Distributed database

Logically connected single databases dispersed on different sites, notes or servers, although sites are connected through communication networks.

Middleware

Is specially designed software where two different programs can interact with each other. It allows several processes to be runa and allows data to pass across a network.

It is transparent as it is not seen by the client and runs in the background.

Examples could be:

Two and three client servers

Object request brokers

Skeleton and stubs.

Replication

Is a process in which data is cloned in multiple servers and locations to improve the availability of data. It speeds up the query evaluation.

Types are of data replications are:

Synchronous replication: no difference between original and cloned data.

Asynchronous replication: Data will be modified after commit is fired on db.

Replicas

Data cluster which is replicated.

Fragmentation

The process of dividing the database into smaller multiple parts is called fragmentation. Fragments are also stored in different locations. The data fragmentation should be carried in such a way that the reconstruction of adata could be

possible.Fragments are also created for performance and reliability reasons.
Data fragmentation could be.
Horizontal data fragmentation: subset of columns
Vertical data fragmentation: subset of rows.

Transparency
Transparency means that users should be made unaware of behind operations of a particular database.

The three dimensions of distribution transparency are -

Location transparency: hide location

Fragmentation transparency: conceal the fragment's existence.

Replication transparency: hide the replication of the database.

Concurrent processing
Concurrent processing is sometimes similar with parallel processing. Tasks are broken down into subtasks and then assigned to separate processors to perform simultaneously, instead of sequentially as they would have to be carried out by a single processor.

Distributed db vs Centralized db
Centralized db:
1. Single db is located at one site on network.
2. Since there is only a database it is easier to get a complete view.
3. Slower because all users are accessing the same node.

Distributed db:
1. Contain 2 or more files located at different sites on network

2. Users won't be able to communicate with each other.
3. Much faster because data is received from the nearest node.