COMPUTER SECURITY

PRINCIPLES AND PRACTICE

SECOND EDITION

William Stallings | Lawrie Brown

Chapter 5

Database Security

Databases

- structured collection of data stored for use by one or more applications
- contains the relationships between data items and groups of data items
- can sometimes contain sensitive data that needs to be secured
- database management system (DBMS)
  - suite of programs for constructing and maintaining the database
  - offers ad hoc query facilities to multiple users and applications
- query language
  - provides a uniform interface to the database

DBMS Architecture

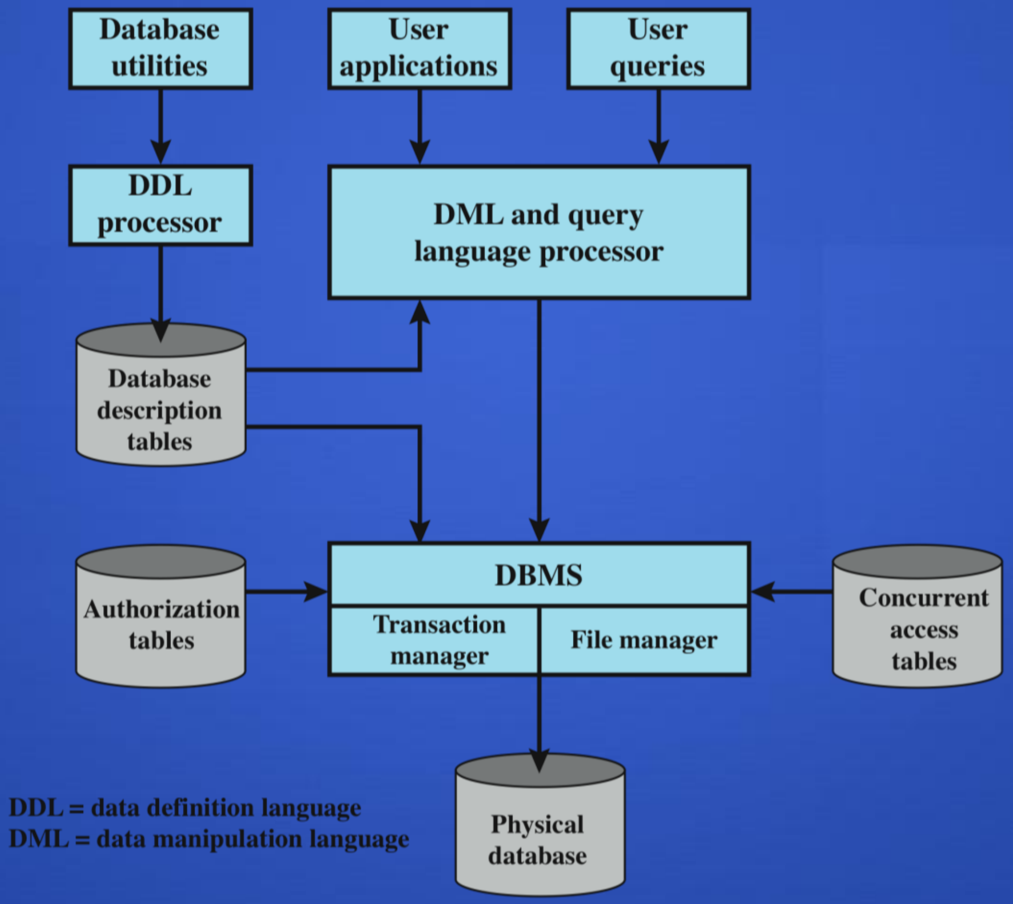


Figure 5.1 DBMS Architecture

Relational Databases

- table of data consisting of rows and columns
  - each column holds a particular type of data
  - each row contains a specific value for each column
  - ideally has one column where all values are unique, forming an identifier/key for that row
- enables the creation of multiple tables linked together by a unique identifier that is present in all tables
- use a relational query language to access the database
  - allows the user to request data that fit a given set of criteria

Figure 5.2

Relational Database Example

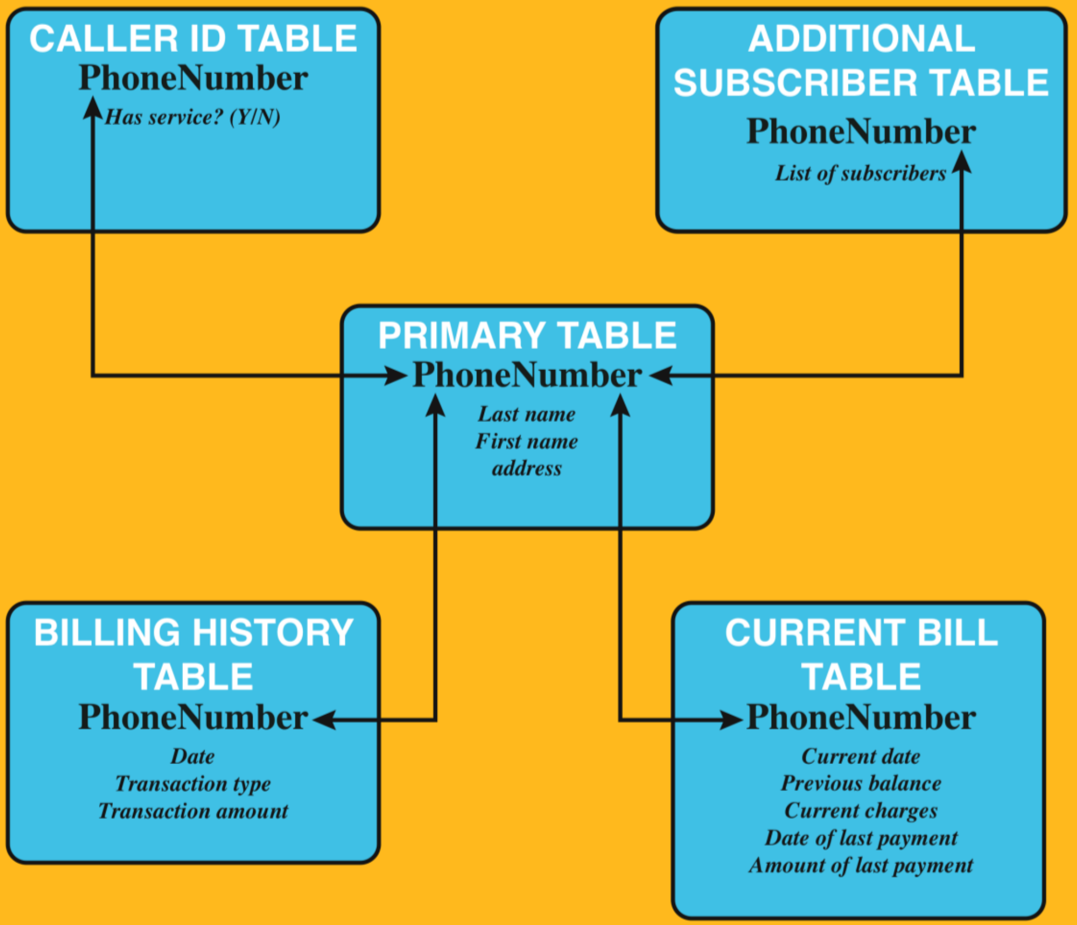


Figure 5.2 Example Relational Database Model. A relational database uses multiple tables related to one another by a designated key; in this case the key is the PhoneNumber field.

Relational Database Elements

- **relation / table / file**
- **tuple / row / record**
- **attribute / column / field**

primary key

- uniquely identifies a row
- consists of one or more column names

foreign key

- links one table to attributes in another

view / virtual table

- result of a query that returns selected rows and columns from one or more tables

Figure 5.3

Relational Database Example

Department Table

Did	Dname	Dacctno
4	human resources	528221
8	education	202035
9	accounts	709257
13	public relations	755827
15	services	223945

primary key (Did)

Employee Table

Ename	Did	Salarycode	Eid	Ephone
Robin	15	23	2345	6127092485
Neil	13	12	5088	6127092246
Jasmine	4	26	7712	6127099348
Cody	15	22	9664	6127093148
Holly	8	23	3054	6127092729
Robin	8	24	2976	6127091945
Smith	9	21	4490	6127099380

foreign key (Did); primary key (Eid)

(a) Two tables in a relational database

Dname	Ename	Eid	Ephone
human resources	Jasmine	7712	6127099348
education	Holly	3054	6127092729
education	Robin	2976	6127091945
accounts	Smith	4490	6127099380
public relations	Neil	5088	6127092246
services	Robin	2345	6127092485
services	Cody	9664	6127093148

(b) A view derived from the database

Figure 5.3 Relational Database Example

Structured Query Language (SQL)

- originally developed by IBM in the mid-1970s
- standardized language to define, manipulate, and query data in a relational database
- several similar versions of ANSI/ISO standard

SQL statements can be used to:
• create tables • insert and delete data in tables • create views • retrieve data with query statements

Database Access Control

database access control system determines:

- if the user has access to the entire database or just portions of it
- what access rights the user has (create, insert, delete, update, read, write)

can support a range of administrative policies

centralized administration

- small number of privileged users may grant and revoke access rights

ownership-based administration

- the creator of a table may grant and revoke access rights to the table

decentralized administration

- the owner of the table may grant and revoke authorization rights to other users, allowing them to grant and revoke access rights to the table

SQL Access Controls

- two commands for managing access rights:
  - grant
    - used to grant one or more access rights or can be used to assign a user to a role
  - revoke
    - revokes the access rights
- typical access rights are:
  - select, insert, update, delete, references

Cascading Authorizations

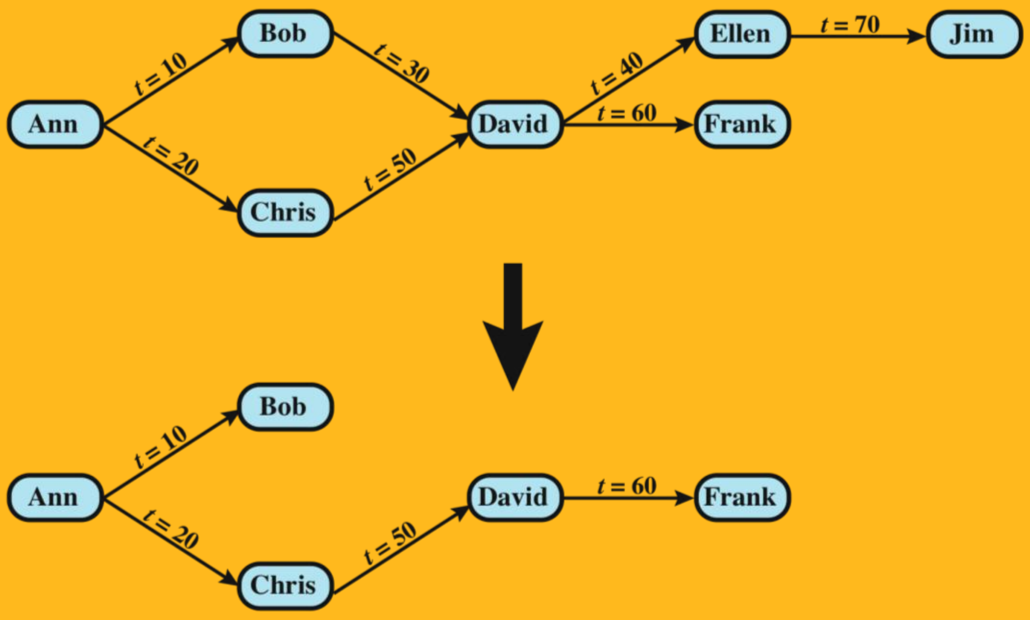


Figure 5.4 Bob Revokes Privilege from David

Role-Based Access Control (RBAC)

- **role-based access control eases administrative burden and improves security**

categories of database users:

- application owner
  - an end user who owns database objects as part of an application
- end user
  - an end user who operates on database objects via a particular application but does not own any of the database objects
- administrator
  - user who has administrative responsibility for part or all of the database

- **a database RBAC needs to provide the following capabilities:**
  - **create and delete roles**
  - **define permissions for a role**
  - **assign and cancel assignment of users to roles**

Table 5.2

Fixed Roles in Microsoft SQL Server

Role	Permissions
Fixed Server Roles	
sysadmin	Can perform any activity in SQL Server and have complete control over all database functions
serveradmin	Can set server-wide configuration options, shut down the server
setupadmin	Can manage linked servers and startup procedures
securityadmin	Can manage logins and CREATE DATABASE permissions, also read error logs and change passwords
processadmin	Can manage processes running in SQL Server
dbcreator	Can create, alter, and drop databases
diskadmin	Can manage disk files
bulkadmin	Can execute BULK INSERT statements
Fixed Database Roles	
db_owner	Has all permissions in the database
db_accessadmin	Can add or remove user IDs
db_datareader	Can select all data from any user table in the database
db_datawriter	Can modify any data in any user table in the database
db_ddladmin	Can issue all Data Definition Language (DDL) statements
db_securityadmin	Can manage all permissions, object ownerships, roles and role memberships
db_backupoperator	Can issue DBCC, CHECKPOINT, and BACKUP statements
db_denydatareader	Can deny permission to select data in the database
db_denydatawriter	Can deny permission to change data in the database

Inference

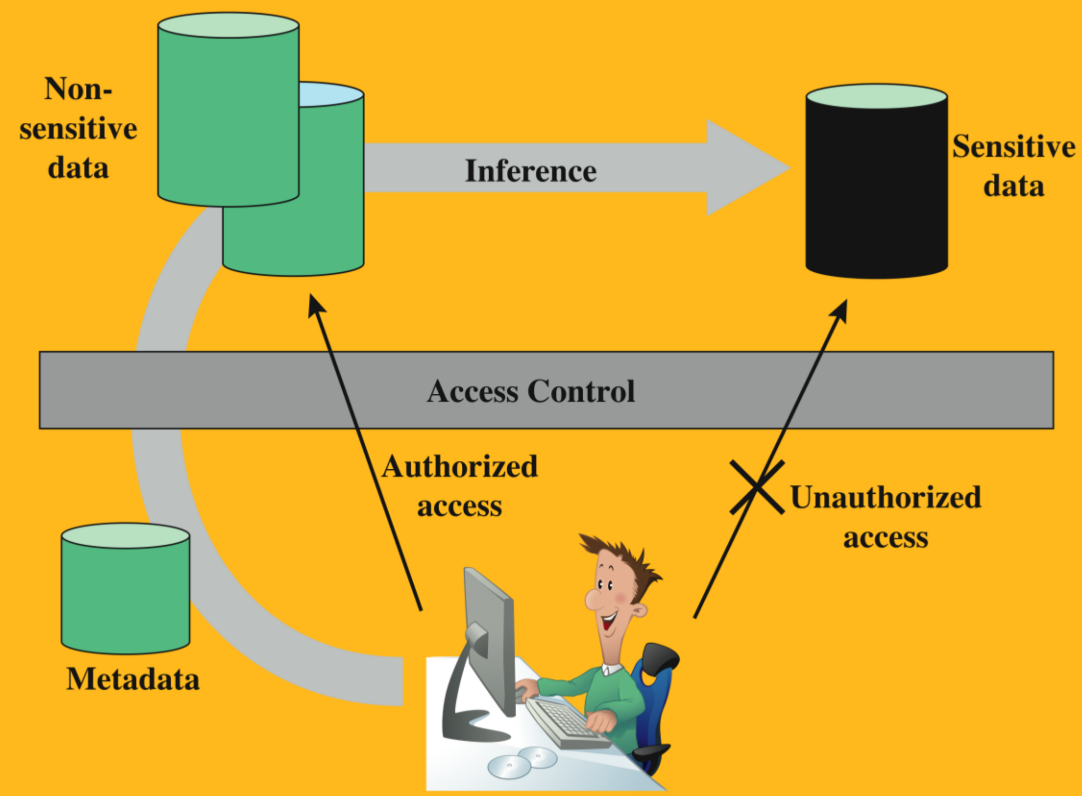


Figure 5.5 Indirect Information Access Via Inference Channel

- the process of performing queries and deducing unauthorized information from the legitimate responses received
- *inference channel* is the information transfer path by which unauthorized data is obtained

Inference Example

Name	Position	Salary ($)	Department	Dept. Manager
Andy	senior	43,000	strip	Cathy
Calvin	junior	35,000	strip	Cathy
Cathy	senior	48,000	strip	Cathy
Dennis	junior	38,000	panel	Herman
Herman	senior	55,000	panel	Herman
Ziggy	senior	67,000	panel	Herman

(a) Employee table

Position	Salary ($)
senior	43, 000
junior	35,000
senior	48,000

Name	Department
Andy	strip
Calvin	strip
Cathy	strip

(b) Two views

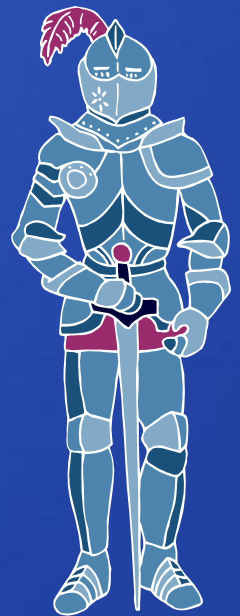
Name	Position	Salary ($)	Department
Andy	senior	43,000	strip
Calvin	junior	35,000	strip
Cathy	senior	48,000	strip

(c) Table derived from combining query answers

Figure 5.6 Inference Example

Inference Countermeasures

- inference detection at database design
  - alter the database structure or change the access control regime
- inference detection at query time
  - monitor and alter or reject the query
- an inference detection algorithm is needed for either approach
  - difficult
  - subject of ongoing research



Statistical Databases (SDB)

- provides data of a statistical nature such as counts and averages
- two types:
  - pure statistical database
    - only stores statistical data
  - ordinary database with statistical access
    - contains individual entries
    - uses DAC, MAC, and RBAC
- access control objective is to provide users with the needed information without compromising the confidentiality of the database
- security problem is one of inference

Abstract Model of a Relational Database

Attributes

Records

	A_1	• • •	A_j	• • •	A_M
1	x_{11}	• • •	x_{1j}	• • •	x_{1M}
•	•		•		•
•	•		•		•
•	•		•		•
i	x_{i1}	• • •	x_{ij}	• • •	x_{iM}
•	•		•		•
•	•		•		•
•	•		•		•
N	x_{N1}	• • •	x_{Nj}	• • •	x_{NM}

Figure 5.7 Abstract Model of a Relational Database

Table 5.3

Statistical Database Example

(a) Database with Statistical Access with $N = 13$ Students

Name	Sex	Major	Class	SAT	GP
Allen	Female	CS	1980	600	3.4
Baker	Female	EE	1980	520	2.5
Cook	Male	EE	1978	630	3.5
Davis	Female	CS	1978	800	4.0
Evans	Male	Bio	1979	500	2.2
Frank	Male	EE	1981	580	3.0
Good	Male	CS	1978	700	3.8
Hall	Female	Psy	1979	580	2.8
Iles	Male	CS	1981	600	3.2
Jones	Female	Bio	1979	750	3.8
Kline	Female	Psy	1981	500	2.5
Lane	Male	EE	1978	600	3.0
Moore	Male	CS	1979	650	3.5

(b) Attribute Values and Counts

Attribute A_j	Possible Values	$\lvert A_j \rvert$
Sex	Male, Female	2
Major	Bio, CS, EE, Psy, …	50
Class	1978, 1979, 1980, 1981	4
SAT	310, 320, 330, …, 790, 800	50
GP	0.0, 0.1, 0.2, …, 3.9, 4.0	41

Statistical Database Security

- use a characteristic formula C
  - a logical formula over the values of attributes
  - e.g. (*Sex*=Male) AND ((*Major*=CS) OR (*Major*=EE))
- query set X(*C*) of characteristic formula *C*, is the set of records matching C
- a statistical query is a query that produces a value calculated over a query set

Table 5.4

Some Queries of a Statistical Database

Name	Formula	Description
count(*C*)	$\lvert X(C) \rvert$	Number of records in the query set
sum(*C*, A_j)	$\sum_{i \in X(C)} x_{ij}$	Sum of the values of numerical attribute A_j over all the records in $X(C)$
rfreq(*C*)	$\frac{\text{count}(C)}{N}$	Fraction of all records that are in $X(C)$
avg(*C*, A_j)	$\frac{\text{sum}(C, A_j)}{\text{count}(C)}$	Mean value of numerical attribute A_j over all the records in $X(C)$
median(*C*, A_j)		The $\lceil \lvert X(C) \rvert / 2 \rceil$ largest value of attribute over all the records in $X(C)$. Note that when the query set size is even, the median is the smaller of the two middle values. $\lceil x \rceil$ denotes the smallest integer greater than *x*.
max(*C*, A_j)	$\text{Max}_{i \in X(C)} (x_{ij})$	Maximum value of numerical attribute A_j over all the records in $X(C)$
min(*C*, A_j)	$\text{Min}_{i \in X(C)} (x_{ij})$	Minimum value of numerical attribute A_j over all the records in $X(C)$

Note: C = a characteristic formula, consisting of a logical formula over the values of attributes.
$X(C)$ = query set of C, the set of records satisfying C.

Protecting Against Inference

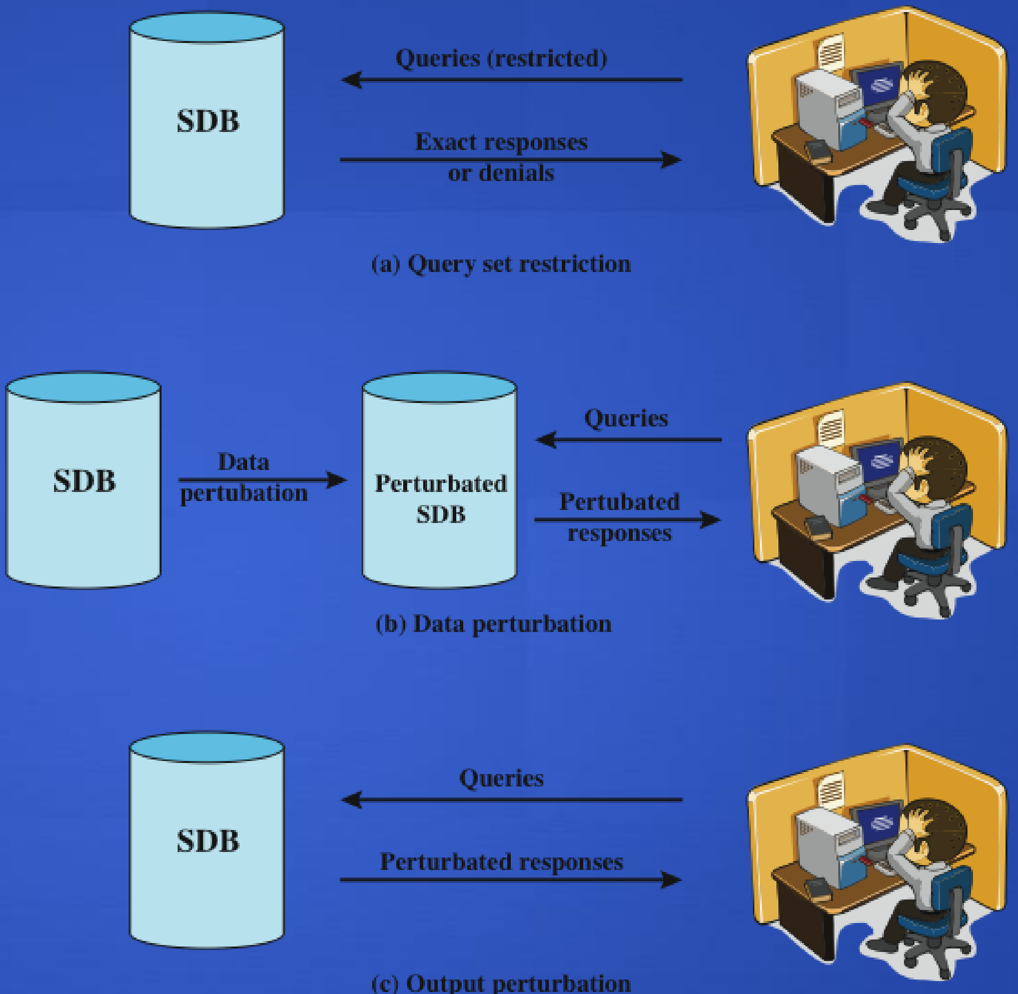


Figure 5.8 Approaches to Statistical Database Security
(based on [ADAM89])

Tracker Attacks

- divide queries into parts
  - C = C1.C2
  - count(C.D) = count(C1) - count (C1. ~*C2)*
- combination is called a tracker
- each part acceptable query size
- overlap is desired result

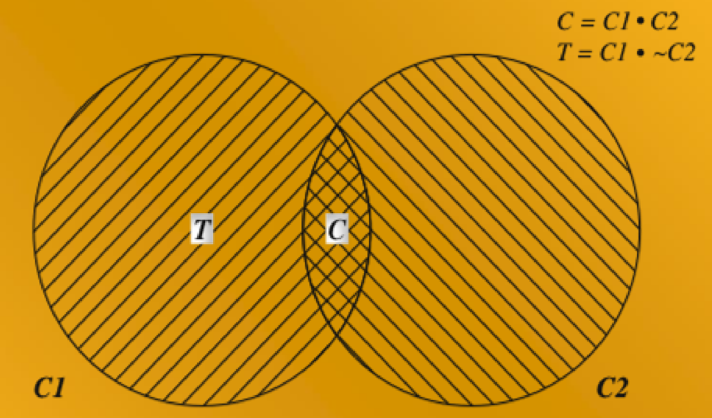


Figure 5.9 Example of Tracker

Other Query Restrictions

- query set overlap control
  - limit overlap between new and previous queries
  - has a number of problems
- partitioning
  - cluster records into a number of mutually exclusive groups
  - query the statistical properties of each group as a whole
- query denial and information leakage
  - denials can leak information
  - to counter must track queries from user

Perturbation

- add noise to statistics generated from original data
- data perturbation technique
  - data can be modified to produce statistics that cannot be used to infer values for individual records
- output perturbation technique
  - system generates statistics that are modified from those that the original database would provide
  - random-sample query
- goal is to minimize the differences between original results and perturbed results
- main challenge is to determine the average size of the error to be used

Data Perturbation Techniques: Data Swapping

Table 5.6 Example of Data Swapping

	D			D'		
Record	**Sex**	**Major**	**GP**	**Sex**	**Major**	**GP**
1	Female	Bio	4.0	Male	Bio	4.0
2	Female	CS	3.0	Male	CS	3.0
3	Female	EE	3.0	Male	EE	3.0
4	Female	Psy	4.0	Male	Psy	4.0
5	Male	Bio	3.0	Female	Bio	3.0
6	Male	CS	4.0	Female	CS	4.0
7	Male	EE	4.0	Female	EE	4.0
8	Male	Psy	3.0	Female	Psy	3.0

Database Encryption

- the database is typically the most valuable information resource for any organization
  - protected by multiple layers of security
    - firewalls, authentication, O/S access control systems, DB access control systems, database encryption
- encryption is often implemented with particularly sensitive data
  - can be applied to the entire database at the record level, the attribute level, or level of the individual field
- disadvantages to encryption:
  - key management
  - inflexibility

Database Encryption

Data owner – organization that produces data to be made available for controlled release

User – human entity that presents queries to the system

Client – frontend that transforms user queries into queries on the encrypted data stored on the server

Server – an organization that receives the encrypted data from a data owner and makes them available for distribution to clients

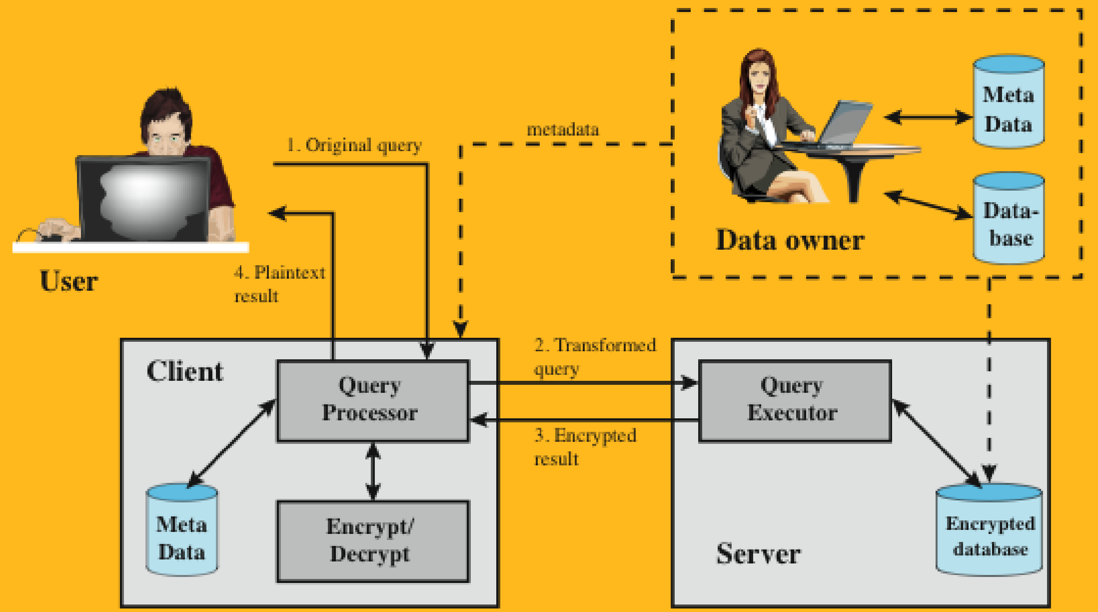


Figure 5.10 A Database Encryption Scheme

Encryption Scheme for Database of Figure 5.7

$E(k, B_1)$	I_{11}	• • •	I_{1j}	• • •	I_{1M}
•	•		•		•
•	•		•		•
•	•		•		•
$E(k, B_i)$	I_{i1}	• • •	I_{ij}	• • •	I_{iM}
•	•		•		•
•	•		•		•
•	•		•		•
$E(k, B_N)$	I_{N1}	• • •	I_{Nj}	• • •	I_{NM}

$B_i = (x_{i1} \,\|\, x_{i2} \,\|\, \ldots \,\|\, x_{iM})$

Figure 5.11 Encryption Scheme for Database of Figure 5.7

Table 5.7 Encrypted Database Example

(a) Employee Table

eid	ename	salary	addr	did
23	Tom	70K	Maple	45
860	Mary	60K	Main	83
320	John	50K	River	50
875	Jerry	55K	Hopewell	92

(b) Encrypted Employee Table with Indexes

E(*k*, *B*)	**I(eid)**	**I(ename)**	**I(salary)**	**I(addr)**	**I(did)**
1100110011001011…	1	10	3	7	4
0111000111001010…	5	7	2	7	8
1100010010001101…	2	5	1	9	5
0011010011111101…	5	5	2	4	9

Cloud Security

NIST defines cloud computing as follows [MELL11]:

> "A model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model promotes availability and is composed of five essential characteristics, three service models, and four deployment models."

Cloud Computing Elements

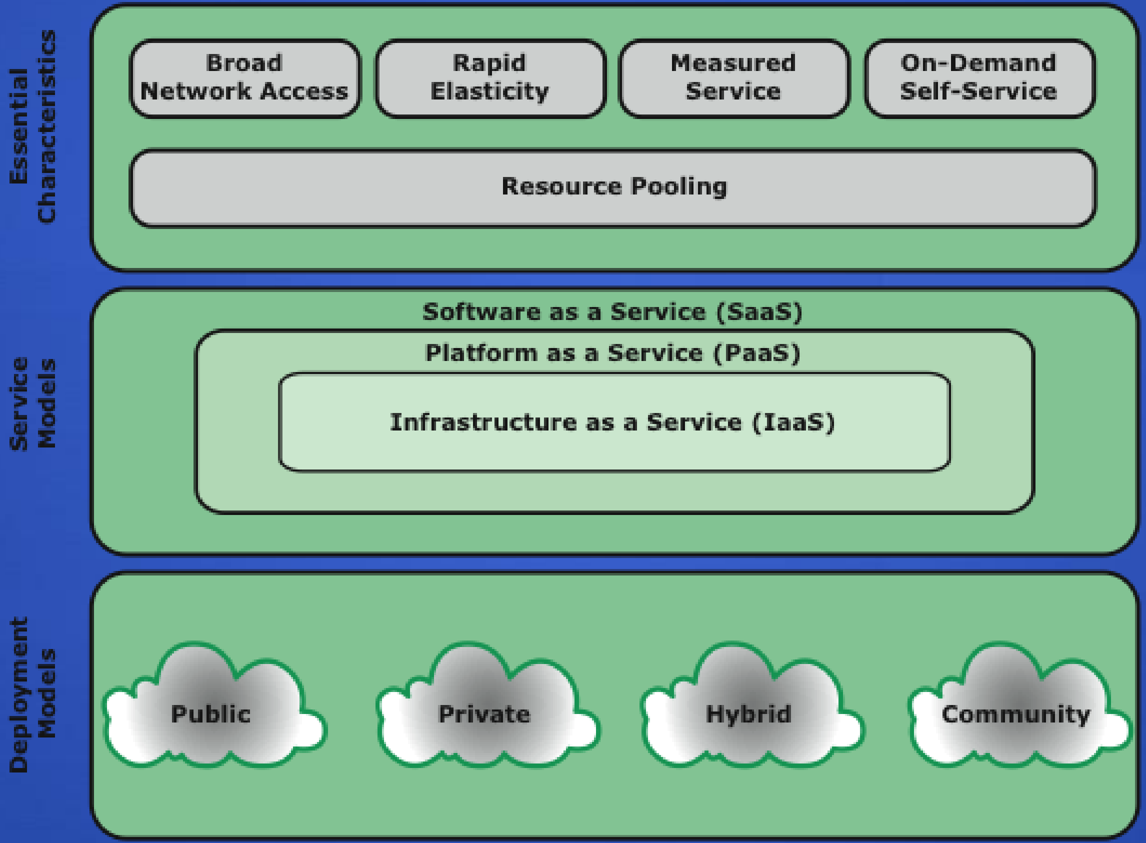


Figure 5.12 Cloud Computing Elements

Figure 5.13

Cloud Computing Context

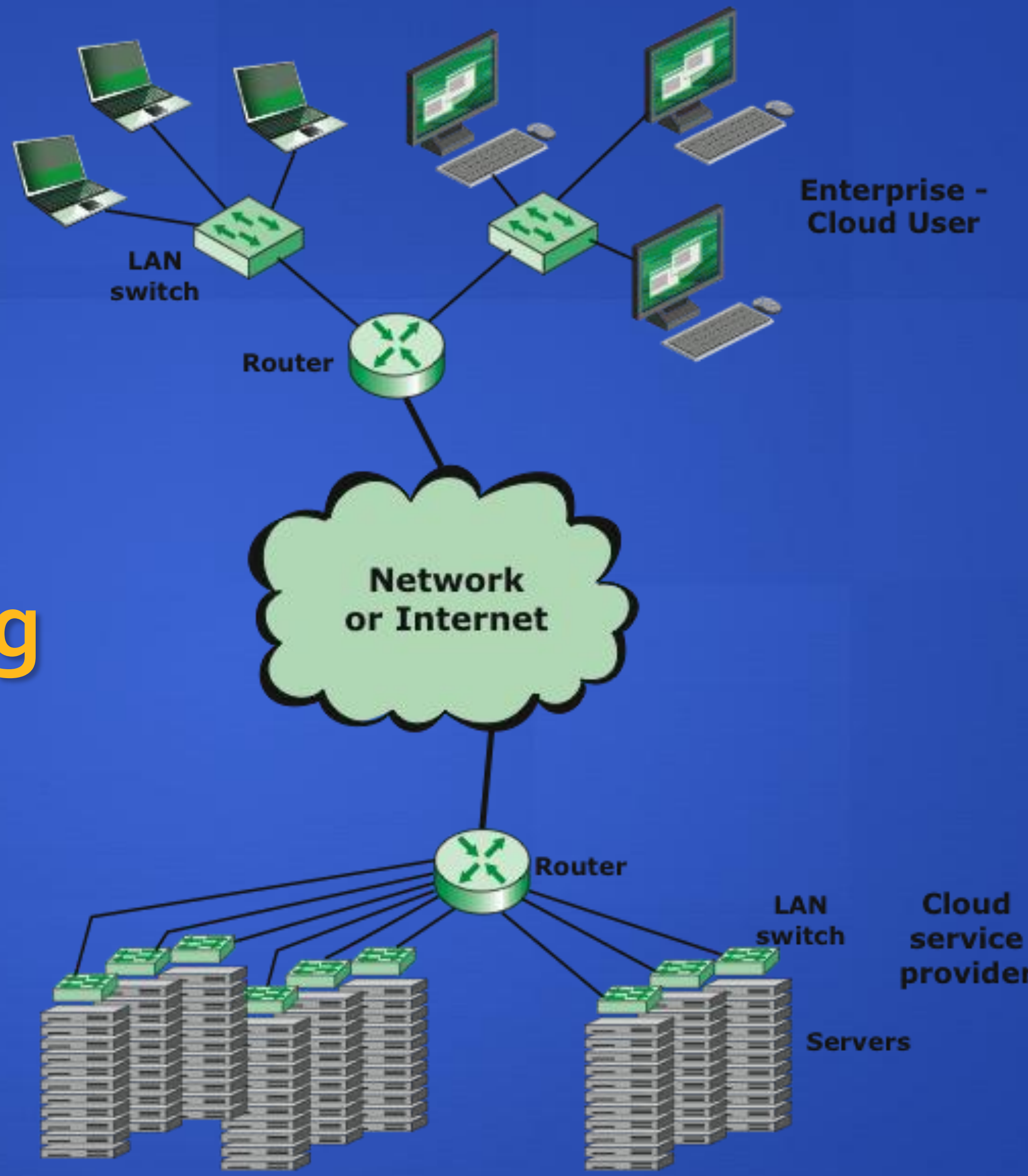


Figure 5.13 Cloud Computing Context

Cloud Security Risks

The Cloud Security Alliance (CSA10) lists the following as the top cloud specific security threats:

- abuse and nefarious use of cloud computing
- insecure interfaces and APIs
- malicious insiders
- shared technology issues
- data loss or leakage
- account or service hijacking
- unknown risk profile

Data Protection in the Cloud

the threat of data compromise increases in the cloud

- risks and challenges that are unique to the cloud
- architectural or operational characteristics of the cloud environment

multi-instance model

- provides a unique DBMS running on a virtual machine instance for each cloud subscriber
- gives the subscriber complete control over administrative tasks related to security

multi-tenant model

- provides a predefined environment for the cloud subscriber that is shared with other tenants typically through tagging data with a subscriber identifier
- gives the appearance of exclusive use of the instance but relies on the cloud provider to establish and maintain a secure database environment

Summary

- database
  - structured collection of data
- database management system (DBMS)
  - programs for constructing and maintaining the database
- structured query language (SQL)
  - language used to define schema/manipulate/query data in a relational database
- relational database
  - table of data consisting of rows (tuples) and columns (attributes)
  - multiple tables tied together by a unique identifier that is present in all tables
- database access control
  - centralized/ownership-based/decentralized administration
- role-based access control (RBAC)
  - application owner/end user other than application owner/administrator
- inference channel
  - information transfer path by which unauthorized data is obtained
- statistical database (SDB)
  - query restriction/perturbation/data swapping/random-sample query
- database encryption
- cloud computing/security/ data protection
  - multi-instance/ multi-tenant model