**MongoDB – A comparison with NoSQL databases – A Literature Review**

Due to massive cost of storing and manipulating data in classical relational database systems, NoSQL databases have been developed. This paper describes the advantages of MongoDB when compared to other NoSQL databases and its applications in sentiment analysis. MongoDB is a document based NoSQL database designed for Internet and web based applications. (Hema Krishnan, 2016)

**Introduction**

Common features of NoSQL products are the divergence from the relational data model, simplification of transactional model and transaction processing and most importantly the shift to the imperative programming model from the declarative style SQL language. But applications those require data and or functional partitioning, either because of the sheer size of the data or for the purpose of load balancing have to rely on custom built relations or utilize alternative database systems. (Hema Krishnan, 2016)

Relational data bases are often being replaced by other viable alternatives, such as NoSQL databases, for reasons of scalability and heterogeneity.

**No SQL Data Models**

1. Key-values Stores: The main idea here is using a hash table where there is a unique key and a pointer to a particular item of data.
2. Column Family Stores: These were created to store and process very large amounts of data distributed over many machines.
3. Document Databases: These were inspired by Lotus Notes and are similar to key-value stores. Document databases are essentially the next level of Key/value, allowing nested values associated with each key.

**MongoDB**

The increasing popularity of MongoDB and the large amounts of user-related sensitive information stored in these databases raise the concern for the confidentiality and privacy of the data and the security provided by these systems. This allows MongoDB to build indexes and match objects against query expressions on both top-level and nested BSON keys. MongoDB’s document data model makes it easy to build on, since it has inherent support for unstructured data and does not require costly and time consuming migrations when application requirements change. BSON at first seems BLOB like, but there exists an important difference: the MongoDB database understands BSON internals. This means that MongoDB can reach inside BSON objects, even nested ones, using dot notation. (Hema Krishnan, 2016)

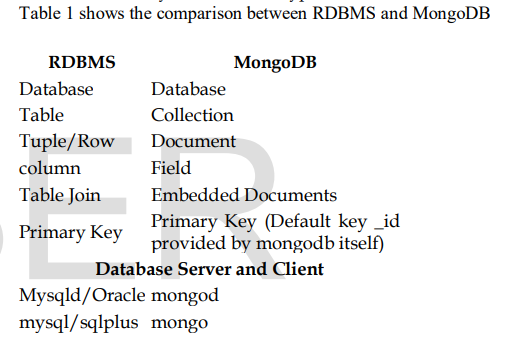
MongoDB uses BSON as network transfer format for documents.

**Components of MongoDB**

MongoDB is a cross-platform, document oriented database that provides, high performance, high availability, and easy scalability. MongoDB works on concept of collection and document. Database is a physical container for collections. (Hema Krishnan, 2016)

Collection: Collection is a group of MongoDB documents. Typically, all documents in a collection are of similar or related purpose.

Document: Dynamic schema means that documents in the same collection do not need to have the same set of fields or structure, and common fields in a collection's documents may hold different types of data.



**Features of MongoDB**

Document Oriented Storage: Data is stored in the form of JSON style documents Index on any attribute. (Hema Krishnan, 2016)

Replication & High Availability: Replica set can have only one primary node. Replica set is a group of two or more nodes (generally minimum 3 nodes are required). In a replica set one node is primary node and remaining nodes are secondary. In a replica one node is primary node that receives all write operations. At the time of automatic failover or maintenance, election establishes for primary and a new primary node is elected. (Hema Krishnan, 2016)

All data replicates from primary to secondary node.

Auto – Sharding: Sharding is the process of storing data records across multiple machines and it is MongoDB's approach to meeting the demands of data growth. As the size of the data increases, a single machine may not be sufficient to store the data nor provide an acceptable read and write throughput. (Hema Krishnan, 2016)

* Rich Queries
* Fast In-Place Updates

**Conclusion**

NoSQL systems offer much less functionality than traditional relation database management systems, especially in transaction isolation and scan operations. MongoDB is an effective document oriented database which can be used for tweet analysis and other applications. (Hema Krishnan, 2016)

# Bibliography

Hema Krishnan, M. E. (2016). *MongoDB – a comparison with NoSQL databases*.

Related Links- https://www.researchgate.net/publication/327120267\_MongoDB\_-\_a\_comparison\_with\_NoSQL\_databases