OS Assignment

Chetan Patel
21119029
CPatel

Q.1 Types of distributed system storages?

Ans. ① distributed storage systems enable data to be stored across multiple nodes/servers, improving availability, fault tolerance and scalability.

② Unlike traditional centralised storage distributed storage systems ensure that data remains accessible even if some nodes fail.

Types

① Distributed file system :- Store data in multiple blocks across nodes allowing for parallel data access and fault tolerance.

② Object storage :- Organises data as objects within buckets making it ideal for unstructured data and storage.

③ NOSQL database :- Uses a distributed architecture that supports high-throughput of read and write operation.

Q2 Replication and Recovery

21/10/2024
Model

Ans. (1) Replication is the process of creating copies of data across multiple nodes in distributed system, enhancing fault tolerance.

(2) By replicating DS can enhance continuity even if individual nodes fail.

(3) 2 - strab

a) Synchronous replication :- require all copies to be updated immediately, ensuring data consistency but higher latency.

b) Asynchronous replication : allow change to be gradual providing lower latency but higher inconsistency on data.

(4) Recovery mechanism

(1) They complement replication by helping systems restore data and resume operation after failure

(2) Techniques allow a new node to take over a leader if they fail.

(3) Data rebalancing helps distribute data evenly across node after recovery.

Q.1 Commit protocols

2111902B
Model

Ans.. (1) Two phase commit:-

It is a distributed algorithm designed to ensure that all node involved in a transaction agree to 'commit' or 'abort'. In the first phase the coordinator sends a 'prepare' message to all nodes. If all agree we go 2nd phase.

(2) Three phase commit:-

It is built over 2PC to reduce the risk of blocking by adding an intermediate 'prepare to commit' stage. After this nodes move to prepare state if all vote 'yes'. This extra steps prevent blocking.

Q.2 Real world applications

Ans. Distributed systems have become essential across many industries to provide high availability, fault tolerance.

(1) In the finance sector they provide high frequency trading platforms

(2) In healthcare, distributed system allows seamless sharing and storage of electronic health record across hospital records

(3) The ecommerce industry uses distributed system to manage inventory, handle transaction

Q.5 Scalability and reliability

21/10/2024

Ans. (1) DS are fundamental to Scalel cloud computing enabling platforms like AWS, google cloud and MS Azure to deliver scalable app.

(2) In a distributed cloud env, data is replicated across multiple centers and regions ensuring redundancy and high availability

(3) This redundancy allow cloud services to scale dynamically

(4) Additionally, DS support moving data closer to user, reducing latency.

Q.6 Fault tolerance and load balancing.

(1) DS attain fault tolerance and load balancing through redundancy, replication and automated recovery mech.

(2) Fault tolerance ensures by duplicating data across nodes if one fail another replies

(3) they also employ monitoring tool to detect real time failure, automatically rerouting request and initiating recovery phase

Q.7 Decentralised architecture

Ans. ① decentralised architecture is distributed system have enabled technologies like Blockchain that empower crypto-currency.

② Here there is no need of an central omnipotent authority.

③ This prevents tampering and improves integrity.

④ Decentralised architecture ensure immutability, autonomy and transparency, empowering user to have control over their data and transaction

Q.8 High performance computing

Ans. ① DS play an important role in high performance computing, by adding parallel processing of large datasets.

② Framework like Hadoop and Apache are specially designed for distributed data processing

③ Hadoops mapReduce programming model allows tasks to be split into smaller sub-tasks

④ Apache spark enhances performance by processing data on memory rather than disk

Q.9 Distributed File system

2/11/2024
Page

Ans. (1) Distributed file systems are storage architecture designed to store files across multiple nodes in a network.

(2) In a DFS, files are divided into blocks and replicated across nodes to ensure redundancy.

(3) DFS support parallel data access, enabling faster read and write speed.

(4) DFS also incorporate data balancing mechanisms to evenly distribute files across nodes.

Q.10 HDFS and Hadoop MapReduce

Ans. (1) HDFS and Hadoop MapReduce together provide a powerful framework for processing and storing vast amounts of data in distributed environ-

(2) HDFS allow storage of data across multiple node, splitting file into block and replicating them across different nodes.

(3) This allows Hadoop to process petabyte scale data effectively making it apt. for web indexing, log analysis and engine for recommendation