NoSQL

NoSQL Database Types

Different Architectural Patterns in NoSQL

– Key-Value databases examples : Riak,Redis, Memcached, BerkeleyDB, upscaledb, Amazon DynamoDB.

– Document databases examples : MongoDB, CouchDB, Terrastore, OrientDB , RavenDB

– Column family stores examples : Cassendra, HBase, HyperTable.

– Graph Databases examples : Neo4j, InfiniteGraph, FlockDB.

1. Key-value store databases

– This is very simple NoSQL database.

– It is specially designed for storing data as a schema free data.

– Such data is stored in a form of data along with indexed key.

Examples

– Cassandra

– Azure Table Storage (ATS)

– DyanmoDB

Use Cases

This type is generally used when you need quick performance for basic Create-Read-Update-Delete operations and

data is not connected.

Example

– Storing and retrieving session information for a Web pages.

– Storing user profiles and preferences

– Storing shopping cart data for ecommerce

Limitations

– It may not work well for complex queries attempting to connect multiple relations of data.

– If data contains lot of many-to-many relationships, a Key-Value store is likely to show poor performance.

Column store database

– Instead of storing data in relational tuples (table rows), it is stored in cells grouped in columns.

– It offers very high performance and a highly scalable architecture.

Examples

(i) HBase (ii) Big Table (iii) Hyper Table

Use Cases

− Some common examples of Column-Family database include event logging and blogs like document databases,

but the data would be stored in a different fashion.

− In logging, every application can write its own set of columns and have each row key formatted in such a way to

promote easy lookup based on application and timestamp.

− Counters can be a unique use case. It is possible to design application that needs an easy way to count or

increment as events occurs.

Document database

– Document databases works on concept of key-value stores where "documents" contains a lot of complex data.

– Every document contains a unique key, used to retrieve the document.

– Key is used for storing, retrieving and managing document-oriented information also known as semi-structured

data.

Examples

(i) MongoDB (ii) Couch DB

Use Cases

The example of such system would be event logging system for an application or online blogging.

In online blogging user acts like a document; each post a document; and each comment, like, or action would

be a document.

All documents would contain information about the type of data, username, post content, or timestamp of document creation.

Limitations

– It's challenging for document store to handle a transaction that on multiple documents.

– Document databases may not be good if data is required in aggregation.

4. Graph database

Data is stored as a graph and their relationships are stored as a link between them whereas entity acts like a node.

Examples

(i) Neo4j (ii) Polyglot

Use Cases

– The very important and popular application would be social networking sites can benefit by quickly locating friends, friends of friends, likes, and so on.

– The Google Maps can help you to use graphs to easily model their data for finding close locations or building shortest routes for directions.

– Many recommendation systems makes effective use of this model.

Limitations

– Graph Databases may not be offering better choice over other NoSQL variations.

– If application needs to scale horizontally this may introduces poor performance.

– Not very efficient when it needs to update all nodes with a given parameter.

Comparison of NoSQL variations

Database model	Performance	Scalability	Flexibility
Key value store database	High	High	High
Column store database	High	High	Moderate
Document store database	High	Variable (High)	High
Graph database	Variable	Variable	High

Benefits of NoSQL

1. Big data analytics

− Big data is one of main feature promotes growth and popularity of NoSQL.

− NoSQL has good provision to handle such big data.

2. Better data availability

− NoSQL database works with distributed environments.

− NoSQL database environments should provide good availability across multiple data servers.

− NoSQL databases supply high performance.

3. Location independence

NoSQL data base can read and write database regardless of location of database operation.

Tools Used for Big Data

1. Map Reduce

Hadoop, Hive, Pig, Cascading, Cascalog, mrjob, Caffeine, S4, MapR, Acunu, Flume, Kafka, Azkaban, Oozie, Greenplum

2. Storage

S3, Hadoop Distributed File System

3. Servers

EC2, Google App Engine, Elastic, Beanstalk, Heroku

4. NoSQL

Zookeeper, MongoDB, Cassandra, Redis, Big Table, Hbase, Hyper table, Voldemort, Riak, Couch DB

5. Processing

R, Yahoo! Pipes, Mechanical Turk, Solr/Lucene, ElasticSearch, Datameer, BigSheets, Tinkerpop