

## **Security issues in distributed database**

Poonam Sharma

Computer Science Department, Amity School of Engineering and Technology,  
Amity University, Haryana, India  
e-mail:poonamsharma.2289@gmail.com

***Abstract: This paper will inspect the basic highlights of the conveyed database engineering. Taking in the errand of circulated database administration framework will lead us to an effective outline. The outline will enhance versatility, availability and adaptability while getting to different kinds of information. Building up a fruitful conveyed database framework requires to address the significance of security issues that may emerge and conceivably trade off the entrance control and the respectability of the framework. We propose a few answers for some security angles, for example, multilevel access control, privacy, unwavering quality, respectability and recuperation that relate to a circulated database framework.***

### **1. Introduction**

A database is an ordered collection of related data that is built for a specific purpose. For example, a company database may include tables for projects, employees, departments, products and financial records. A distributed database is a set of interconnected databases that is distributed over the computer network or internet or a distributed database is a collection of multiple interconnected databases, which are spread physically across various locations that communicate via a computer network. Security means protection of information and information system from unauthorized access, modification and misuse of information. The purpose of distributed database security is to deal with protecting data from people or, software having malicious intension [1].

A distributed system needs additional security measures than centralized system, since there are many users, diversified data, multiple sites and distributed control.

In distributed communication systems, there are two types of intruders –

**Passive eavesdroppers** – They monitor the messages and get hold of private information.

**Active attackers** – They not only monitor the messages but also corrupt data by inserting new data or modifying existing data.

Security measures encompass security in communications, security in data and data auditing.

### **2. Communications or Network Security**

In a distributed database, a lot of data communication takes place owing to the diversified location of data, users and transactions. So, it demands secure communication between users and databases and between the different database environments.

Security in communication encompasses the following –

- Data should not be corrupt during transfer.
- The communication channel should be protected against both passive eavesdroppers and active attackers.
- In order to achieve the above stated requirements, well-defined security algorithms and protocols should be adopted.

Two popular, consistent technologies for achieving end-to-end secure communications are discussed below:

## **2.1 Secure Socket Layer Protocol or Transport Layer Security Protocol**

Secure Sockets Layer (SSL) is a standard security technology for establishing an encrypted link between a server and a client. SSL allows sensitive information or data to be transmitted securely [2]. Normally, data sent between browsers and web servers is sent in plain text—leaving you vulnerable to eavesdropping. If an attacker is able to intercept all data being sent between a browser and a web server, they can see and use that information.

More specifically, SSL is a security protocol. Protocols describe how algorithms should be used. In this case, the SSL protocol determines variables of the encryption for both the link and the data being transmitted.

All browsers have the capability to interact with secured web servers using the SSL protocol. However, the browser and the server need what is called an SSL Certificate to be able to establish a secure connection.

SSL secures millions of peoples' data on the Internet every day, especially during online transactions or when transmitting confidential information. Internet users have come to associate their online security with the lock icon that comes with an SSL-secured website or green address bar that comes with an Extended Validation SSL-secured website.

When a browser attempts to access a website that is secured by SSL, the browser and the web server establish an SSL connection using a process called an "SSL Handshake".

Essentially, three keys are used to set up the SSL connection: the public, private, and session keys. Anything encrypted with the public key can only be decrypted with the private key, and vice versa.

Because encrypting and decrypting with private and public key takes a lot of processing power, they are only used during the SSL Handshake to create a symmetric session key.

After the secure connection is made, the session key is used to encrypt all transmitted data.

Browser connects to a web server (website) secured with SSL (https). Browser requests that the server identify itself. Server sends a copy of its SSL Certificate, including the server's public key.

Browser checks the certificate root against a list of trusted CAs and that the certificate is unexpired, unrevoked, and that its common name is valid for the website that it is connecting to.

If the browser trusts the certificate, it creates, encrypts, and sends back a symmetric session key using the server's public key. Server decrypts the symmetric session key using its private key and sends back an acknowledgement encrypted with the session key to start the encrypted session.

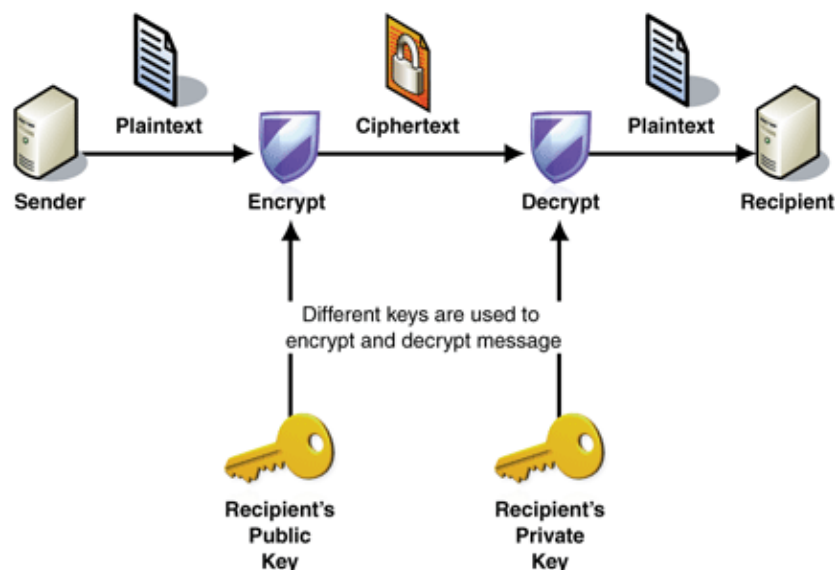


**Fig 1. Encryption process in database**

## **2.2 Virtual Private Networks (VPN).**

It secures your computer's internet connection to guarantee that all of the data you're sending and receiving is encrypted and secured from prying eyes [3].

Remote-access virtual private networks (VPNs) allow secure access by establishing an encrypted tunnel across the Internet. Tunneling, encryption, and authenticated connections these are the three fundamental features that make VPN security work.

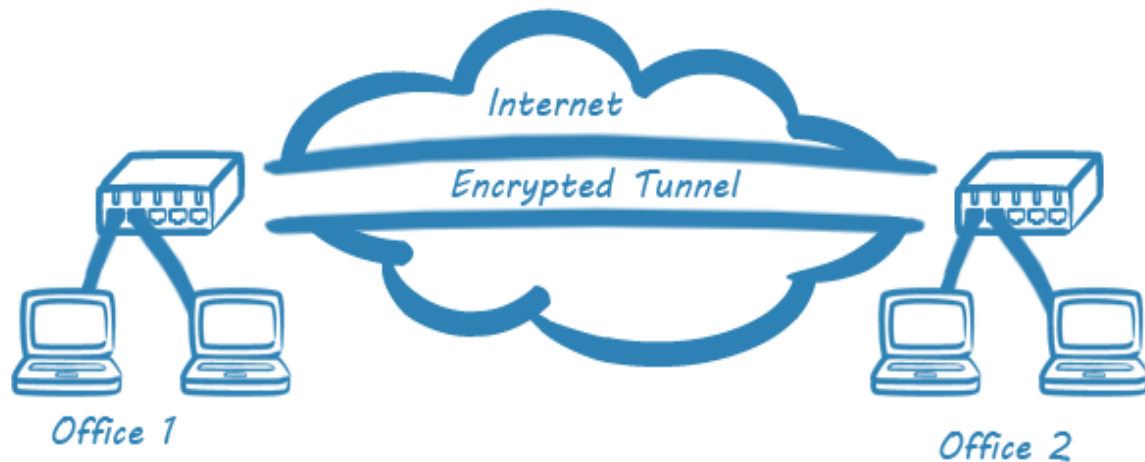


Fig 2. Tunneled connections

### **Tunneled connections**

Tunneling is a process by which data is sent privately over the internet, via a VPN.

As we know that all data transmitted over the internet is split into small pieces called “packets.” Every packet also carries additional information, including the protocol (such as HTTP, Telnet, Bittorrent and so on) it's being used for and the sender's IP address.

On a VPN's tunneled connection, every data packet is placed inside another data packet before it is sent over the internet. The process is called encapsulation.

The outer packet provides a layer of security that keeps the contents safe from public view.

### **Encrypting the packets**

Next layer of security is encryption, where data is encoded so that packets can only be read by your VPN client and server, which are securely connected together.

VPNs can use a number of security protocols to encrypt data. The most common are IPSec (Internet Protocol Security) and OpenVPN.

They work by:

Encrypting each encapsulated data packet's contents with an encryption key. The key is shared only between the VPN's server and clients.

Using a sub-protocol called Encapsulation Header to hide certain packet information, including the sender's identity, during transmission.

### **Advantages:**

Privacy is the real goal of VPN security, privacy is the condition of being free to use and search the web in a manner in which no one is looking over your shoulder, collecting your information or stealing your data.

- VPN secures traffic between your devices and the VPN servers, so hackers and spies can't see your data while it's in transmission.
- Use of VPN Internet connection became private, so it helps to maintain your anonymity and to repel malicious influences including viruses, censors and hackers by allowing you and only you to access your computer.
- VPN provide confidentiality for the users, authentication of senders and message integrity for data transfers. This is done by replacing or altering your routing and other identifying information so others cannot track you, and cannot access your connection or device.
- VPN also provide encryption means that even if your connection is hacked and information is obtained, no one will be able to read it.
- When we use VPN over internet is become totally Anonymity. Anonymity means there is no identifying information on the Internet.

### **3. Database Security**

Data security is an imperative aspect of any database system. It is of particular importance in distributed systems because of large number of users, fragmented and replicated data, multiple sites and distributed control.

#### **3.1 Threats in a Database**

**Availability loss** – Availability loss refers to non-availability of database objects by legitimate users.

**Integrity loss** – Integrity loss occurs when unacceptable operations are performed upon the database either accidentally or maliciously. This may happen while creating, inserting, updating or deleting data. It results in corrupted data leading to incorrect decisions.

**Confidentiality loss** – Confidentiality loss occurs due to unauthorized or unintentional disclosure of confidential information. It may result in illegal actions, security threats and loss in public confidence.

#### **Protection**

##### **Authentication and authorization**

To provide authentication digital certificates are used.

A Digital Signature (DS) is an authentication technique based on public key cryptography. It associates a unique mark to an individual within the body of his message. This helps others to authenticate valid senders of messages.

Typically, a user's digital signature varies from message to message in order to provide security against counterfeiting.

The method is as follows –

- The sender takes a message, calculates the message digest of the message and signs it digest with a private key.
- The sender then appends the signed digest along with the plaintext message.
- The message is sent over communication channel.
- The receiver removes the appended signed digest and verifies the digest using the corresponding public key.
- The receiver then takes the plaintext message and runs it through the same message digest algorithm.
- If the results of step 4 and step 5 match, then the receiver knows that the message has integrity and authentic.

### **Data encryption**

The two approaches for data encryption in distributed systems are –

Internal to distributed database approach: The user applications encrypt the data and then store the encrypted data in the database. For using the stored data, the applications fetch the encrypted data from the database and then decrypt it [4].

External to distributed database: The distributed database system has its own encryption capabilities. The user applications store data and retrieve them without realizing that the data is stored in an encrypted form in the database.

### **Validated input**

In this security measure, the user application checks for each input before it can be used for updating the database. An un-validated input can cause a wide range of exploits like buffer overrun, command injection, cross-site scripting and corruption in data.

## **4. Data Auditing**

A database security system needs to detect and monitor security violations, in order to ascertain the security measures it should adopt. It is often very difficult to detect breach of security at the time of occurrences [5]. One method to identify security violations is to examine audit logs. Audit logs contain information such as –

- Date, time and site of failed access attempts.
- Details of successful access attempts.
- Vital modifications in the database system.
- Access of huge amounts of data, particularly from databases in multiple sites.

## **5. Conclusion**

All the above information gives an insight of the activities in the database. A periodical analysis of the log helps to identify any unnatural activity along with its site and time of occurrence. This log is ideally stored in a separate server so that it is inaccessible to attackers.

### **References**

- [1] Bell, David and Jane Grisom, Distributed Database Systems. Workinham, England: Addison Isley, 1992.
- [2] Charles P. Pfleeger and Shari Lawrence Pfleeger, Security in Computing, Prentice Hall Professional Technical Reference, Upper Saddle River, New Jersey, 2003.
- [3] Haigh, J. T. et al., “The LDV Secure Relational DBMS Model,” In Database Security, IV: Status andProspects, S. Jajodia and C.E. LandIhr eds., pp. 265-269, North Holland: Elsevier, 1991.
- [4] İlker Köse, GYTE, Veri ve Ağ Güvenliği, Distributed Database Security, Spring 2002.
- [5] James F. Kurose and Keith W. Ross, Computer Networking: A Top-Down Approach Featuring the Internet, Pearson Education, Inc, New York, 2003.