

Redundancy and Replication

Why Redundancy and Replication?

Redundancy and replication are essential for maintaining availability and data integrity. Let's see how.

Faults and data loss is prevalent in distributed systems because of:

1. Commodity Hardware
2. Network Fault

Due to the above two factors, chances of failure of distributed systems are very high because the commodity hardware and network are highly unreliable.

Therefore, distributed systems are **fault-prone systems**.

We need to handle the faults and failures gracefully, known as Partition Tolerance/Failure Tolerance.

Partition/Failure Tolerance can be achieved by:

1. Replication
2. Redundancy

Now, let's understand each term.

What is Redundancy?

Redundancy is simply the duplication of nodes or components so that when a node or component fails, the duplicate node is available to service customers. This is how redundancy helps in maintaining availability, failure recovery or failure management. The idea behind redundancy is to make backup paths speed, efficient and available.

Example: Global Positioning System (GPS)

Types of Redundancy

There are two types of redundancy:

1. Active Redundancy:

Active Redundancy is considered when each unit is operating/active and responding to the action. Multiple nodes are connected to a load balancer, and each unit receives an equal load.

Example: A building requires five generators to power the complete building. Now, for failure tolerance, we can use additional generators according to the principle of redundancy. We can have $n+1$ (6 generators) for supporting single

node failure as even if one generator of all the working six generators fail, the power won't cut. This is how active redundancy works; the system keeps operating even in the case of failure of any node.

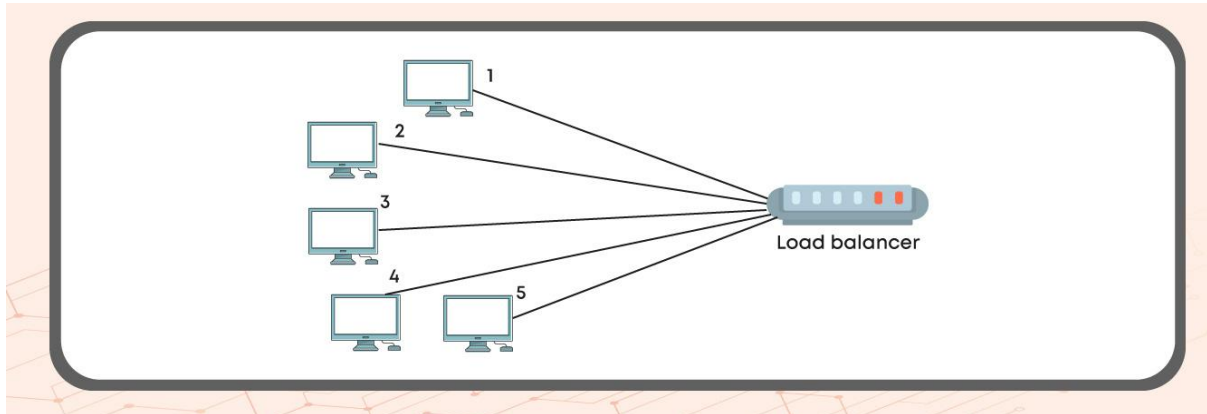


Fig: Active Redundancy

2. Passive Redundancy:

Passive Redundancy is considered when one node is active or operational and the other is not operating. During the breakdown of the active node, the passive node maintains availability by becoming the active node.

Example: A car has four tires and one spare tire. Four of the tires are operational and keep the vehicle moving. In the case of one flat tire, the spare tire can be used, which was earlier acting as a passive component.

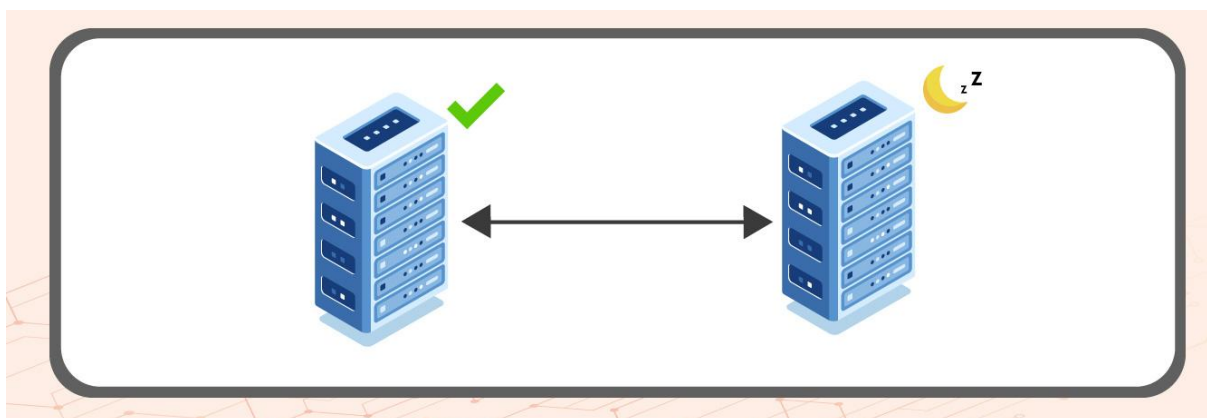


Fig: Passive Redundancy

What is Replication?

Replication is the management process of different data storage where each element is stored in multiple copies hosted on other servers. Simply, it is the copying of data on various machines.

It is the synchronisation of different machines.

Replication helps to ensure consistency between redundant resources to improve fault tolerance and reliability.

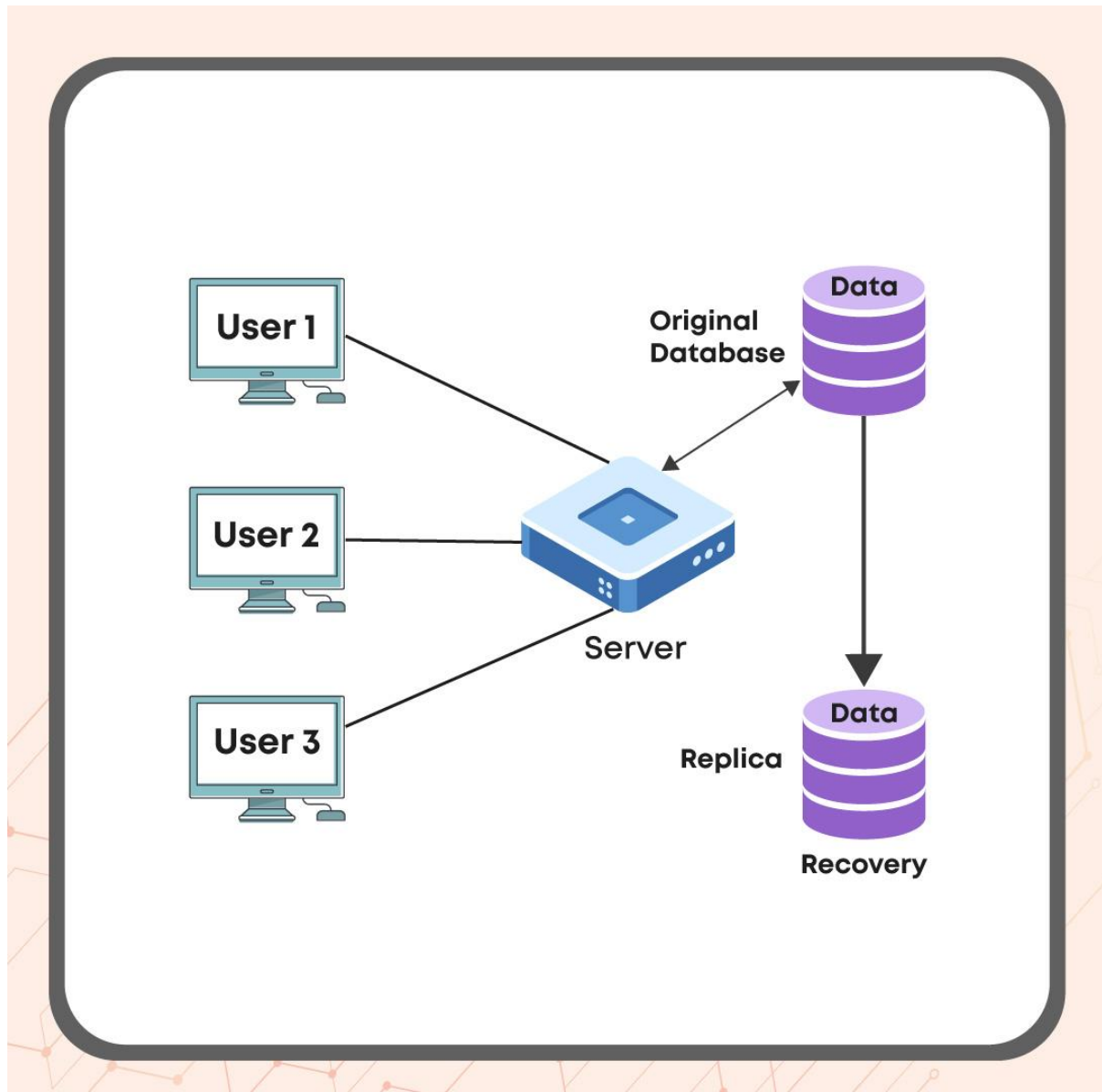


Fig: Replication

Types of Replication

There are two types of Replication.

1. Active Replication:

Active replication is when all the nodes of a cluster are connected, and the data is replicated on every cluster node. It is performed by processing the same request on all the replica machines.

Active replication is used mostly when a single master cannot tolerate all writes, so we need all nodes to accept writes. Then they sync data among themselves. Reads should preferably be below in such cases but could be the same as writes. An example of such a system is "how many times x item was viewed on amazon"

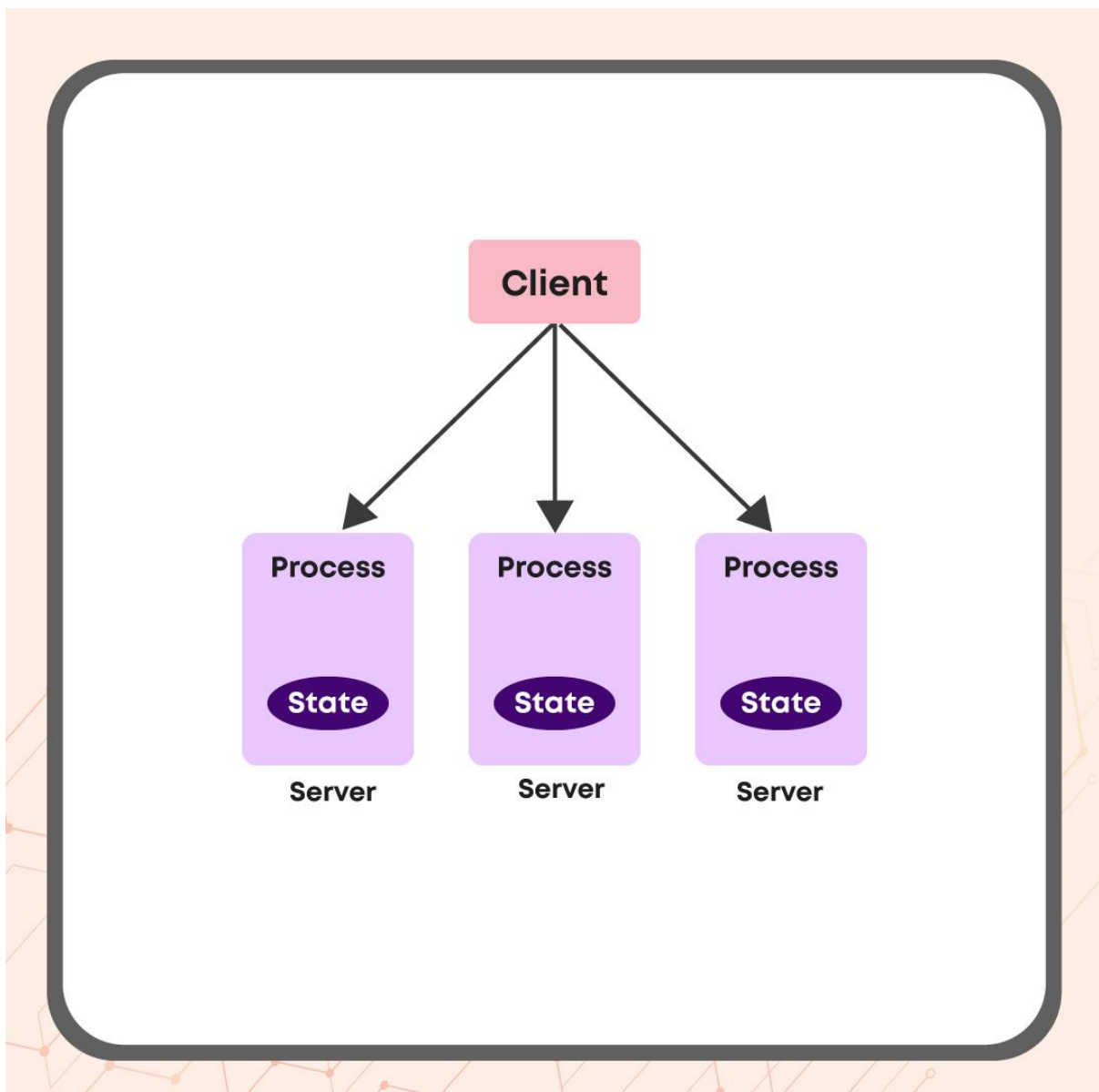


Fig: Active Replication

2. Passive Replication:

Passive replication is when a cluster of nodes has one machine as the master machine and other machines as the slave machine. The read and write operations are completed on the master machine, but the data is replicated on all the slave machines in offline sessions. If the master machine goes down, any of the slave machines can be promoted as the master machine.

Passive replication is used in systems that have low write rates by high read rates. So a single master is enough to accept all writes and can replicate it to all the slaves (sync/async).

Examples of such systems could be Wikipedia.

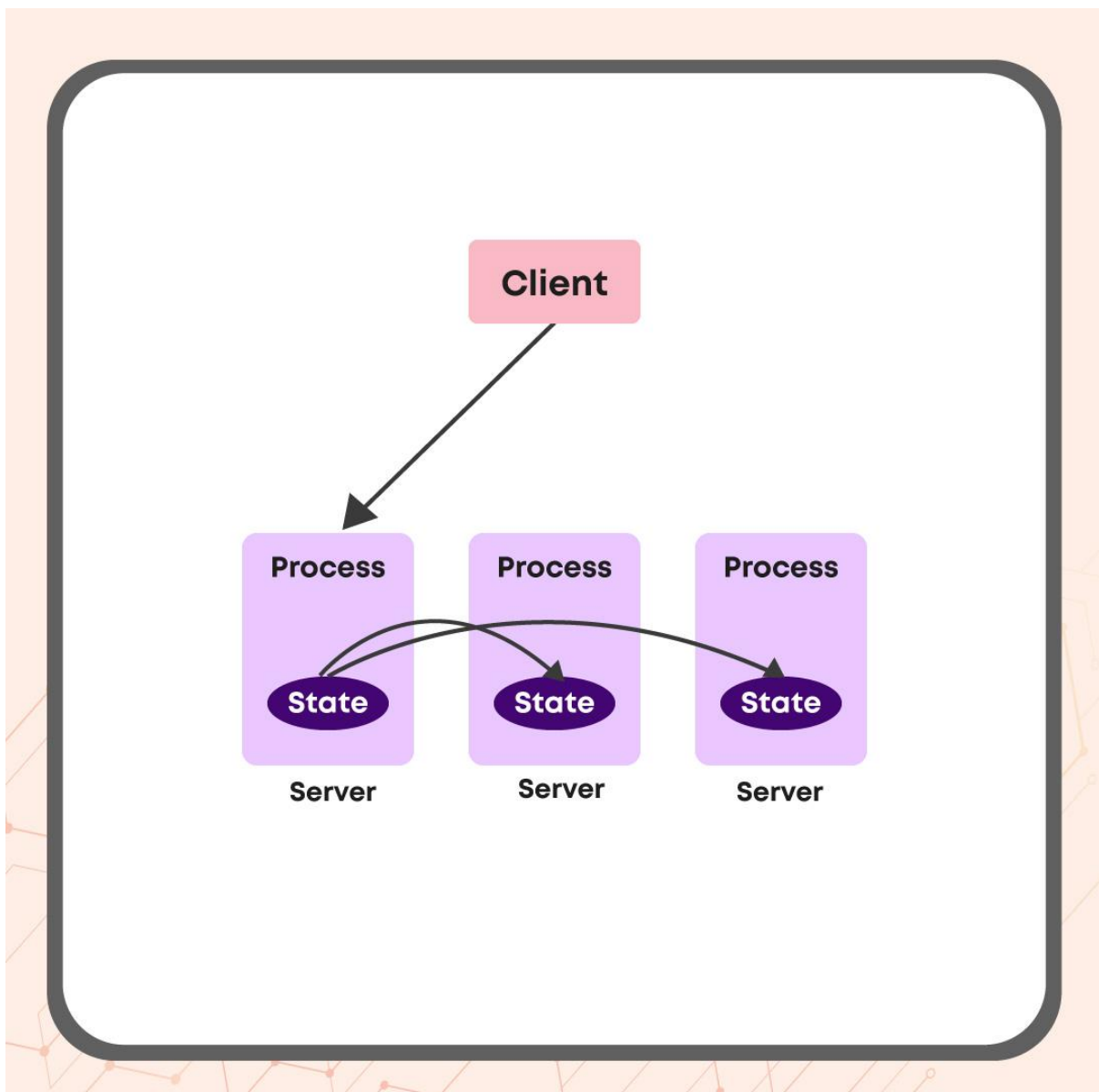


Fig: Passive Replication