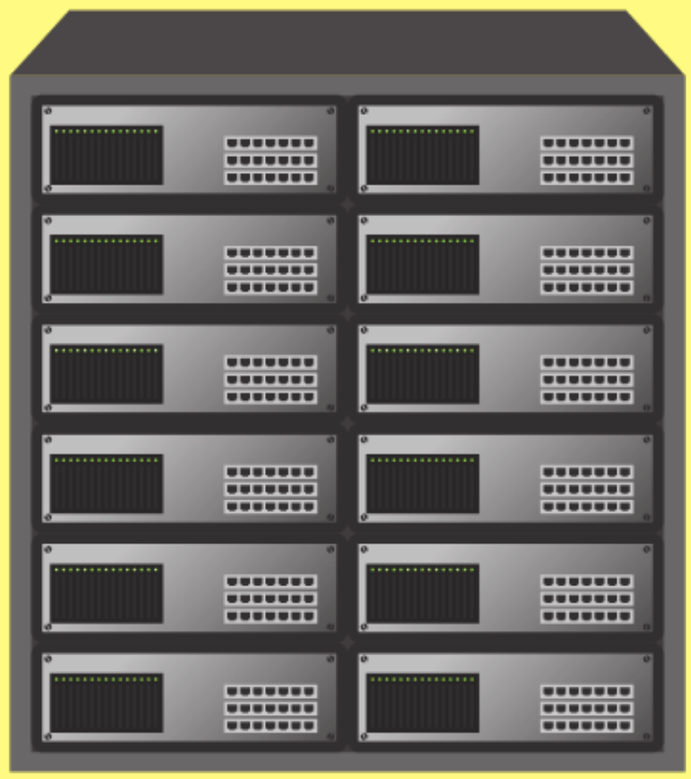
brijpandeyji

System Design Terminology

The Ultimate Guide to System Design Terms and Glossary



swipe >

brijpandeyji

Contents

swipe >

brijpandeyji

Contents

swipe >

brijpandeyji

Contents

- Consistent Hashing
- Count Min Sketch
- Data Warehousing
- Row Oriented Storage
- Column Oriented Storage
- Data Cube
- Denormalization
- Deserialization
- Disaster Recovery
- Distributed File Storage
- DNS (Domain Name System)
- Document Store
- Enterprise Service Bus (ESB)
- Event-Driven Architecture (EDA)
- Event Sourcing
- Federation

swipe >

brijpandeyji

Contents

- Geohashing
- GRPC
- GraphQL
- Graph Databases
- Indexes
- Implicit Indexes
- Composite Indexes
- Key-Value Store
- Caching
- Session Management
- High-Speed Data Storage
- Large Scale Systems
- Latency
- Linearizability
- Load Balancer
- Long Polling

swipe >

brijpandeyji

Contents

swipe >

brijpandeyji

Contents

- Vertical Sharding
- Performance
- Publish-Subscribe
- Quadtrees
- REST
- Read API
- Read Replicas
- RDBMS
- Relational Databases
- Reliability
- Remote Procedure Call (RPC)
- Replication
- Request Coalescing
- Response Time
- Reverse Proxy
- Rate Limiting

swipe >

brijpandeyji

Contents

- Scalability
- Serialization
- Server-Sent Events (SSE)
- Service Discovery
- SLA (Service Level Agreement)
- SLO (Service Level Objective)
- SLI (Service Level Indicator)
- Single Sign-On (SSO)
- SQL Tuning
- Indexing
- Query Optimization
- Partitioning
- Denormalization
- SSL (Secure Sockets Layer)
- Stream Processing
- TCP

swipe >

brijpandeyji

Contents

- Three Tier Caching
- Throughput
- TLS (Transport Layer Security)
- Transaction
- Atomicity
- Consistency
- Isolation
- Durability
- UDP
- Virtual Machines (VMs)
- Wide Column Store
- Web Sockets

END

swipe >

brijpandeyji

API Gateway

API Gateway refers to a server that acts as an entry point for client requests to access web services or APIs.

Asynchronism

Asynchronism refers to a programming paradigm where multiple tasks can be executed concurrently, without waiting for each other to complete.

Async API

Async APIs enable multiple clients to initiate and manage requests concurrently, improving system performance and scalability.

swipe >

brijpandeyji

Availability

Availability refers to the ability of a system to remain operational and accessible to users, even in the event of failures or outages.

Availability Patterns

Availability patterns are the design techniques and practices used to ensure high availability of a system. These patterns include replication, failover and availability in numbers.

Fail Over

Fail-over is an availability pattern that involves switching to a backup system when the primary system fails.

swipe >

brijpandeyji

Replication

Replication is an availability pattern that involves maintaining multiple copies of data in different locations.

Availability in Numbers

Availability in numbers is an availability pattern that involves using redundancy to ensure that there are multiple instances of a system available to handle requests.

Batch Processing

Batch processing is the execution of a series of jobs or tasks in a single batch or group, rather than individually or in real-time.

swipe >

brijpandeyji

Bloom Filter

Bloom filter is a probabilistic data structure used to test whether an element is a member of a set.

Cache

Cache is a temporary storage location that stores frequently accessed data to improve system performance.

Client Caching

Client caching is the process of storing frequently used data on the client-side.

swipe >

brijpandeyji

CDN Caching

CDN caching involves storing frequently accessed data on a Content Delivery Network.

Web Server Caching

Web server caching is the process of storing web pages on the server-side.

Database Caching

Database caching is the process of storing frequently used data on the database.

swipe >

brijpandeyji

Application Caching

Application caching is the process of storing frequently used data in the application.

Cache Aside

Cache-aside is a cache management technique that involves storing data in cache only when it´s needed.

Write Through

Write-through involves updating the cache and the original data source at the same time.

swipe >

Write Behind

Write-behind or write-back is a cache management technique involves updating the cache first and then updating the original data source later.

Refresh Ahead

Refresh-ahead involves updating the cache with anticipated data before it is requested.

Cache Stampede

Cache stampede is where multiple clients simultaneously request data that is not currently in the cache, resulting in a surge of requests to the original data source.

swipe >

CDN

CDN (Content Delivery Network) is a distributed network of servers that are geographically distributed to provide fast and reliable content delivery to users.

Push CDN

A CDN that pushes content to a network of servers distributed around the world before the content is requested, enabling faster delivery to users.

Pull CDN

A CDN that pulls content from the origin server when it is requested, and caches it at edge servers distributed around the world, enabling faster delivery to users.

swipe >

brijpandeyji

CQRS

CQRS (Command and Query Responsibility Segregation) is an architectural pattern that separates read and write operations into distinct models, each with its own data store.

Clocks

Clocks refer to a mechanism used to keep track of time and synchronize events across distributed systems.

Communication

Communication in system design refers to the exchange of information between different components of a distributed system.

swipe >

Consensus

Consensus is a mechanism used to reach agreement among distributed components on a single value or decision.

Consistency Patterns

Consistency patterns are techniques used to ensure that different components of a distributed system share a consistent view of data.

Weak Consistency

Data may be inconsistent for a short period of time in a weak consistency pattern.

swipe >

brijpandeyji

Eventual Consistency

Data will eventually become consistent, but there may be a delay.

Strong Consistency

Clocks refer to a mechanism used to keep track of time and synchronize events across distributed systems.

Consistent Hashing

Consistent hashing is a technique to partition and distribute data across multiple nodes in a way that minimizes the amount of data movement when nodes are added or removed from the system.

swipe >

brijpandeyji

Count Min Sketch

Count Min Sketch is a probabilistic data structure used for frequency counting and approximate querying.

Data Warehousing

Data warehousing is a technique used in system design to store and analyze large amounts of data from multiple sources in a centralized repository.

Row Oriented Storage

The database is partitioned horizontally and with this approach writes are performed easily as compared to reads.

swipe >

Column Oriented Storage

The database is partitioned vertically and with this approach reads are performed easily as compared to writes.

Data Cube

A data cube allows data to be modelled and viewed in multiple dimensions.

Denormalization

Denormalization is a technique used in database design to optimize query performance by adding redundant data to a database schema.

swipe >

Deserialization

Deserialization is the process of converting data in a serialized format back into its original form, such as converting binary data into an object.

Disaster Recovery

Disaster recovery is a process of preparing for and recovering from an unexpected event that causes a system outage or data loss.

Distributed File Storage

Distributed file storage is a system design approach that allows files to be stored across multiple nodes or servers in a distributed network.

swipe >

brijpandeyji

DNS (Domain Name System)

DNS (Domain Name System) is a hierarchical distributed system that translates human-readable domain names into IP addresses that machines can understand.

Document Store

A document store is a NoSQL database that stores and retrieves data in JSON, XML, or other document formats, providing flexible and schemaless data modeling.

Enterprise Service Bus (ESB)

Enterprise Service Bus (ESB) is a software architecture that provides a messaging infrastructure to facilitate communication between disparate applications, services, and systems within an organization.

swipe >

brijpandeyji

Event-Driven Architecture (EDA)

Event-Driven Architecture (EDA) is a software design pattern that emphasizes the use of events to trigger and communicate between different parts of a system.

Event Sourcing

Event Sourcing is a design pattern in which the state of a system is derived from a sequence of events.

Federation

Federation is a system design concept in which multiple autonomous systems are combined into a single cohesive system.

swipe >

brijpandeyji

Geohashing

Geohashing is a system design concept used for spatial indexing of geographic data.

gRPC

gRPC is a high-performance open-source Remote Procedure Call (RPC) framework that uses protocol buffers to enable communication between distributed systems.

GraphQL

GraphQL is a query language for APIs that allows clients to specify exactly what data they need and provides a predictable and efficient approach for fetching data.

swipe >

brijpandeyji

Graph Databases

Graph databases are NoSQL databases that store data in nodes and edges, enabling the representation and querying of complex, interconnected data.

Indexes

Indexes are data structures used in databases to speed up the retrieval of data by allowing for faster lookup of data based on specific fields.

Implicit Indexes

Implicit indexes are created by databases to internally to store, retrieve faster and efficiently.

swipe >

Composite Indexes

Composite indexes are created by using multiple columns to uniquely identify the data points.

Key-Value Store

A key-value store is a type of NoSQL database that stores data as key-value pairs, enabling high-speed data retrieval and storage.

Caching

Key-value stores are often used as an in-memory cache to store frequently accessed data, such as user sessions, to improve the performance of web applications.

swipe >

Session Management

Key-value stores are often used to store session data, such as user information and shopping cart contents, to keep track of the current state of a user's session.

High-speed Data Storage

Key-value stores are often used to store large amounts of data quickly and efficiently, such as in real-time analytics or distributed systems.

Large Scale Systems

Large scale systems refer to complex software systems that need to handle large amounts of data or traffic, and require scalable, fault-tolerant architecture.

swipe >

Latency

Latency is the time it takes for a request to be sent and a response to be received, and is an important factor in system performance and user experience.

Linearizability

Linearizability is a consistency model used in distributed systems to ensure that all nodes see the same order of events, even when requests are processed concurrently.

Load Balancer

A load balancer is a network device or software component that distributes traffic across multiple servers to improve system performance and prevent overloading of individual servers.

swipe >

Long Polling

Long polling is a technique used in web applications to allow for real-time updates without constantly refreshing the page, by holding open a connection until new data is available.

Maintainability

Maintainability is a software quality attribute that refers to how easy it is to modify and maintain a system over time.

Master-Master Replication

Master-master replication is a database replication technique where two or more nodes can act as both the master and the slave, allowing for data to be updated on any node and then propagated to all other nodes.

swipe >

Master-Slave Replication

Master-slave replication is a database replication technique where one node (the master) accepts updates and propagates them to one or more other nodes (the slaves).

Memory Cache

Memory cache is a type of cache that stores data in the memory of a system, enabling high-speed data retrieval.

Message Brokers

Message brokers are software components that enable communication between different parts of a system by routing messages between different nodes.

swipe >

brijpandeyji

Message Queues

Message queues are a type of message broker that allows for asynchronous communication between different parts of a system by storing messages until they can be processed.

Microservices

Microservices are a software architecture pattern where a system is broken down into small, independent services that can be developed and deployed separately.

mTLS

mTLS (Mutual Transport Layer Security) is a security protocol that uses certificates to establish a secure connection between two nodes, ensuring that both parties can verify each other´s identity.

swipe >

brijpandeyji

N-tier Architecture

N-tier architecture is a software architecture pattern where a system is divided into multiple layers, with each layer responsible for a specific set of functions.

Non-Relational Databases

NoSQL databases are a type of database that do not use a traditional relational data model, allowing for greater flexibility and scalability.

Object Store

Object store is a type of data storage system that stores data in a flat namespace, using unique identifiers for each object.

swipe >

brijpandeyji

Ordering

Ordering refers to the sequence in which events occur in a system, and is an important consideration in distributed systems to ensure consistency and correctness.

OAuth 2.0

An authorization framework that enables third-party applications to access user data on a resource server without the user's credentials.

OpenID Connect (OIDC)

An authentication protocol that enables single sign-on (SSO) between different systems using OAuth 2.0 framework.

swipe >

brijpandeyji

Partitioning/Sharding

A database scaling technique where data is distributed across multiple nodes to improve performance, availability, and scalability.

Horizontal Sharding

Horizontal sharding is a database partitioning technique where data is distributed across multiple nodes based on a shard key

Vertical Sharding

Vertical sharding involves splitting a table's columns into multiple physical tables.

swipe >

Performance

A measure of how well a system or component accomplishes its intended function, usually measured in terms of response time, throughput, or resource utilization.

Publish-Subscribe

A messaging pattern where a message producer sends messages to multiple consumers who are interested in receiving them, without requiring the producer to know the consumers' identities.

Quadtrees

A data structure that represents a two-dimensional space partitioned into smaller regions to efficiently perform spatial queries.

swipe >

brijpandeyji

REST

A architectural style for designing web services that uses HTTP methods to perform operations on resources.

Read API

An API designed for reading data from a system or database, usually optimized for high throughput and low latency.

Read Replicas

Duplicate copies of a database that are used to offload read traffic from the primary database, improving performance and scalability.

swipe >

RDBMS

A software system that manages relational databases, providing tools for data storage, retrieval, and modification.

Relational Databases

A type of database that stores data in tables with predefined relationships between them, usually managed by an RDBMS.

Reliability

The ability of a system or component to perform its intended function with a certain level of confidence and under specific conditions.

swipe >

brijpandeyji

Remote Procedure Call (RPC)

A protocol that enables a program to execute a procedure on a remote system over a network, as if it were local.

Replication

The process of copying data from one database to another for redundancy, availability, and scalability purposes.

Request Coalescing

A technique that groups multiple small requests into a single larger request to reduce overhead and improve performance.

swipe >

Response Time

The time it takes for a system or component to respond to a request or event.

Reverse Proxy

A server that acts as an intermediary between clients and a back-end server, usually used for load balancing, security, and caching purposes.

Rate Limiting

A technique used to limit the amount of traffic or requests that a system or API can handle over a given period of time, to prevent overload and maintain performance.

swipe >

Scalability

The ability of a system or component to handle increasing amounts of work or traffic by adding resources or nodes to the system.

Serialization

The process of converting an object into a format that can be transmitted over a network or stored in a file.

Server-Sent Events (SSE)

A protocol for real-time, bi-directional communication between a server and clients over HTTP, used for streaming data or notifications.

swipe >

Service Discovery

The process of automatically discovering and registering services in a distributed system, usually using a centralized registry or a peer-to-peer protocol.

SLA (Service Level Agreement)

A contractual agreement that specifies the expected level of service and performance that a provider will deliver to a client.

SLO (Service Level Objective)

A measurable goal or target for a service's performance or availability, usually based on an SLA or business requirements.

swipe >

brijpandeyji

SLI (Service Level Indicator)

A measurable metric used to track and monitor a service's performance or availability, usually based on an SLO or business requirements.

Single Sign-On (SSO)

A process that enables users to authenticate once and access multiple applications or systems without requiring additional authentication steps.

SQL Tuning

The process of optimizing SQL queries to improve their performance and efficiency.

swipe >

brijpandeyji

Indexing

Creating indexes on columns that are frequently used in WHERE clauses and JOINs can significantly improve query performance.

Query Optimization

The database optimizer analyzes the query and selects the most efficient execution plan.

Partitioning

Partitioning large tables into smaller, more manageable chunks can improve query performance, especially for large data sets.

swipe >

Denormalization

Denormalization is the process of purposely adding redundant data to one or more tables to improve query performance.

SSL (Secure Sockets Layer)

A security protocol that encrypts data transmitted over the internet to ensure secure communication.

Stream Processing

A method of processing continuous streams of data in real-time to derive insights and take action.

swipe >

brijpandeyji

TCP (Transmission Control Protocol)

A reliable, connection-oriented protocol used for transmitting data over the internet.

Three Tier Caching

A caching strategy that uses three levels of cache: client, application server, and database server, to improve application performance.

Throughput

The amount of data or transactions that a system can process in a given amount of time.

swipe >

brijpandeyji

TLS (Transport Layer Security)

A security protocol that provides encryption and authentication of data transmitted over the internet.

Transaction

A logical unit of work that comprises one or more database operations that must be performed as a single, indivisible unit.

Atomicity

It means either all the operations of a transaction are properly reflected in the database or none of them.

swipe >

Consistency

It means the execution of a transaction should be isolated so that data consistency be maintained.

Isolation

In situations where multiple transactions are executing, each transaction should be unaware of the other executing transaction and should be isolated.

Durability

It means after the transaction is complete, the changes that are made to the database should persists even in case of system failures.

swipe >

brijpandeyji

UDP (User Datagram Protocol)

A lightweight, connectionless protocol used for transmitting data over the internet.

Virtual Machines (VMs)

A software emulation of a physical computer that allows multiple operating systems to run on a single physical machine.

Wide Column Store

A type of NoSQL database that uses a column-family data model to store data in a distributed and scalable manner.

swipe >

brijpandeyji

Web Sockets

A protocol that provides bi-directional, full-duplex communication channels over a single, long-lived connection between a client and server.

swipe >

brijpandeyji

For More Interesting Content

Brij Kishore Pandey

Follow Me On LinkedIn

https://www.linkedin.com/in/brijpandeyji/