Data Replication in Distributed System

Data replication is the process in which the data is copied at multiple locations (diff comp or servers) to improve the availability of data.

Goals of data replication

- Increase the availability of data
- speed up the query evaluation.

Types of data replication:-

① Synchronous Replication :- The replica will be modified

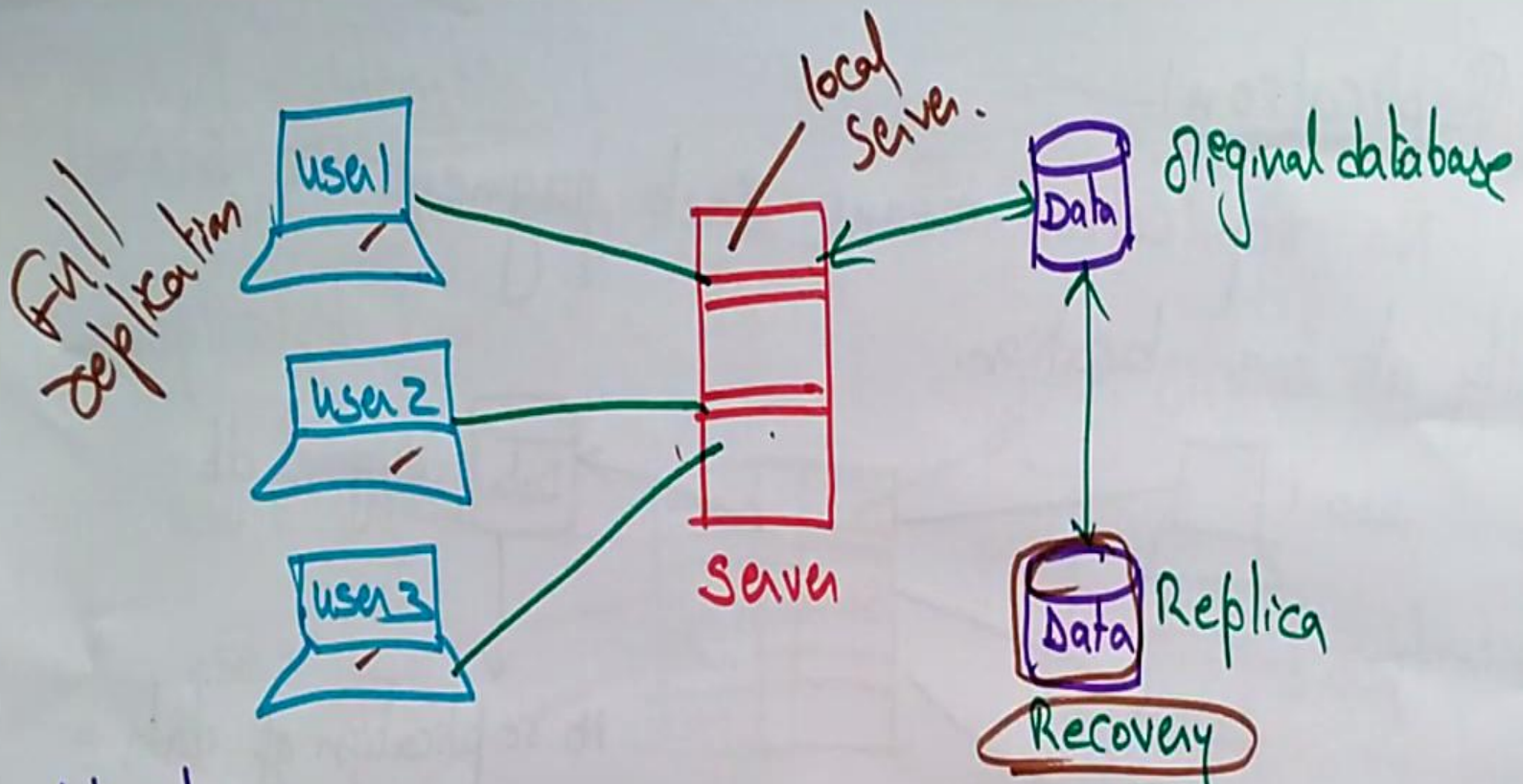
immediately after some changes are made in the relation table. So, there is no difference between original data and replica.

② Asynchronous replication:-

The replica will be modified after commit is fired on the db.

Replication Schemes :- (3 schemes)

① Full replication :- The db is available to almost every location or user in communication n/w.



Advantages :-

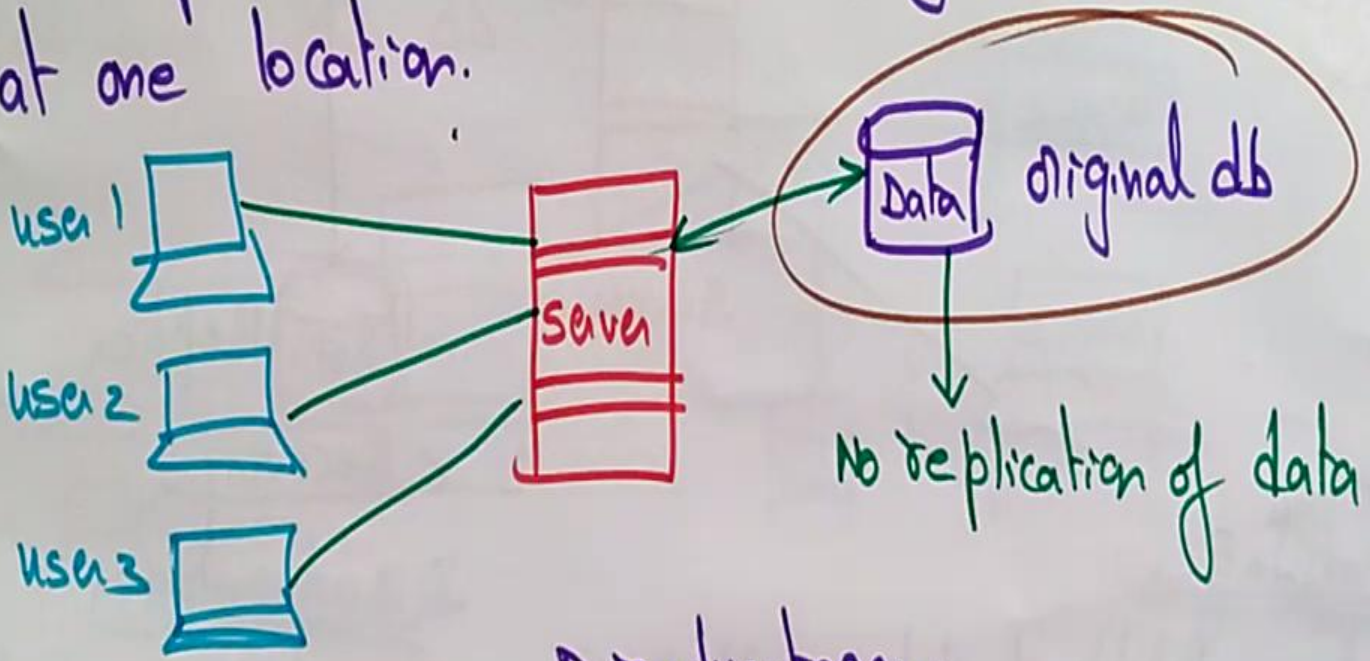
1. High availability of data, as db is available to almost every location
2. Faster execution of queries

Disadvantages

1. Concurrency control is difficult to achieve in full replication
2. update operation is slower

② No Replication :-

No replication means, Each fragment is stored exactly at one location.



Advantages

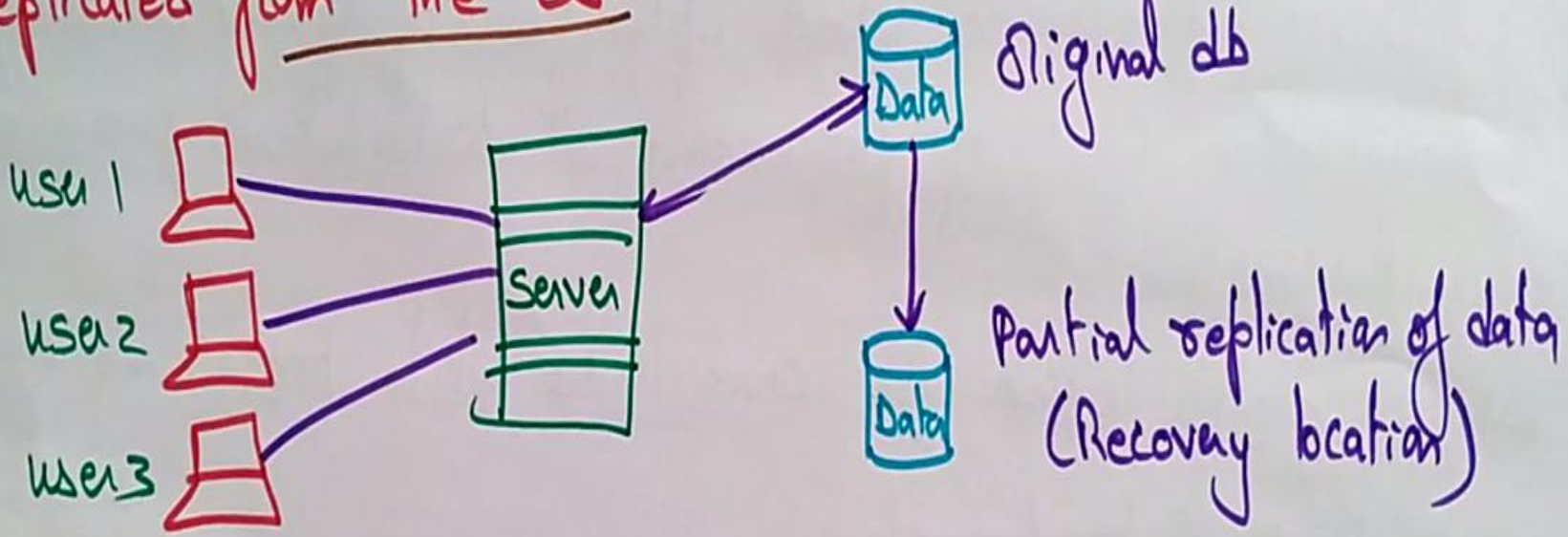
- Concurrency can be minimized
- Easy recovery of data

Disadvantages :-

- Poor availability of data
- Slows down the query execution process as multiple clients accessing same server.

③ Partial replication :-

It means only some & fragments are replicated from the db.



Advantages :-

→ No: of replicas created for ~~data~~ fragments depend upon the importance of data in that fragment.