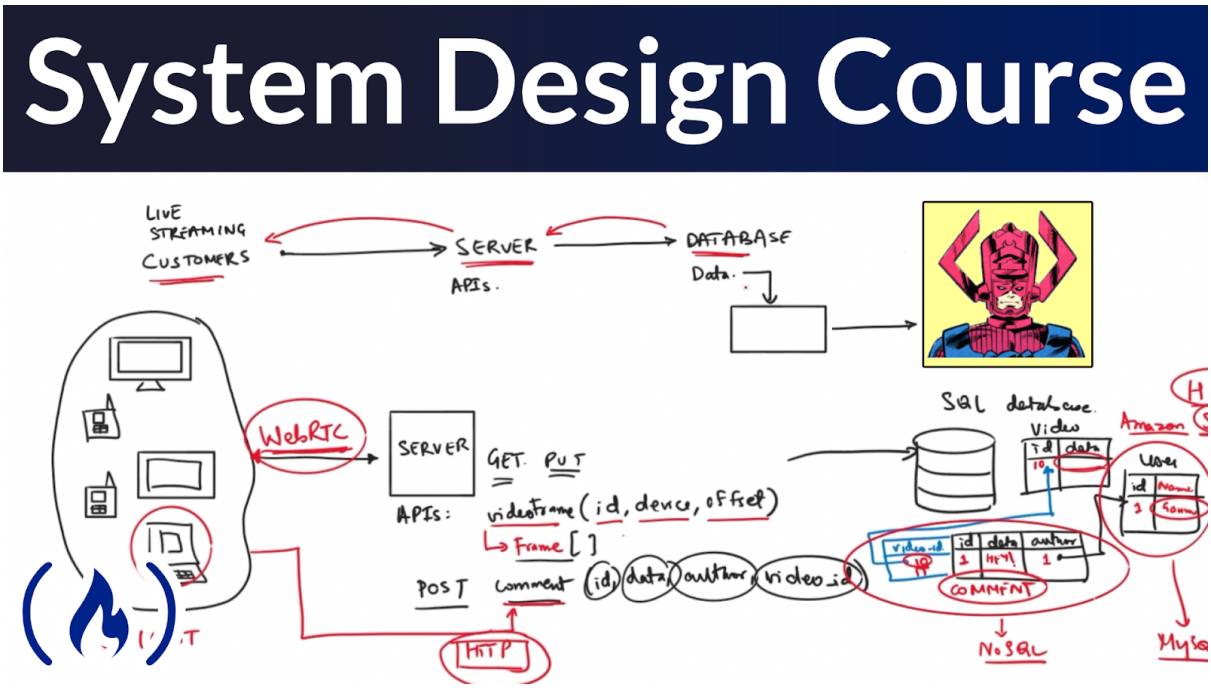
System Design

Define System Design ?

System design is the process of defining the elements of a system such as the architecture, modules and components of a system and the data that goes through that system

System design could be seen as the application of system theory to product development

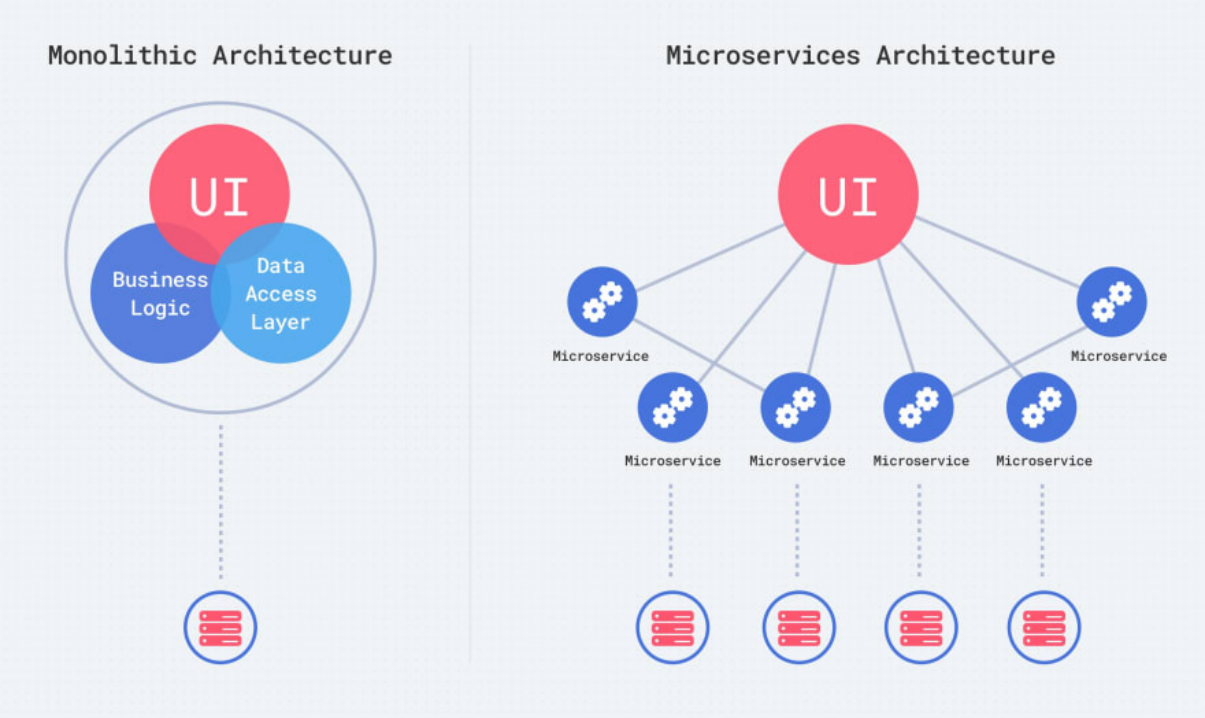


Why do we need system design ?

To build scalable applications

Difference between monolithic vs microservices ?

A monolithic application is built as a single unified unit while a microservices architecture is a collection of smaller, independently deployable services.



Define DNS ?

DNS => Domain Name System
IP => Internet Protocol

All computers on the Internet, from your smart phone or laptop to the servers that serve content for massive retail websites, find and communicate with one another by using numbers. These numbers are known as IP addresses. When you open a web browser and go to a website, you don't have to remember and enter a long number. Instead, you can enter a domain name like example.com and still end up in the right place.

A DNS service such as Amazon Route 53 is a globally distributed service that translates human readable names like www.example.com into the numeric IP addresses like 192.0.2.1 that computers use to connect to each other. The Internet's DNS system works much like a phone book by managing the mapping between names and numbers. DNS servers translate requests for names into IP addresses, controlling which server an end user will reach when they type a domain name into their web browser. These requests are called queries

Define availability ?

Availability is the amount of time that a system is able to respond as well as your system is up and running, that is the ratio of Uptime / (Uptime + Downtime).

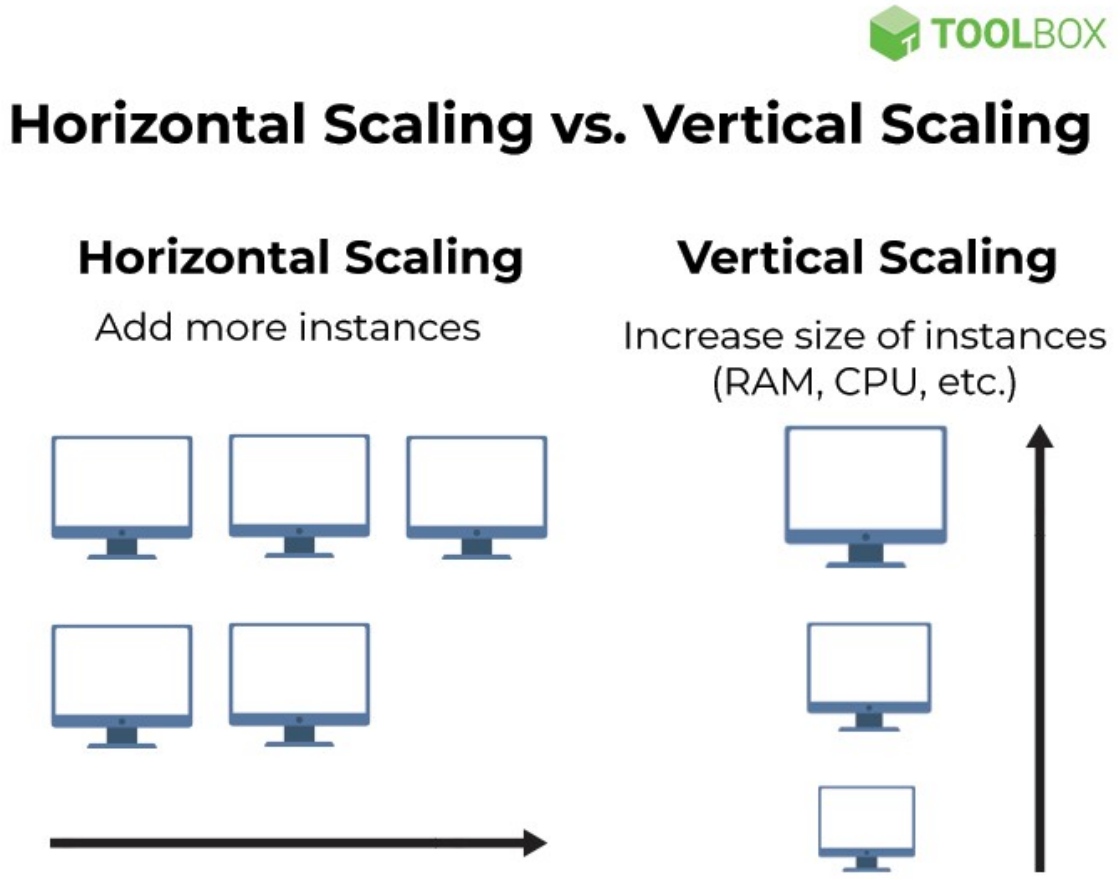
One way to look at it is how resistant a system is to failures

Define scalability ?

Scalability is the property of a system to handle a growing amount of work / traffic. A scalable system is one that can handle rapid changes to workloads and user demands, here workloads and user demands means traffic. Scalability is the measure of how well that system responds to changes by adding or removing resources to meet demands

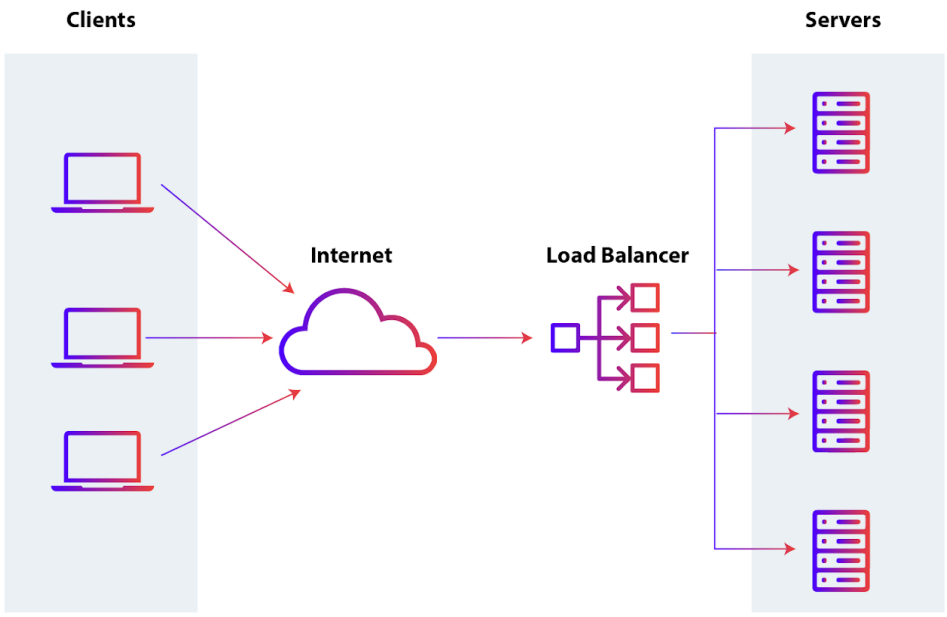
Horizontal scaling vs vertical scaling ?

The primary difference between horizontal scaling and vertical scaling is that horizontal scaling involves adding more machines or nodes to a system, while vertical scaling involves adding more power (CPU, RAM, storage, etc.) to an existing machine.



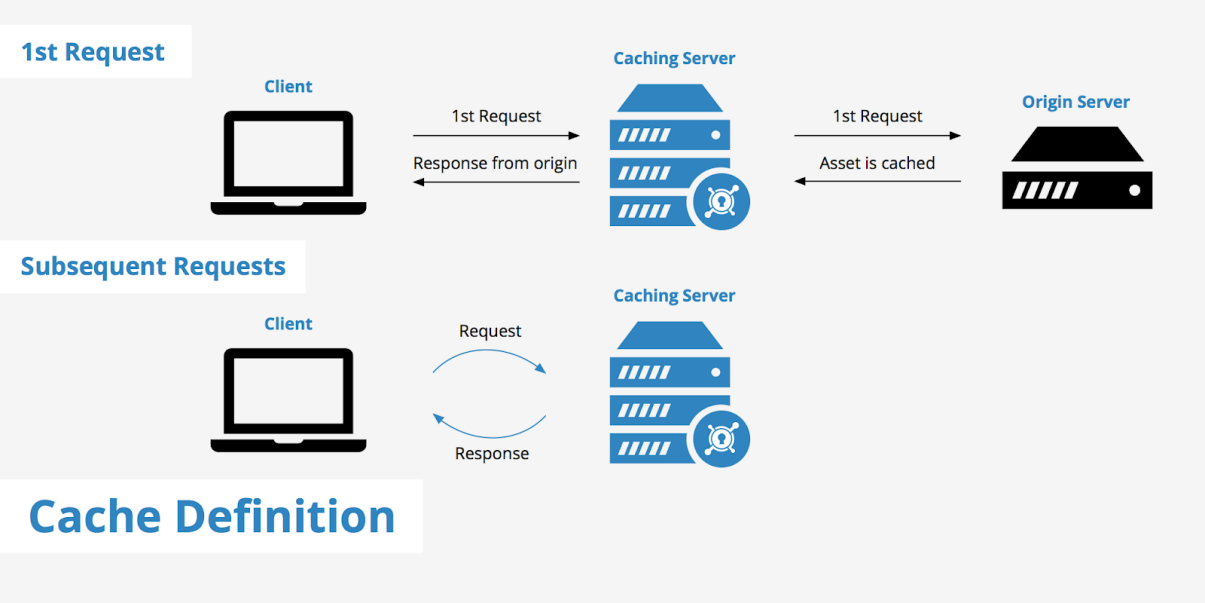
Define load balancing ?

In computing, load balancing is the process of distributing a set of tasks over a set of resources, with the aim of making their overall processing more efficient. Load balancing can optimise the response time and avoid unevenly overloading of some compute nodes while other compute nodes are left idle



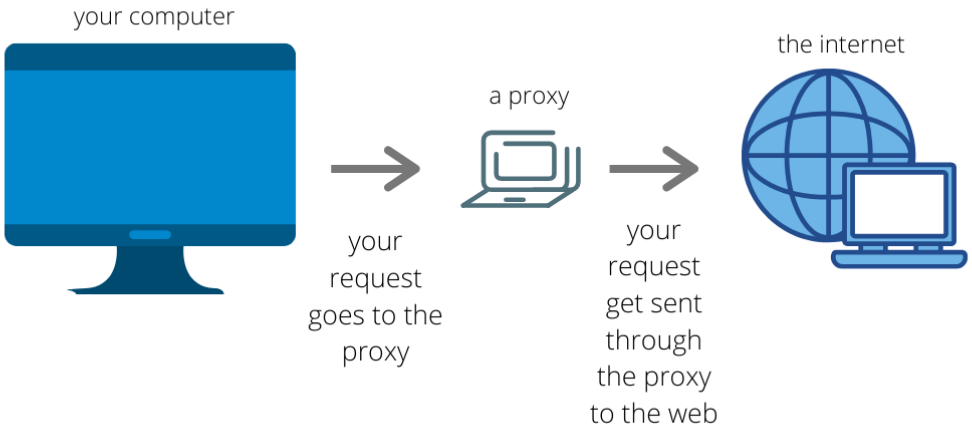
Define caching ?

In computing, a cache is a hardware or software component that stores data so that future requests for that data can be served faster; the data stored in a cache might be the result of an earlier computation or a copy of data stored elsewhere**.**



Define proxy ?

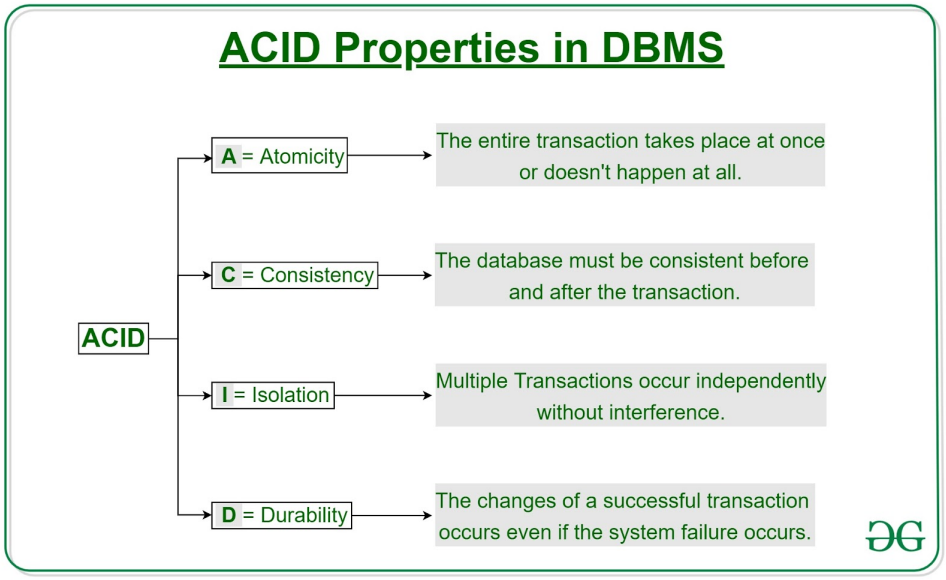
A proxy server is an intermediary piece of hardware/software sitting between the client and the backend server (acting as a middleman).It receives requests from clients and relays them to the origin servers. Typically, proxies are used to filter requests, log requests, or sometimes transform requests (by adding/removing headers, encrypting/decrypting, or compression.)



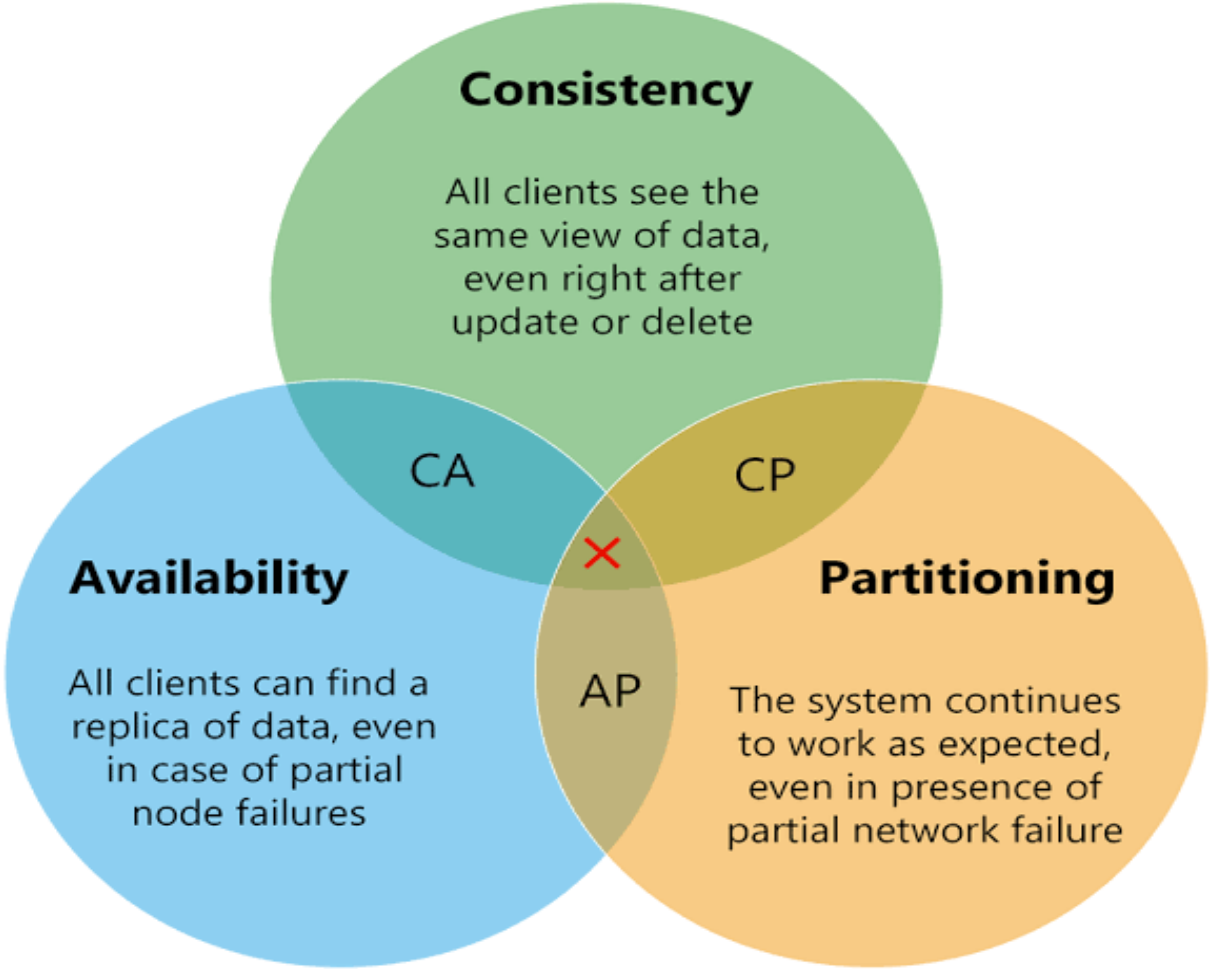
Mention five critical differences between SQL and NoSQL ?

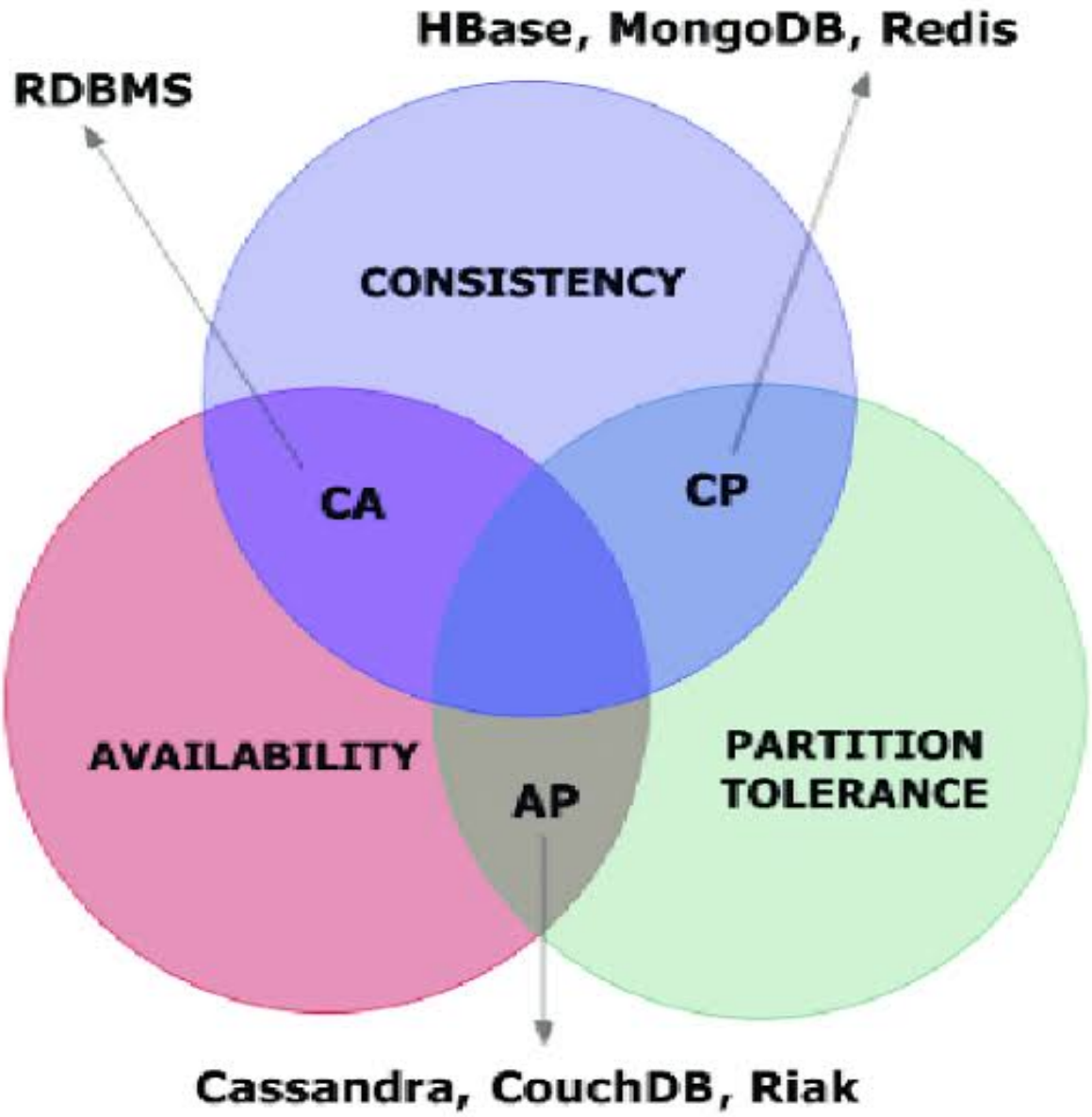
- SQL databases are relational, and NoSQL databases are non-relational.
- SQL databases use structured query language (SQL) and have a predefined schema for structured data. NoSQL databases have dynamic schemas for unstructured data.
- SQL databases are vertically scalable, while NoSQL databases are horizontally scalable.
- SQL databases are table-based, while NoSQL databases are document, key-value, graph, or wide-column stores.
- SQL databases are better for multi-row transactions, while NoSQL is better for unstructured data like documents or JSON.

ACID Properties


ACID Properties in DBMS
ACID
A = Atomicity
The entire transaction takes place at once or doesn't happen at all.
C = Consistency
The database must be consistent before and after the transaction.
I = Isolation
Multiple Transactions occur independently without interference.
D = Durability
The changes of a successful transaction occurs even if the system failure occurs.

CAP Theorem


Consistency
All clients see the same view of data, even right after update or delete
CA
CP
Availability
All clients can find a replica of data, even in case of partial node failures
AP
Partitioning
The system continues to work as expected, even in presence of partial network failure



Database Index

A database index is a data structure that improves the speed of data retrieval operations on a database table at the cost of additional writes and storage space to maintain the index data structure.

Index structure

Search key	Data Reference

Fig: Structure of Index

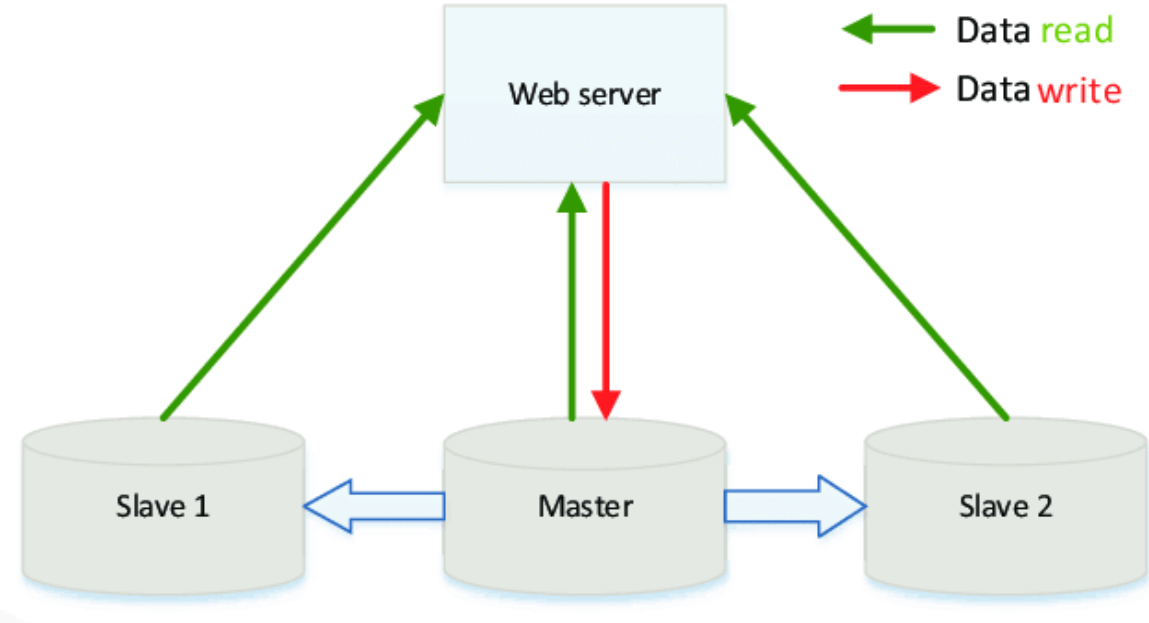
Indexes can be created using some database columns.

The first column of the database is the search key that contains a copy of the primary key or candidate key of the table. The values of the primary key are stored in sorted order so that the corresponding data can be accessed easily.

The second column of the database is the data reference. It contains a set of pointers holding the address of the disk block where the value of the particular key can be found

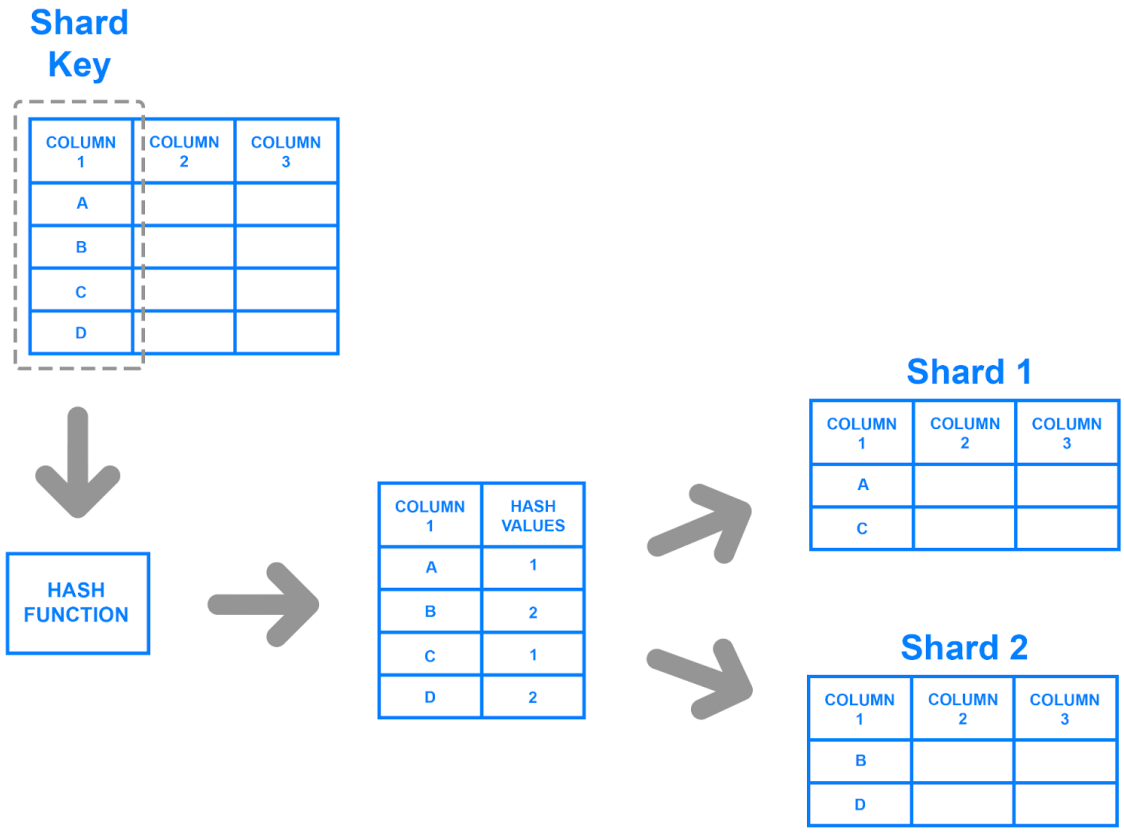
Database replication ?

Database replication refers to the process of copying data from a primary database to one or more replica databases in order to improve data accessibility, availability, system fault-tolerance and reliability across a network.



Database sharding ?

Database sharding is the process of storing a large database across multiple machines. A single machine, or database server, can store and process only a limited amount of data.



PRODUCT	PRICE
WIDGET	$118
GIZMO	$88
TRINKET	$37
THINGAMAJIG	$18
DOODAD	$60
TCHOTCHKE	$999

($0-$49.99)

PRODUCT	PRICE
TRINKET	$37
THINGAMAJIG	$18

($50-$99.99)

PRODUCT	PRICE
GIZMO	$88
DOODAD	$60

($100+)

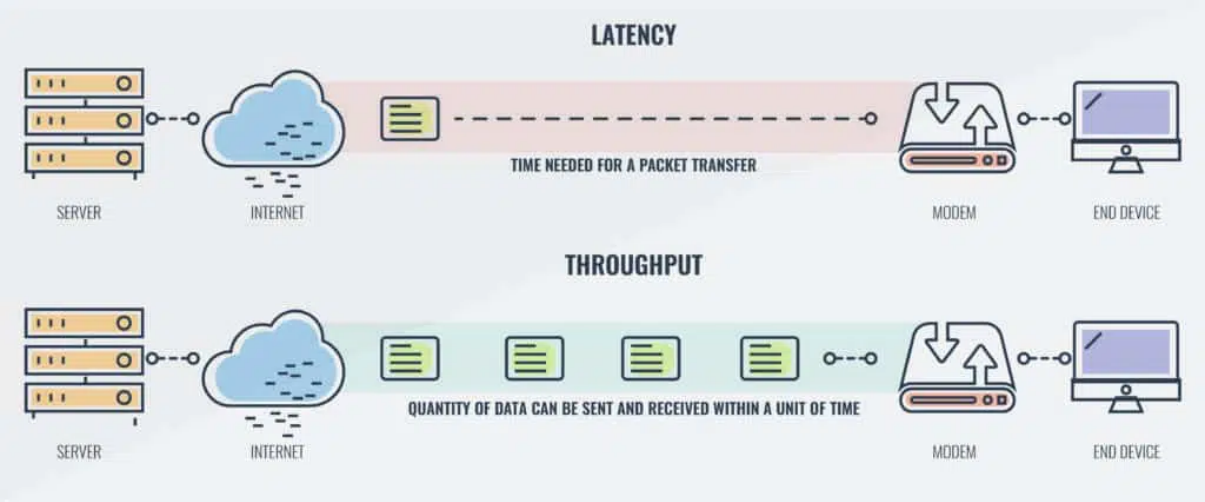
PRODUCT	PRICE
WIDGET	$118
TCHOTCHKE	$999

Latency vs Throughput ?

Latency is the time required to perform some action or to produce some result. Latency is measured in units of time -- hours, minutes, seconds, nanoseconds or clock periods. Throughput is the number of such actions executed or results produced per unit of time

Latency is the amount of time in milliseconds (ms) it takes a single message to be delivered.

Throughput is the amount of data that is successfully transmitted through a system in a certain amount of time, measured in bits per second (bps).

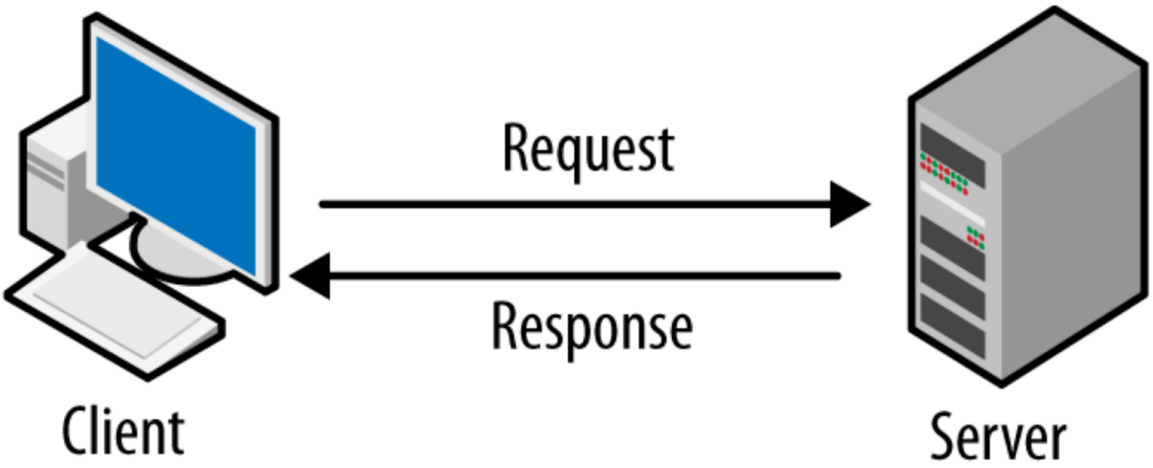


Define throttling ?

API throttling is the process of limiting the number of API requests a user can make in a certain period.

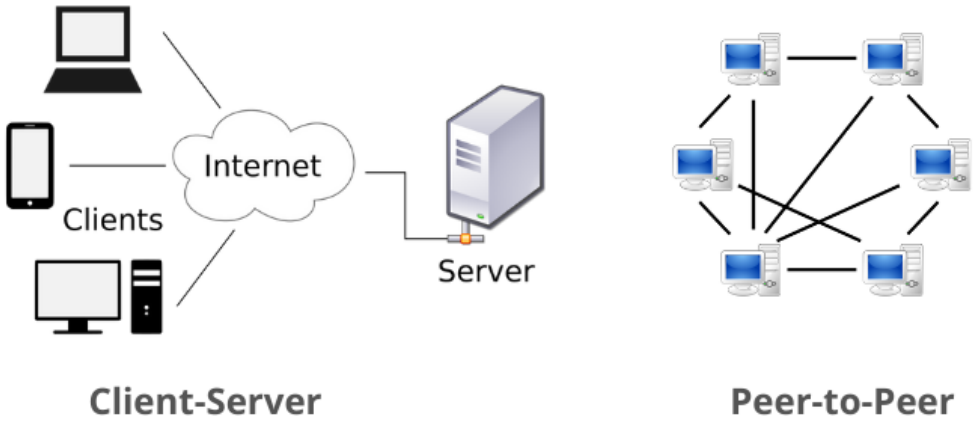
Define client server model ?

The client–server model is a distributed application structure that partitions tasks or workloads between the providers of a resource or service, called servers, and service requesters, called clients



Define peer-to-peer network ?

Peer-to-peer computing or networking is a distributed application architecture that partitions tasks or workloads between peers. Peers are equally privileged, equipotent participants in the network. They are said to form a peer-to-peer network of nodes

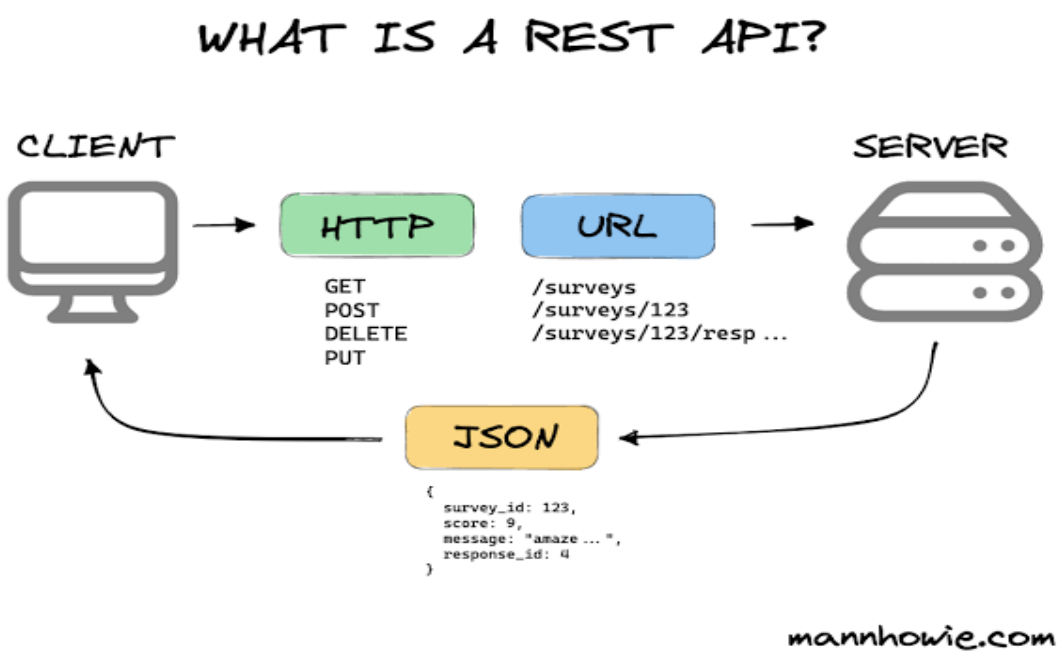


Define api ?

API stands for “application programming interface.” An API is essentially a set of rules that dictate how two machines/applications talk to each other. Some examples of API-based interactions include a cloud application communicating with a server, servers pinging each other, or applications interacting with an operating system.

Define REST api ?

RESTful API is an interface that two computer systems use to exchange information securely over the internet / HTTP protocol.



Method
Description

GET
Retrieve information about the REST API resource

POST
Create a REST API resource

PUT/PATCH
Update a REST API resource

DELETE
Delete a REST API resource or related component

Code
Status

Description

200
OK
The request was successfully completed.

201
Created

A new resource was successfully created.

400
Bad Request
The request was invalid.

401
Unauthorised
The request did not include an authentication token or the authentication token expired.

403
Forbidden
The client did not have permission to access the requested resource.

404
Not Found
The requested resource was not found.

405
Method Not Allowed
The HTTP method in the request was not supported by the resource. For example, the DELETE method cannot be used with the Agent API.

409
Conflict
The request could not be completed due to a conflict. For example, POST ContentStore Folder API cannot complete if the given file or folder name already exists in the parent location.

500
Internal Server Error
The request was not completed due to an internal error on the server side.

503
Service Unavailable
The server was unavailable

Define CRUD operations in Rest API's ?

CRUD is an acronym that comes from the world of computer programming and refers to the four functions that are considered necessary to implement a persistent storage application: create, read, update and delete.

Define CDN (Content Delivery Network) ?

A CDN is a network of servers that distributes content from an “origin” server throughout the world by caching content close to where each end user is accessing the internet via a web-enabled device. The content they request is first stored on the origin server and is then replicated and stored elsewhere as needed