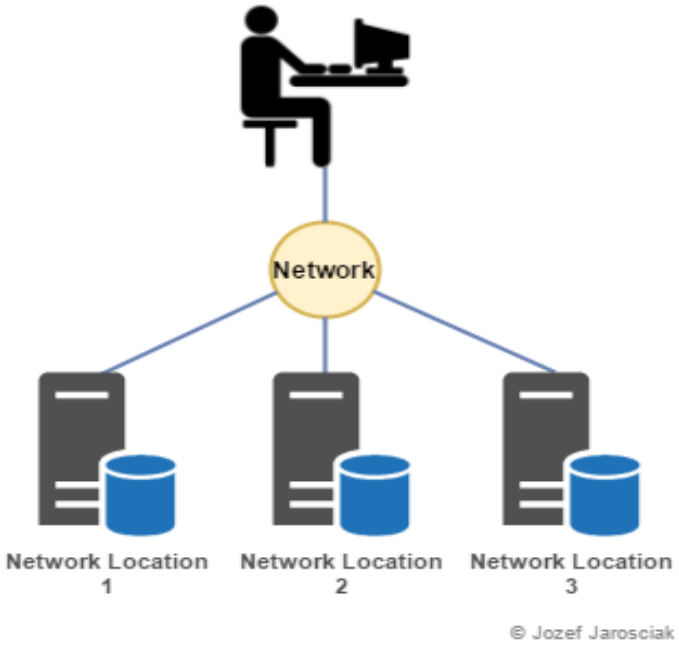
Local/Centralized – A centralized database is one where the DBMS is stored in one single location, which could mean just one server or one computer and users of it can be connected via a several ways, the most common one being a computer network.



Generally, this method is preferred when the system does not need to be accessed from a remote location and in a case where increased network traffic will hurt the performance. But it has other advantages, such as: being easier to set up and it restricts hacker's options to tamper with one's system, as it does not allow outsider systems.

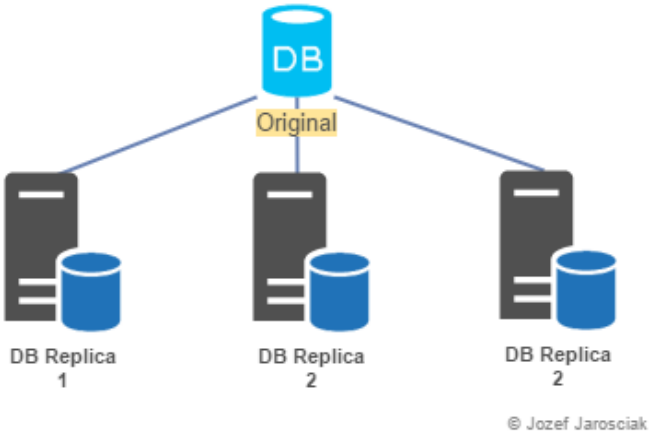
[1] https://www.rankred.com/common-encryption-techniques/

Distributed – A distributed system is one far more complex than a locally based one, but it comes with many benefits. Most impactful pros are that a distributed DMBS can be accessed from multiple locations via network, and the system itself can be stored in several locations.



While the access that remote computers have to the system proves to be a detriment to security, it opens up many more options for users aside from the non-local access, along with implementations that can avoid data loss.

Database mirroring – Database mirroring is when one uses the network splitting capabilities of a distributed system and stores a copy of the database on multiple network devices:



This method is highly recommended to avoid data loss and very much simplify data recovery, but only if the system has enough resources for that, otherwise it could become a detriment to performance.

V. Conclusion

Enforcing data protection in our DMBS means safeguarding information from unauthorized or improper actions. In this paper we have explored several core principles of reaffirm the defense of a system, such as identification, authentication, authorization and encryption and the danger of not following those principles.

Along with that a focus was given on the options of initializing a database, their overall benefits, and detriments, of making it based in a local or distributed environment and gave examples of how they can be organized to be more secure.

Citations:

Ng, A. (2019, April 03). Millions of Facebook records were exposed on public Amazon server. Retrieved November 20, 2020, from https://www.cnet.com/news/millions-of-facebook-records-were-exposed-on-public-amazon-server

Cimpanu, C. (2019, April 01). Indian govt agency left details of millions of pregnant women exposed online. Retrieved November 20, 2020, from https://www.zdnet.com/article/indian-govt-agency-left-details-of-millions-of-pregnant-women-exposed-online

Stoianov, N. (2019, July 15). От НАП са изтекли лични данни на милиони български граждани и фирми / Millions of personal data belonging to Bulgarian citizens and firms leaked from the NRA. Retrieved November 26, 2020, from https://www.capital.bg/politika_i_ikonomika/bulgaria/2019/07/15/3938624_ot_nap_sa_iztekli_lichni_danni_na_milioni_bulgarski/

Stoianov, N. (2019, July 15). От НАП са изтекли лични данни на милиони български граждани и фирми / Millions of personal data belonging to bulgarian citizens and firms leaked from the NRA. Retrieved November 26, 2020, from https://www.capital.bg/politika_i_ikonomika/bulgaria/2019/07/15/3938624_ot_nap_sa_iztekli_lichni_danni_na_milioni_bulgarski/

Jarosciak, J. (2016, October 08). Local vs Distributed Databases. Retrieved November 26, 2020, from https://www.joe0.com/2016/10/08/local-vs-distributed-databases/

Bertino, E., Jajodia, S., & Samarati, P. (1995, May 16). Database Security: Research and Practice. Retrieved November 26, 2020, from https://www.sciencedirect.com/science/article/abs/pii/0306437995000294