CLASS PRESENTATION
2023-2024

DOCUMENT DATABASE

JEDDAH UNIVERCITY

NAMES OF TEAM GROUP:

Salma mohammed	2110956
Dania Abdelmonem	2112241
Somaya Almalki	2111499
Ramah Hassan Alharbi	2115249

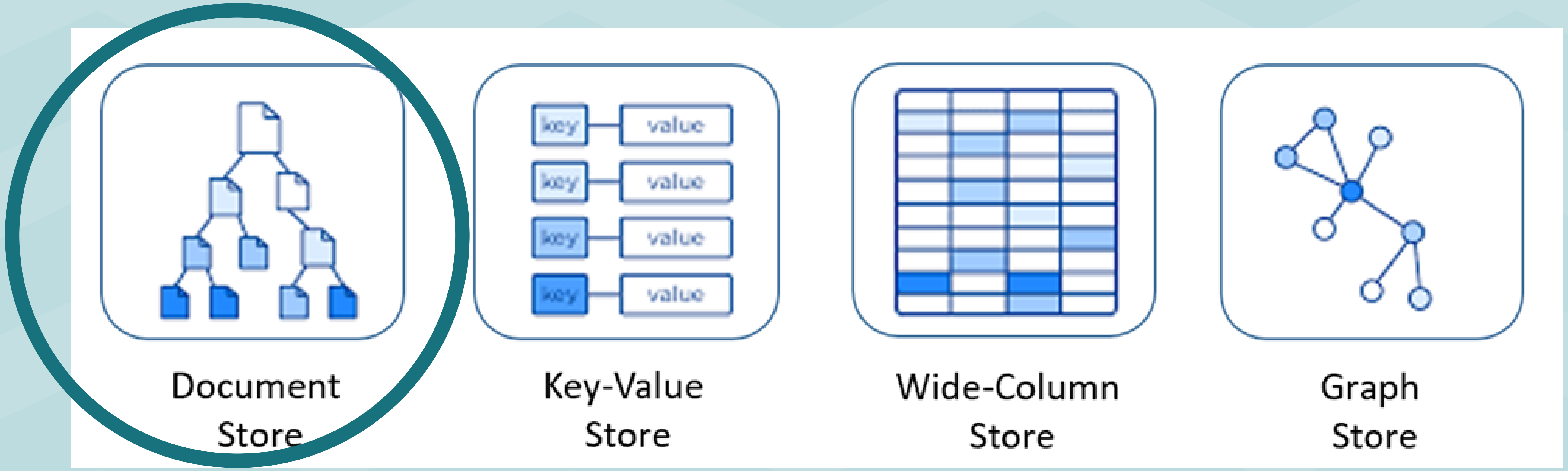
OUTLINE

- Introduction
- CouchBD
- MongoDB
- CAP theorem
- Comparison of CouchDB with RDB
- Comparison of MongoDB with RDB
- CRUD Operations
- Queries

INTRODUCTION TO NO-SQL

NoSQL is a non-relational database that is used to store the data in the nontabular form. NoSQL stands for Not only SQL. the main types are documents, key-value, wide-column, and graphs.

TODAY OUR TOPIC IS ABOUT?



WHAT IS A DOCUMENTS DATABASE?

The document-based database is a non-relational database, Instead of storing the data in rows and columns (tables), it uses the documents to store the data in the database. A document database stores data in JSON, BSON, or XML documents.

Also Documents can be stored and retrieved in a form that is much closer to the data objects used in applications which means less translation is required to use these data in the applications.
the particular elements can be accessed by using the index value that is assigned for faster querying.

OTHER DEATILS ABOUT DOCUMENTS DB:

Key features of documents database:

-Flexible schema

-Faster creation and maintenance

-No foreign keys

-Open formats

Strengths:

-Powerful data model – able to express nested structure

-Good scaling

-No database maintenance required to add or remove 'columns'.

weakness:

-Unsuited for interconnected data

-Query model limited to keys.

MONGODB

DEFINITION

*What is MongoDB?
Scalable High-Performance Open-source, Document-orientated database, Built for Speed.

FEATURES

1. Rich Document based queries for Easy readability.
2. Full Index Support for High Performance.
3. Replication and Failover for High Availability.
4. Auto Shading for Easy Scalability.
5. Map / Reduce for Aggregation

FUNCTIONALITY

- Dynamic schema (No DDL)
- Document-based database
- Secondary indexes.
- Query language via an API
- Atomic writes and fully-consistent reads(if system configured that way)
- Master-slave replication with automated failover (replica sets)
- Built-in horizontal scaling via automated range-based
- partitioning of data (sharding)
- No joins nor transactions

WHY USE MONGODB?

Simple queries
Functionality provided applicable to most web applications
Easy and fast integration of data

- No ERD diagram

WHEN USE MONGODB?

1. You're using cloud computing. MongoDB is ideal for cloud computing
2. You need your data fast and easily accessible.
3. You don't have a database administrator.
4. You have lots of unstructured data.
5. You're using Agile methodologies for development.
6. You have schema issues.

EXAMPLE:

EBAY

*is a multinational company that provides a platform for the customer to customer sales.

METLIFE

is a leading company in employee benefit programs, annuities, and insuranceng for Easy Scalability.

SHUTTERFLY

(PROS&CONS)

PROS:

RDBMS replacement for Web Applications. Semi-structured Content Management.

Real-time Analytics & High-Speed Logging. Caching and High Scalability

CONS:

Highly Transactional Applications. Problems requiring SQL.

Not well suited for heavy and complex transactions systems.

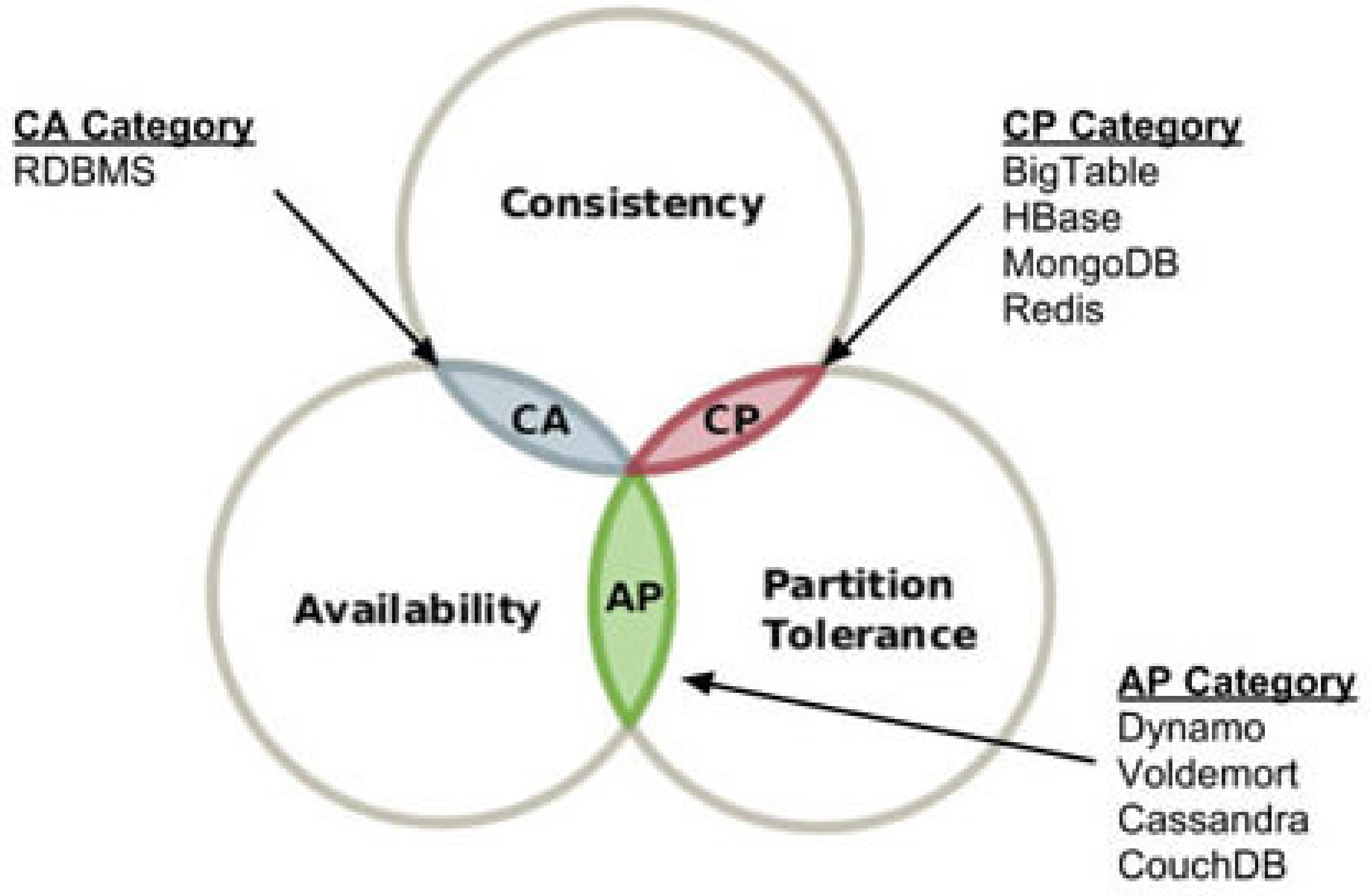
CAP THEOREM

CAP THEOREM IS ALSO CALLED BREWER'S THEOREM. IT STATES THAT IS IMPOSSIBLE FOR A DISTRIBUTED DATA STORE TO OFFER MORE THAN TWO OUT OF THREE GUARANTEES :

1) CONSISTENCY
2) AVAILABILITY
3) PARTITION TOLERANCE

EXPLANATION:

CAP Theorem



CAP THEOREM

NOSQL DATABASES ARE CLASSIFIED BASED ON THE TWO CAP CHARACTERISTICS THEY SUPPORT:

CP DATABASE: A CP DATABASE DELIVERS CONSISTENCY AND PARTITION TOLERANCE AT THE EXPENSE OF AVAILABILITY.

AP DATABASE: AN AP DATABASE DELIVERS AVAILABILITY AND PARTITION TOLERANCE AT THE EXPENSE OF CONSISTENCY.

CA DATABASE: A CA DATABASE DELIVERS CONSISTENCY AND AVAILABILITY ACROSS ALL NODES.

MONGODB AND CAP THEOREM:

ACCORDING TO THE CAP THEOREM, MONGODB IS A CP SYSTEM AND CASSANDRA IS AN AP SYSTEM. CAP THEOREM PROVIDES AN OVERLY SIMPLIFIED VIEW OF TODAY'S DISTRIBUTED SYSTEMS SUCH AS MONGODB AND CASSANDRA. UNDER NORMAL OPERATIONS, AVAILABILITY AND CONSISTENCY ARE ADJUSTABLE AND CAN BE CONFIGURED TO MEET SPECIFIC REQUIREMENTS.

COMPARISON OF MONGODB WITH RDB:

MONGODB

- It is a non-relational and document-oriented database.
- It is suitable for hierarchical data storage .
- It has a dynamic schema .
- It centers around the CAP theorem (Consistency, Availability, and Partition tolerance)In terms of performance, it is much

faster than RDBMS.

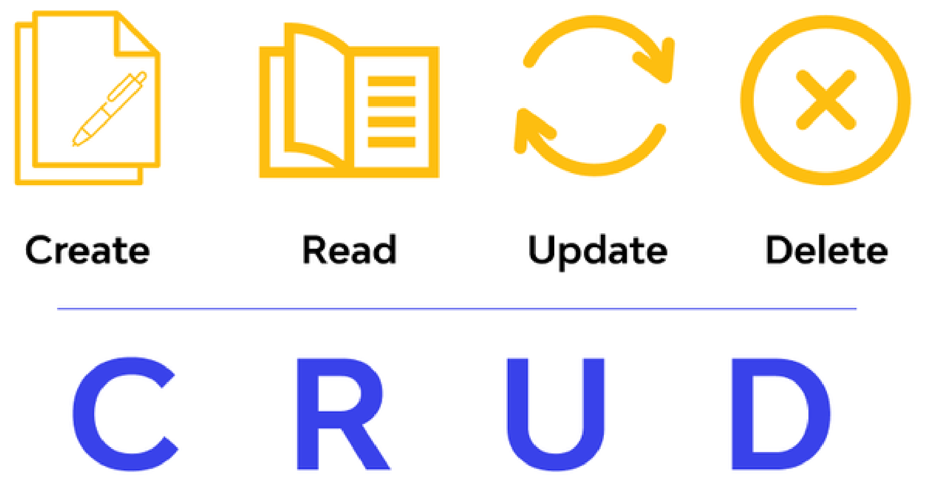
RDB

- It is a relational database.
- It is not suitable for hierarchical data storage .
- It has a predefined schema .
- It centers around ACID properties (Atomicity, Consistency,

Isolation, and Durability)In terms of performance, it is slower than MongoDB.

CRUD OPERATIONS

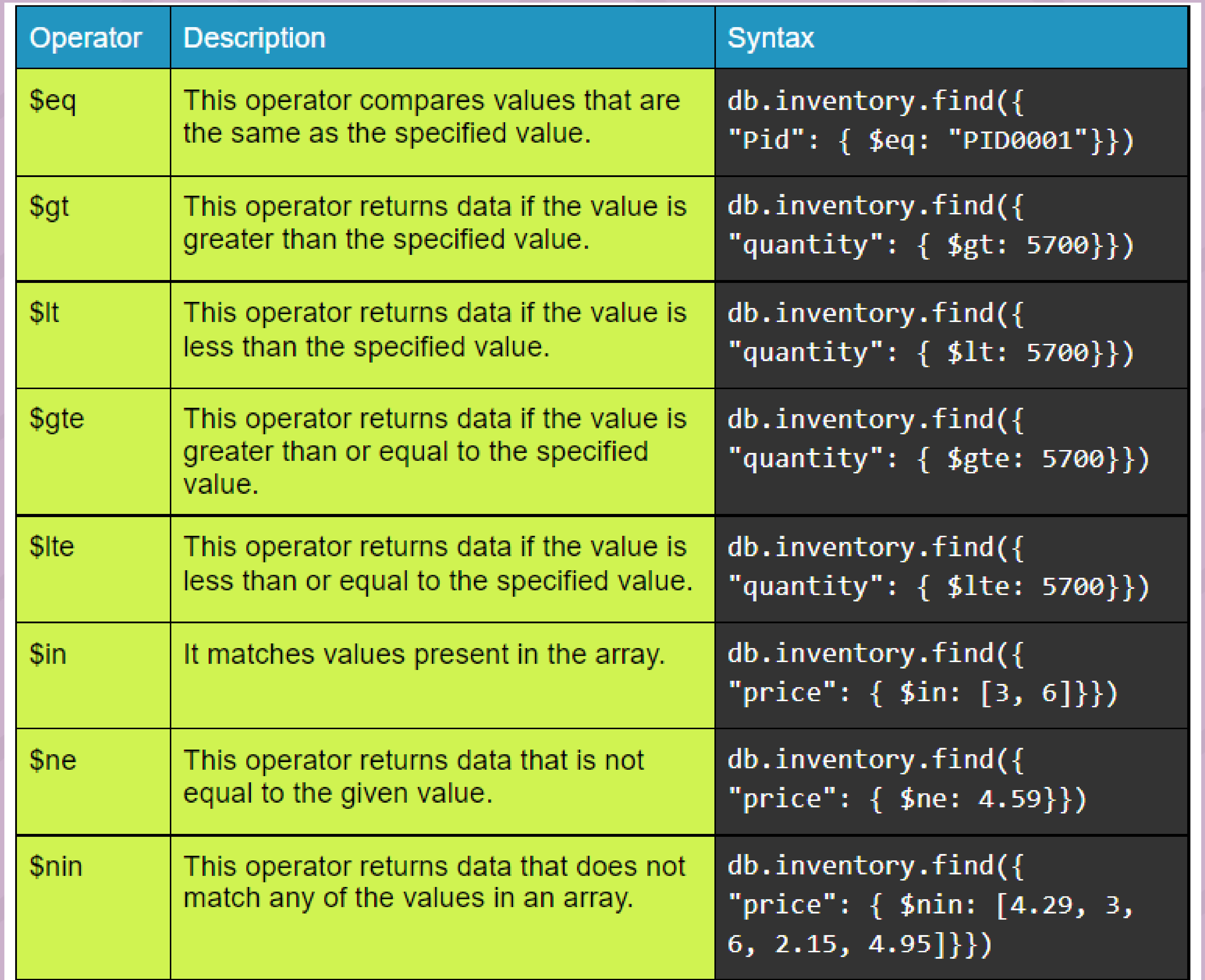
CRUD refers to the four basic operations a software application should be able to perform – Create, Read, Update, and Delete.



WHAT IS CRUD OPERATIONS IN MONGODB?

- THE CREATE OPERATION IS USED TO INSERT NEW DOCUMENTS IN THE MONGODB DATABASE.
- THE READ OPERATION IS USED TO QUERY A DOCUMENT IN THE DATABASE.
- THE UPDATE OPERATION IS USED TO MODIFY EXISTING DOCUMENTS IN THE DATABASE.
- THE DELETE OPERATION IS USED TO REMOVE DOCUMENTS IN THE DATABASE.

EX QUERY OPERATIONS IN MONGODB:



Operator	Description	Syntax
$eq	This operator compares values that are the same as the specified value.	`db.inventory.find({ "Pid": { $eq: "PID0001"}})`
$gt	This operator returns data if the value is greater than the specified value.	`db.inventory.find({ "quantity": { $gt: 5700}})`
$lt	This operator returns data if the value is less than the specified value.	`db.inventory.find({ "quantity": { $lt: 5700}})`
$gte	This operator returns data if the value is greater than or equal to the specified value.	`db.inventory.find({ "quantity": { $gte: 5700}})`
$lte	This operator returns data if the value is less than or equal to the specified value.	`db.inventory.find({ "quantity": { $lte: 5700}})`
$in	It matches values present in the array.	`db.inventory.find({ "price": { $in: [3, 6]}})`
$ne	This operator returns data that is not equal to the given value.	`db.inventory.find({ "price": { $ne: 4.59}})`
$nin	This operator returns data that does not match any of the values in an array.	`db.inventory.find({ "price": { $nin: [4.29, 3, 6, 2.15, 4.95]}})`

INTRO OF COUCHDB:

-IT IS A NOSQL DOCUMENT STORE DATABASE
-THE EMPHASIS IS ON USABILITY AND EMBRACING THE WEB
-THE APACHE SOFTWARE FOUNDATION CREATED THE OPEN-SOURCE DATABASE COUCHDB
-YOU MAY TRANSFER, EXCHANGE, AND SYNCHRONIZE DATA BETWEEN DATABASES AND MACHINES BY UTILIZING COUCHDB'S SIMPLE REPLICATION.
-ADDITIONALLY, THE STRAIGHTFORWARD HTTP RESOURCE AND METHOD STRUCTURES (GET, PUT, DELETE) MAKE THEM SIMPLE TO COMPREHEND AND USE

FUNCTIONALITY AND DESIGN:

-FAST INDEXING AND RETRIEVAL
-REST-LIKE INTERFACE FOR DOCUMENT INSERTION, UPDATES, RETRIEVAL, AND DELETION REST IS AN ACRONYM FOR REPRESENTATIONAL STATE TRANSFER
-IT OFFERS THE ABILITY TO STORE DOCUMENTS WITH UNIQUE NAMES AND A RESTFUL HTTP API THAT CAN BE USED TO READ AND UPDATE (ADD, MODIFY, AND REMOVE) DATABASE DOCUMENTS.

WHY AND WHEN YOU USE IT?

To-do app or list of items, Comments & Replies on blogs, Service logs, Data from instruments and/or GPS are enough good use-cases for CouchDB. Good use of CouchDB and other NoSQL repositories is for synchronizing data in an occasionally connected which makes sense in industries like finance, data collection, log analysis, Iot fields. the model is also beneficial for companies that have a distributed workforce like real estate or construction. High velocity log writes, high performance caching are well suited situations for CouchDB.

CAP THEOREM

AP DATABASE: AN AP DATABASE DELIVERS AVAILABILITY AND PARTITION TOLERANCE AT THE EXPENSE OF CONSISTENCY.

COMPARISON COUCHDB WITH RDB:

COUCHDB

- It is written in the Erlang language
- It does support XML data format
- It is developed by Apache Software Foundation and was initially released in 2005
- It can handle only one connection at a time
- It has no Data Schema
- In CouchDB, there are no predefined datatypes

RDB

- It is developed in C and C++ languages
- It does not support XML data format
- It is developed by Oracle in May 1995
- It can handle multiple connections simultaneously
- It has Data Schema
- It supports predefined data types such as float, date, number, etc..

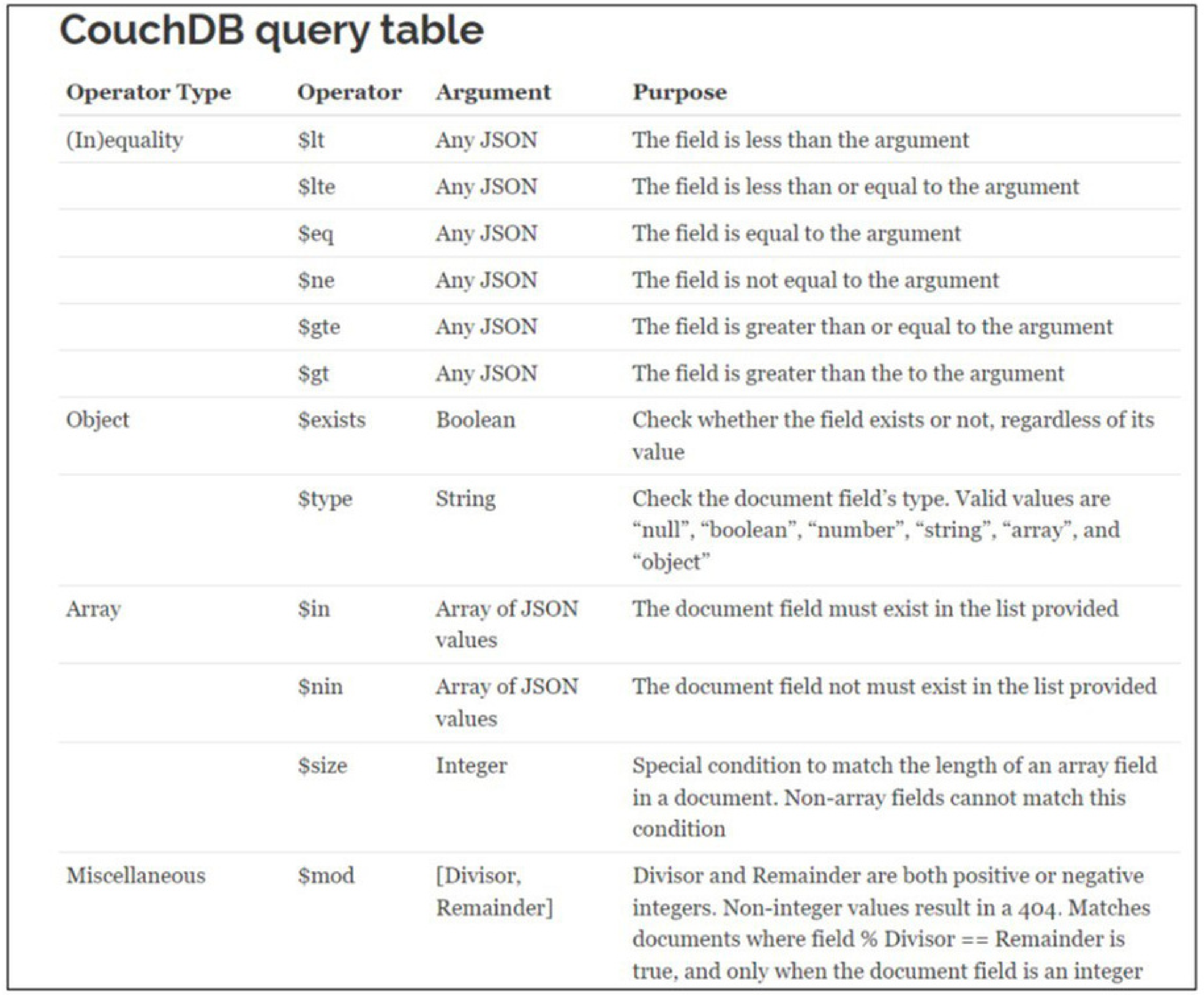
FEATURES:

- HIGH AVAILABILITY: COUCHDB PROVIDES HIGH AVAILABILITY THANKS TO ITS BUILT-IN REPLICATION FEATURE
- SCHEMA-FREE: THERE IS NO NEED TO DEFINE A SCHEMA BEFORE YOU START USING COUCHDB.
- SCALABLE AND REPLICABLE: COUCHDB IS HIGHLY SCALABLE, ALLOWING IT TO STORE AND REPLICATE LARGE AMOUNTS OF DATA EASILY.
- QUERYABLE: DATA STORED IN COUCHDB CAN BE QUERIED USING VIEWS WRITTEN IN JAVASCRIPT.
- RESTFUL API: COUCHDB OFFERS A RESTFUL API THAT MAKES IT EASY TO INTERACT WITH THE DATABASE REMOTELY.

CRUD OPERATIONS:

COUCHDB USES A RESTFUL API TO ACCESS THE DATABASE FROM ANYWHERE, WITH FULL CRUD (CREATE, READ, UPDATE, DELETE) OPERATIONS FLEXIBILITY. THIS SIMPLE AND EFFECTIVE MEANS OF DATABASE CONNECTIVITY MAKES COUCHDB FLEXIBLE, FAST, AND POWERFUL TO USE WHILE REMAINING HIGHLY ACCESSIBLE.

QUERY OPERATION:



CouchDB query table

Operator Type	Operator	Argument	Purpose
(In)equality	$lt	Any JSON	The field is less than the argument
	$lte	Any JSON	The field is less than or equal to the argument
	$eq	Any JSON	The field is equal to the argument
	$ne	Any JSON	The field is not equal to the argument
	$gte	Any JSON	The field is greater than or equal to the argument
	$gt	Any JSON	The field is greater than the to the argument
Object	$exists	Boolean	Check whether the field exists or not, regardless of its value
	$type	String	Check the document field's type. Valid values are "null", "boolean", "number", "string", "array", and "object"
Array	$in	Array of JSON values	The document field must exist in the list provided
	$nin	Array of JSON values	The document field not must exist in the list provided
	$size	Integer	Special condition to match the length of an array field in a document. Non-array fields cannot match this condition
Miscellaneous	$mod	[Divisor, Remainder]	Divisor and Remainder are both positive or negative integers. Non-integer values result in a 404. Matches documents where field % Divisor == Remainder is true, and only when the document field is an integer

INTRO OF COUCHBASEDB:

WHAT IS IT? AN OPEN-SOURCE, DISTRIBUTED MULTI-MODEL NOSQL DOCUMENT-ORIENTED DATABASE SOFTWARE PACKAGE, FORMERLY KNOWN AS MEMBASE, IS DESIGNED FOR INTERACTIVE APPLICATIONS.

FUNCTIONALITY AND DESIGN:

THE DATABASE SUPPORTS CAPABILITIES INCLUDING KEY-VALUE, SQL++, FULL-TEXT SEARCH, ANALYTICS, EVENTING, CROSS DATA CENTER REPLICATION, AND EVEN MORE.

WHY AND WHEN YOU USE IT?

Couchbase has a variety of features and services that most NoSQL competitors can't match. Couchbase is widely used for web, mobile, and Iot applications.

CAP THEOREM

COUCHBASE IS NORMALLY A CP TYPE SYSTEM MEANING IT PROVIDES CONSISTENCY AND PARTITION TOLERANCE, OR IT CAN BE SET UP AS AN AP SYSTEM WITH MULTIPLE CLUSTERS.

COMPARISON COUCHBASEDB WITH RDB:

COUCHBASEDB

- is a non-relational and document-oriented database.
- is field-based.
- is faster.
- is a dynamic schema.

RDB

- is a relational database.
- is column-based.
- has a predefined schema.

FEATURES:

- IT HAS A FLEXIBLE DATA MODEL, IS EASILY SCALABLE, PROVIDES CONSISTENT HIGH PERFORMANCE AND IS CAPABLE OF SERVING APPLICATION DATA WITH 100% UPTIME. IT IS ALSO RELIABLE AND SECURE

CRUD OPERATIONS:

THE COUCHBASE DESTINATION CAN INSERT, UPDATE, DELETE, OR UPSERT DATA. THE DESTINATION WRITES THE RECORDS BASED ON THE CRUD OPERATION DEFINED IN A CRUD OPERATION HEADER ATTRIBUTE OR IN OPERATION-RELATED STAGE PROPERTIES. YOU CAN DEFINE THE CRUD OPERATION IN A CRUD OPERATION RECORD HEADER ATTRIBUTE. THE DESTINATION LOOKS FOR THE CRUD OPERATION TO USE IN THE SDC.OPERATION.TYPE RECORD HEADER ATTRIBUTE.

QUERY OPERATION:

COUCHBASE SERVER CAN BE QUERIED USING N1QL, THE COUCHBASE SERVER QUERY LANGUAGE. AS ITS PRIMARY FUNCTION, THE QUERY SERVICE ENABLES YOU TO ISSUE QUERIES TO EXTRACT DATA FROM COUCHBASE SERVER.
YOU CAN ALSO ISSUE QUERIES FOR DATA DEFINITION (DEFINING INDEXES) AND DATA MANIPULATION (ADDING OR DELETING DATA).

Thanks!

Any questions? Don't hesitate to ask for our help