**SEFP ASSIGNMENT-5(MONGO DB)**

**INTRODUCTION:**

Mongo DB is an open-source document database that provides high performance, high availability, and automatic scaling. Mongo DB obviates the need for an Object Relational Mapping (ORM) to facilitate development. Mongo DB has rapidly grown to become a popular database for web applications and is a perfect fit for Node.JS applications, letting you write Java script for the client, backend and database layer. Its schemaless nature is a better match to our constantly evolving data structures in web applications, and the integrated support for location queries is a bonus that’s hard to ignore. Throw in Replica Sets for scaling, and we’re looking at really nice platform to grow your storage needs now and in the future.

**FEATURES:**

* **General purpose database**
* **High availability**
* **Scalability**
* **Aggregation**
* **Load Balancing**
* **Native Replication**
* **Security**
* File storage
* Load balancing
* Aggregation
* Server-side JavaScript execution

**ARCHITECTURE:**

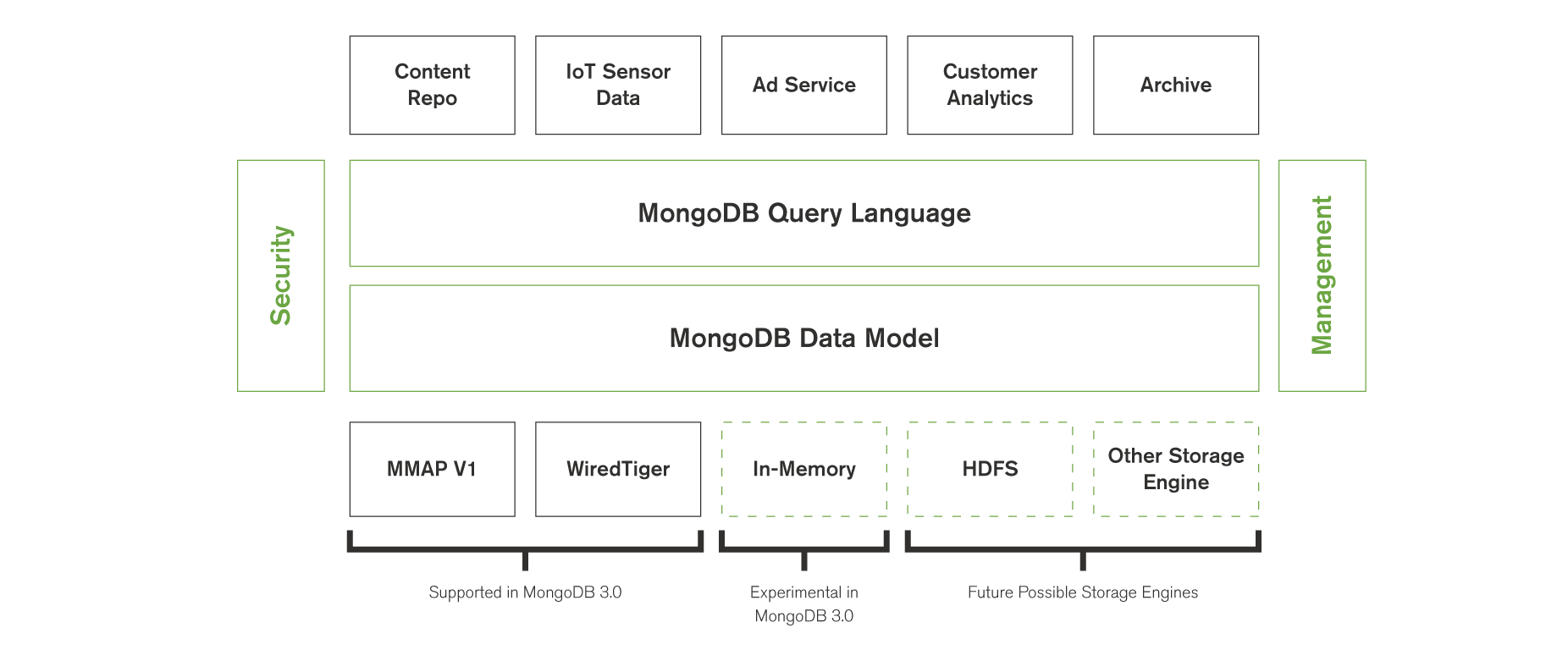
MongoDB embraces two key trends in modern IT:

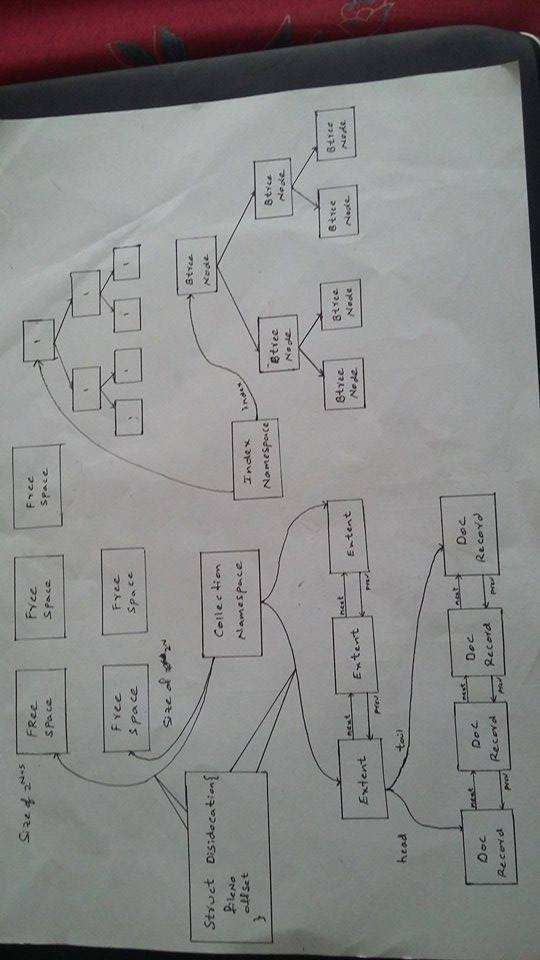
* **New apps.** Organizations are expanding the range of applications they deliver to support the business.
* **Technology rationalization.** CIOs are rationalizing their technology portfolios to a strategic set of vendors they can leverage to more efficiently support their business.

With MongoDB, organizations can address diverse application needs, hardware resources, and deployment designs with a single database technology. Through the use of a pluggable storage architecture, MongoDB can be extended with new capabilities, and configured for optimal use of specific hardware architectures. This approach significantly reduces developer and operational complexity compared to running multiple databases to power applications with unique requirements. Users can leverage the same MongoDB query language, data model, scaling, security and operational tooling across different applications, each powered by different pluggable MongoDB storage engines.

**Storage efficiency with compression:**

MongoDB supports native compression when configured with the WiredTiger storage engine, reducing physical storage footprint by as much as 80%. In addition to reduced storage space, compression enables much higher storage I/O scalability as fewer bits are read from disk. Administrators have the flexibility to configure specific compression algorithms for collections, indexes and the journal.





## MongoDB Consistency & Availability

### Transaction model:

### One or more fields may be written in a single operation, including updates to multiple sub-documents and elements of an array. The ACID guarantees provided by MongoDB ensure complete isolation as a document is updated; any errors cause the operation to roll back so that clients receive a consistent view of the document.

**Replica sets**

MongoDB maintains multiple copies of data called replica sets using native replication. A replica set is a fully self-healing shard that helps prevent database downtime.The number of replicas in a MongoDB replica set is configurable: a larger number of replicas provide increased data availability and protection against database downtime (e.g., in case of multiple machine failures, rack failures, data center failures, or network partitions). Optionally, operations can be configured to write to multiple replicas before returning to the application, thereby providing functionality that is similar to synchronous replication.

In-memory performance with on-disk capacity

MongoDB makes extensive use of RAM to speed up database operations. Reading data from memory is approximately 100,000 times faster than reading data from disk. In MongoDB, all data is read and manipulated through memory-mapped files. Data that is not accessed is not loaded into RAM. Because MongoDB provides in-memory performance, for most applications there is no need for a separate caching layer to scale your database.

### Security:

* **Authentication.** Simplifying access control to the database, MongoDB offers integration with external security mechanisms including LDAP, Windows Active Directory, Kerberos and x.509 PKI certificates.
* **Authorization.** User-defined roles enable administrators to configure granular permissions for a user or application, based on the privileges they need to do their job. Additionally, field-level redaction can work with trusted middleware to manage access to individual fields within a document, allowing the co-location of data with multiple security levels for ease of development and operation.
* **Auditing.** For regulatory compliance, security administrators can use MongoDB's native audit log to track access and operations performed against the database.
* **Encryption.** MongoDB data can be encrypted on the network and on disk. Support for SSL allows clients to connect to MongoDB over an encrypted channel

## Advantages of MongoDB:

* Schema less : MongoDB is document database in which one collection holds different different documents. Number of fields, content and size of the document can be differ from one document to another.
* Structure of a single object is clear
* No complex joins
* Deep query-ability. MongoDB supports dynamic queries on documents using a document-based query language that's nearly as powerful as SQL
* Tuning
* Ease of scale-out: MongoDB is easy to scale
* Conversion / mapping of application objects to database objects not needed
* Uses internal memory for storing the (windowed) working set, enabling faster access of data

**CONCLUSION:**

In summary, MongoDB has been built to be **fast**(no joins but embedded documents),**flexible** (schema less), **scalable**(horizontal no vertical), to minimize administrative tasks (replication, disaster recovery, automatic failover, sharding, load balancing, etc),**easy to learn and with powerful tools for data analysis** (aggregation framework). Big companies like Telefónica, Bosch, Coca Cola, CISCO, Leroy Merlin, ebay, IBM, Forbes and many others already use MongoDB. Cities like Chicago are using a system composed of MongoDB and Hadoop to analyze data and make quick decisions in order to achieve a better quality of life for their citizens.