Database Management System – 50 (Characteristics of Redis, MongoDB, Cassandra and ArangoDB)

Ajay James
Asst. Prof in CSE
Government Engineering College Thrissur

Outline

- Redis
- MongoDB
- Cassandra
- ArangoDB

Redis

- **RE**mote **DI**ctionary **S**erver
- Developed in 2009 (Open-source (BSD licensed))
- Key-value store
- In-memory data structure store
- Allows the user to store vast amounts of data without the limits of a relational database
- Supports various data structures such as strings, hashes, sets, lists, sorted sets, bitmaps, hyperloglogs and geospatial indexes
- Redis can be compiled and used on Linux, OSX, OpenBSD, NetBSD, FreeBSD

Features of Redis

- **Speed**
  - Redis loads the whole dataset in memory
  - Supports Pipelining of commands
  - Getting and setting multiple values in a single command is possible to speed up communication with the client libraries
- **Persistence**
  - While all the data lives in memory, changes are asynchronously saved on disk using flexible policies based on elapsed time and/or number of updates since last save

Features of Redis

- **Data Structures**
  - Supports data structures such as strings, hashes, sets, lists, sorted sets with range queries, bitmaps, hyperloglogs and geospatial indexes with radius queries
- **Atomic Operations**
  - Are atomic
  - So setting or increasing a key, adding and removing elements from a set, increasing a counter will all be accomplished safely.

Features of Redis

- **Supported Languages**
  - ActionScript, C, C++, C#, Clojure, Common Lisp, D, Dart, Erlang, Go, Haskell, Haxe, Io, Java, JavaScript (Node.js), Julia, Lua, Objective-C, Perl, PHP, Pure Data, Python, R, Racket, Ruby, Rust, Scala, Smalltalk and Tcl.
- **Master/Slave Replication**
  - Supports a very simple and fast Master/Slave replication

Features of Redis

- **Sharding**
  - Distributing the dataset across multiple Redis instances is easy in Redis
- **Portable**
  - Written in ANSI C and works in most POSIX systems like Linux, BSD, Mac OS X, Solaris

MongoDB

- Scalable and flexible NoSQL document database platform
- MongoDB is well known for its horizontal scaling and load balancing capabilities

Features of MongoDB

1. Ad-hoc queries for optimized, real-time analytics
   - Ad hoc query is a short-lived command whose value depends on a variable
   - Each time an ad hoc query is executed, the result may be different, depending on the variables in question
   - MongoDB has ad-hoc query support that allows developers to update ad-hoc queries in real time

Features of MongoDB

2. Indexing appropriately for better query executions
   - MongoDB offers a broad range of indices and features with language-specific sort orders that support complex access patterns to datasets
   - MongoDB indices can be created on demand to accommodate real-time, ever-changing query patterns and application requirements.

Features of MongoDB

3. Replication for better data availability and stability
   - Replica sets are employed in MongoDB
   - A primary server or node accepts all write operations and applies those same operations across secondary servers, replicating the data
   - If the primary server should ever experience a critical failure, any one of the secondary servers can be elected to become the new primary node
   - If the former primary node comes back online, it does so as a secondary server for the new primary node.

Features of MongoDB

4. Sharding
   - Sharding—the process of splitting larger datasets across multiple distributed collections, or "shards"
   - Helps the database distribute and better execute what might otherwise be problematic and cumbersome queries
   - Horizontal scaling means that each shard in every cluster houses a portion of the dataset in question, essentially functioning as a separate database

Features of MongoDB

5. Load balancing
   – horizontal scaling features like replication and sharding, MongoDB supports large-scale load balancing
   – can handle multiple concurrent read and write requests for the same data
   – Nest-in-class concurrency control and locking protocols that ensure data consistency

Features of MongoDB

6. Uses JavaScript instead of Procedures.
7. It is a schema-less database written in C++.
8. Provides high performance.
9. Stores files of any size easily without complicating your stack
10. Easy to administer in the case of failures

Apache Cassandra

- Open source distributed
- Decentralized/distributed storage system (database)
- Manage very large amounts of structured data spread out across the world
- Provides highly available service with no single point of failure

Apache Cassandra

- It is scalable, fault-tolerant, and consistent
- It is a column-oriented database
- Its data model on Google's Bigtable
- Created at Facebook
- Cassandra implements a replication model with no single point of failure
- Cassandra is being used by some of the biggest companies such as Facebook, Twitter, Cisco, Rackspace, ebay, Twitter, Netflix, and more

Features of Cassandra

- **Elastic scalability**
  - Highly scalable
  - Allows to add more hardware to accommodate more customers and more data as per requirement
- **Always on architecture**
  - Has no single point of failure and it is continuously available for business-critical applications that cannot afford a failure
- **Fast linear-scale performance**
  - It is linearly scalable, i.e., it increases your throughput as you increase the number of nodes in the cluster

Features of Cassandra

- **Flexible data storage**
  - Accommodates all possible data formats including: structured, semi-structured, and unstructured
- **Easy data distribution**
  - Flexibility to distribute data where you need by replicating data across multiple data centers
- **Transaction support**
  - Supports ACID properties
- **Fast writes**
  - Performs fast writes and can store hundreds of terabytes of data, without sacrificing the read efficiency

ArangoDB

- Called a Multi-model database
- Data can be stored as documents, key/value pairs or graphs

Features of ArangoDB

- Multi-model Paradigm
  - Document model
  - Key/Value model
  - Graph model
- ACID Properties
- HTTP API
  - Allows clients, such as browsers, to interact with the database with HTTP API, the API being resource-oriented and extendable with JavaScript.

Advantages of using ArangoDB

- Consolidation
  - Eliminates the need to deploy multiple databases
  - Decreases the number of components and their maintenance
  - Leads to lower total cost of ownership and increasing flexibility
- Simplified Performance Scaling
  - Independently scaling with different data models
  - Can scale both vertically and horizontally

Advantages of using ArangoDB

- Reduced Operational Complexity
- Strong Data Consistency
- Fault Tolerance
- Lower Total Cost of Ownership
- Transactions

References

- https://www.w3resource.com/redis/
- https://www.mongodb.com/what-is-mongodb/features
- https://www.javatpoint.com/mongodb-features
- https://www.tutorialspoint.com/cassandra/cassandra_introduction.htm
- https://www.tutorialspoint.com/arangodb/arangodb_advantages.htm

Thank you