

Document DBs & MongoDB

Monday, February 10, 2025 9:08 AM

Document Database is a non-relational database that stores data as structured docs, usually in JSON

- Simple, flexible, and scalable design

```
{
  "orders": [
    {
      "orderid": "748745375",
      "date": "June 30, 2008 1:54:23 AM",
      "trackingno": "TMO039291",
      "custid": "11045",
      "customer": {
        "custid": "11045",
        "fname": "Sue",
        "lname": "Sachfield",
        "address": "1409 Silver Street",
        "city": "Ashland",
        "state": "NE",
        "zip": "68003"
      }
    }
  ]
}
```

- **JSON** (JavaScript Object Notation)
 - a lightweight data-interchange format
 - It is easy for humans to read and write.
 - It is easy for machines to parse and generate.
- JSON is built on two structures:
 - A collection of name/value pairs. In various languages, this is operationalized as an object, record, struct, dictionary, hash table, keyed list, or associative array.
 - An ordered list of values. In most languages, this is operationalized as an array, vector, list, or sequence.
- These are two universal data structures supported by virtually all modern programming languages
 - Thus, JSON makes a great data interchange format.
- **BSON** → Binary JSON
 - binary-encoded serialization of a JSON-like document structure
 - supports extended types not part of basic JSON (e.g. Date, BinaryData, etc)
 - **Lightweight** - keep space overhead to a minimum
 - **Traversable** - designed to be easily traversed, which is vitally important to a document DB
 - **Efficient** - encoding and decoding must be efficient
 - Supported by many modern programming languages

```
{ "hello": "world" } →
\x16\x00\x00\x00          // total document size
\x02                       // 0x02 = type String
hello\x00                  // field name
\x06\x00\x00\x00world\x00  // field value
\x00                       // 0x00 = type E00 ('end of object')
```

- XML → eXtensible Markup Language
 - Precursor to JSON as data exchange format
 - XML + CSS → web pages that separated content and formatting
 - Structurally similar to HTML, but tag set is extensible

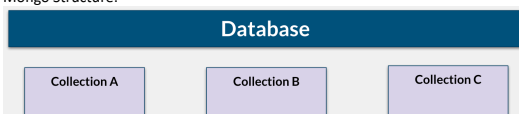
```
<CATALOG>
<CD>
  <TITLE>Empire Burlesque</TITLE>
  <ARTIST>Bob Dylan</ARTIST>
  <COUNTRY>USA</COUNTRY>
  <COMPANY>Columbia</COMPANY>
  <PRICE>10.90</PRICE>
  <YEAR>1985</YEAR>
</CD>
<CD>
  <TITLE>Hide your heart</TITLE>
  <ARTIST>Bonnie Tyler</ARTIST>
  <COUNTRY>UK</COUNTRY>
  <COMPANY>CBS Records</COMPANY>
  <PRICE>9.90</PRICE>
  <YEAR>1988</YEAR>
</CD>
</CATALOG>
```

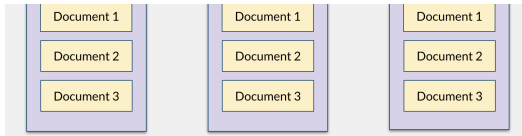
Why Document Databases?

- Document databases address the impedance mismatch problem between object persistence in OO systems and how relational DBs structure data.
 - OOP → Inheritance and Composition of types.
 - How do we save a complex object to a relational database?
 - We basically have to deconstruct it.
- The structure of a document is self-describing.
- They are well-aligned with apps that use JSON/XML as a transport layer

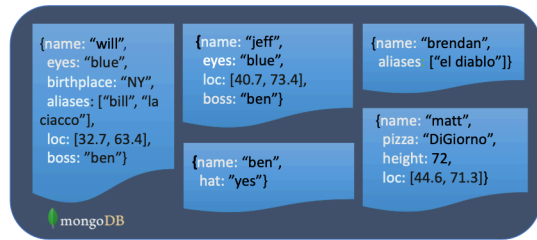
MongoDB

- Started in 2007 after Doubleclick was acquired by Google, and 3 of its veterans realized the limitations of relational databases for serving > 400,000 ads per second
- MongoDB was short for Humongous Database
- MongoDB Atlas released in 2016 → documentdb as a service
- Mongo Structure:





• Mongo Documents:



- No predefined schema for documents is needed
- Every document in a collection could have different data/schema

Relational vs Mongo

RDBMS	MongoDB
Database	Database
Table/View	Collection
Row	Document
Column	Field
Index	Index
Join	Embedded Document
Foreign Key	Reference

Mongo Features

- Rich Query Support - robust support for all CRUD ops
- Indexing - supports primary and secondary indices on document fields
- Replication - supports replica sets with automatic failover
- Load balancing built in

How to Interact

- **mongosh** → MongoDB Shell
 - CLI tool for interacting with a MongoDB instance
- MongoDB Compass
 - free, open-source GUI to work with a MongoDB database
- DataGrip and other 3rd Party Tools
- Every major language has a library to interface with MongoDB
 - PyMongo (Python), Mongoose (JavaScript/node), ...

