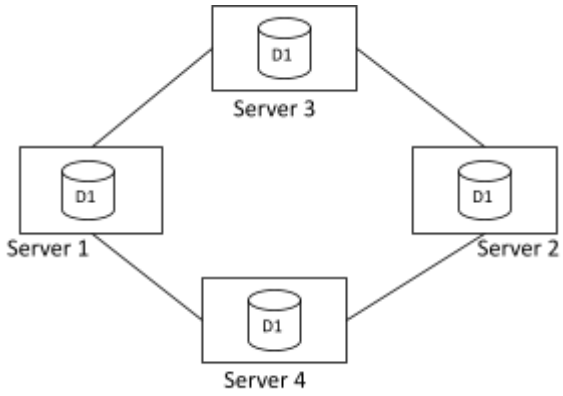
Clustering:
It is the process of combining more than one server or instances holding the same database. Or we can say it is a set of replicated servers.



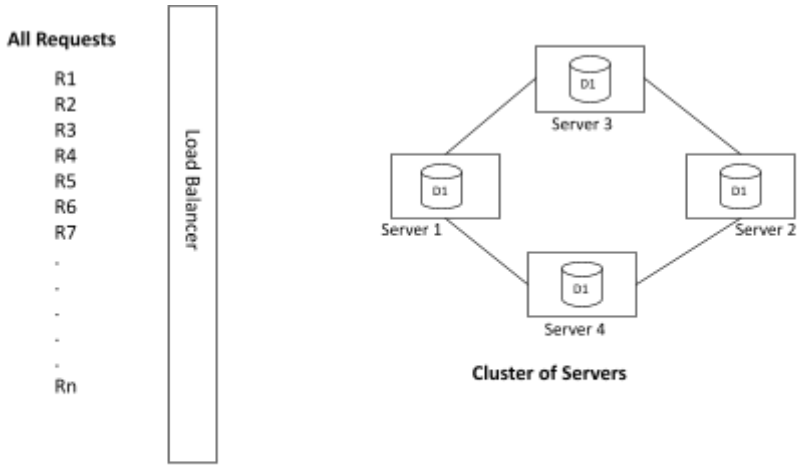
Here, we have cluster of 4 servers, which have been loaded with the same database i.e. D1.

Why do we need Clustering ?
Clustering databases helps servers with a lot of features like,

- **Data Redundancy:** Clustering of databases helps with data redundancy, as we store the same data at multiple servers. Don't confuse this data redundancy as repetition of the same data that might lead to some anomalies or ambiguities. The redundancy that clustering offers is required and is quite certain due to the synchronization.
  In case any of the servers had to face a failure due to any possible reason, the data is available at other servers to access.
  Like in the image above, if the server 3 holding database D1 had any kind of malfunction or issues. Other servers 1,2 and 4 won't be affected and data can be accessed from them.

- **Load Balancing:** Clustering of databases also helps the servers with the load balancing i.e. suppose if there's only one server and there are a lot of requests coming in to access the database then it may become difficult for a server to handle all the requests and the response time for the request will also become huge.

Although, if we introduce more servers holding the same database now the load of the request can be divided among them.
Hence, the response time for the requests decreases too.
So, load balancing refers to the process where the load balancer smartly assign the requests to the servers with the same database.



In the above image, load balancer randomly assigns the ‘N’ requests among the cluster of servers keeping in mind the load they already have, so that the load of requests is equally divided among them and the response time for the request can be improved.

- **Availability:** An Availability of of a database is defined as the time when we can access the database. Now if it wasn’t for clustering, or replication of a database to exist we would have faced this problem very often whenever a database would have been going through a transaction.
  Now as we have clusters of servers available, even if one of the databases is going through a transaction, now the other servers can be used to access the database with the help of load balancer.
  Also, even if a server fails, the database will be available.
  Hence, due to clustering the databases have high availability.

Types of Database Clusters:

On the basis of the requirement of the system there are multiple type of cluster architecture present in the market:

- High-Performance Cluster
- Load Balancing Clusters
- Failover/High Availability Clusters

Hence, Clustering makes the database more available, and improves the performance of the database.

Therefore, improving the scalability of the database.