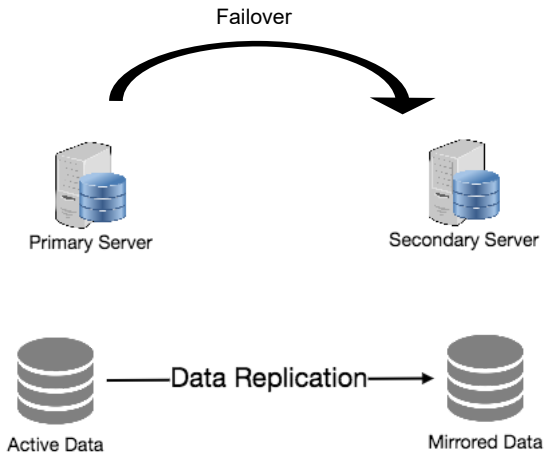
Redundancy and Replication

Redundancy means duplication of critical data or services with the intention of increased reliability of the system. For example, if there is only one copy of a file stored on a single server, then losing that server means losing the file. Since losing data is seldom a good thing, we can create duplicate or redundant copies of the file to solve this problem.

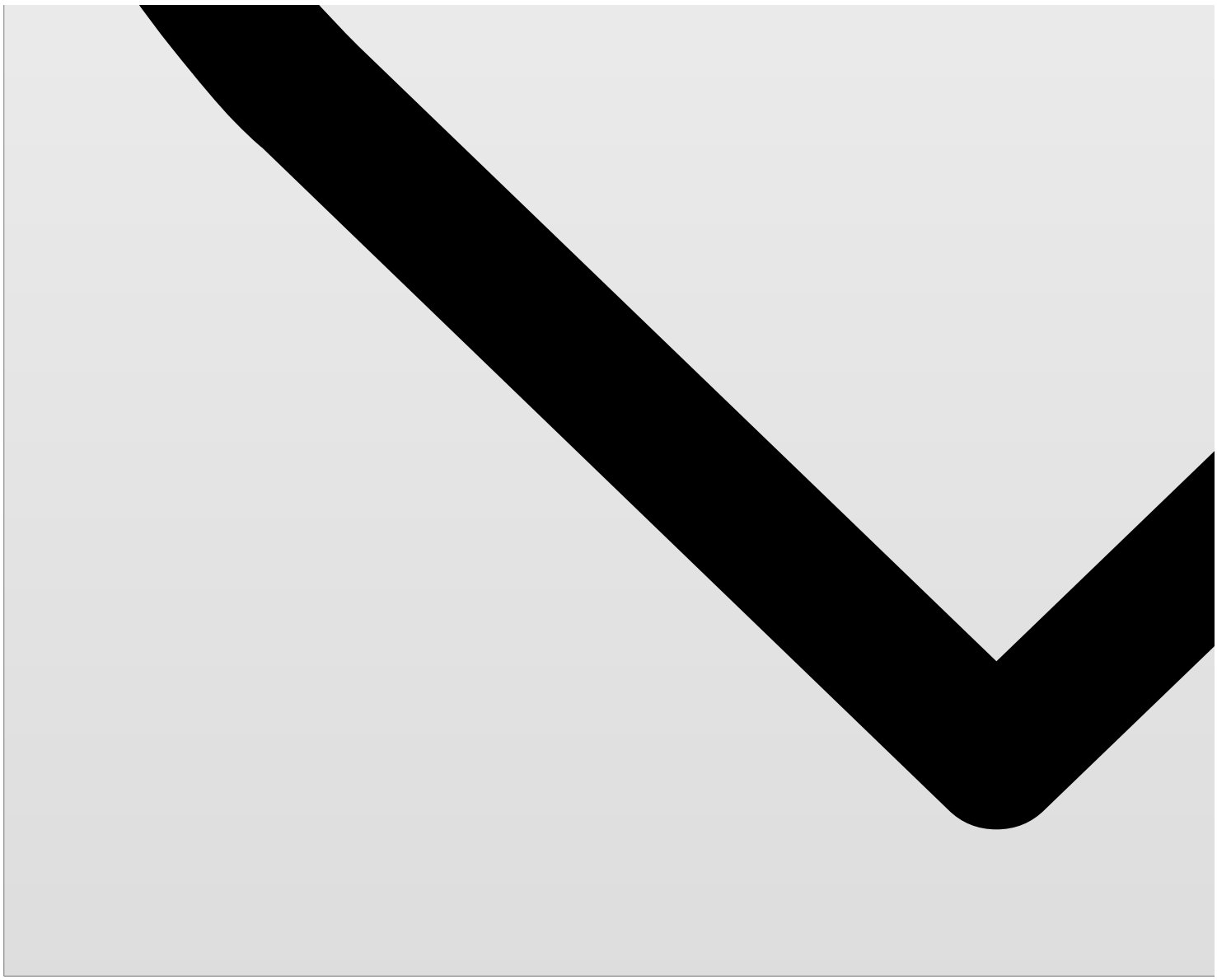
This same principle applies to services too. If we have a critical service in our system, ensuring that multiple copies or versions of it are running simultaneously can secure against the failure of a single node.

Creating redundancy in a system can remove single points of failure and provide backups if needed in a crisis. For example, if we have two instances of a service running in production, and if one fails or degrades, the system can failover to the other one. These failovers can happen automatically or can be done manually.



Another important part of service redundancy is to create a shared-nothing architecture, where each node can operate independently of one another. There should not be any central service managing state or orchestrating activities for the other nodes. This helps a lot with scalability since new servers can be added without special conditions or knowledge and most importantly, such systems are more resilient to failure as there is no single point of failure.

Mark as completed
← *Previous*Queues*Next* →SQL vs. NoSQL
Send feedback or ask a question



13 recommendations

- Home
- Featured
- Team
- Blog
- FAQ
- Terms of Service
- Contact Us