ONE PAGE SUMMARY

IT & Software › NoSQL › Cassandra

Introduction to Cassandra

Cassandra: The Scalable NoSQL Database for Big Data Applications

DEFINITION

> Cassandra is a highly scalable NoSQL database designed for handling large amounts of data across many servers. It offers high availability and no single point of failure.

KEY CONCEPTS

Distributed Architecture

Cassandra uses a distributed architecture, meaning data is spread across multiple nodes for reliability and speed.

Data Model

Cassandra employs a wide-column store data model, allowing for flexible schema design and efficient storage.

Replication

Data in Cassandra is replicated across nodes to ensure durability and availability, configurable based on needs.

CQL

Cassandra Query Language (CQL) is used to interact with the database, similar to SQL but tailored for Cassandra's architecture.

Tunable Consistency

Cassandra allows developers to choose the level of consistency for read and write operations, balancing speed and accuracy.

EXAMPLES

→ Storing user profiles in a social media application.

→ Logging sensor data from IoT devices.

→ Managing product catalogs in e-commerce platforms.

MEMORY TIPS

★ Remember 'Cassandra' as 'Cassandra the Queen of Scalability'.

★ Think of 'CQL' as 'Cassandra's Query Language', like SQL but for Cassandra.

★ Visualize data spread like a 'Cassandra Cake' with layers representing nodes.

COMMON MISTAKES

✗ Confusing Cassandra with traditional relational databases.

✗ Neglecting to configure replication settings properly.

✗ Underestimating the importance of data modeling.

QUICK RECAP

Cassandra is a NoSQL database known for its scalability and high availability. It uses a distributed architecture and allows flexible data modeling with CQL. Remember its key features like replication and tunable consistency for effective data management.