



# MarkLogic Database - Only Enterprise

## NoSQL DB

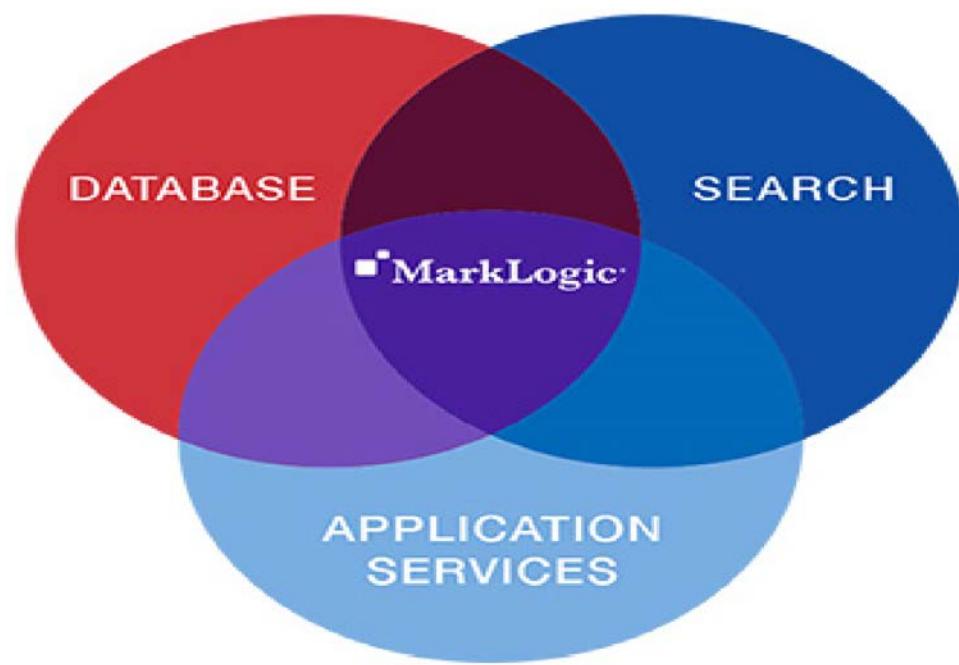
Sanket V. Patel, Aashi Rastogi

Department of Computer Science

University of Bridgeport, Bridgeport, CT

### Abstract

MarkLogic DB is one of the Enterprise NoSQL database that supports multiple-model database design. It is optimized for structured and unstructured data that allows to store, manage, query and search across JSON, XML, RDF (Triplestore) and can handle data with a schema free and leads to faster time-to-results by providing handling of different types of data. It provides ACID Transactions using MVCC (multi-version concurrency control). One of the important key feature of MarkLogic is its Bitemporal behavior by providing data at every point in time. Due to its shared-nothing architecture it is highly available and easily and massively scalable with no single point of failure making structured data integration easier. It also has incremental backup means to only backup the updated data. Marklogic provides Hadoop integration and Hadoop is designed to store large amount of data in Hadoop Distributed File System (HDFS) and works better with the transactional applications.



### Introduction

NoSQL means non-SQL or non-relational databases which provides mechanism to store and retrieve the data other than relational databases. NoSQL database is in use nowadays because of simplicity of design, easy to scale out and control over availability.

Types of NoSQL Databases:

- Key-value based
- Column oriented
- Graph oriented
- Document based
- Multi-model

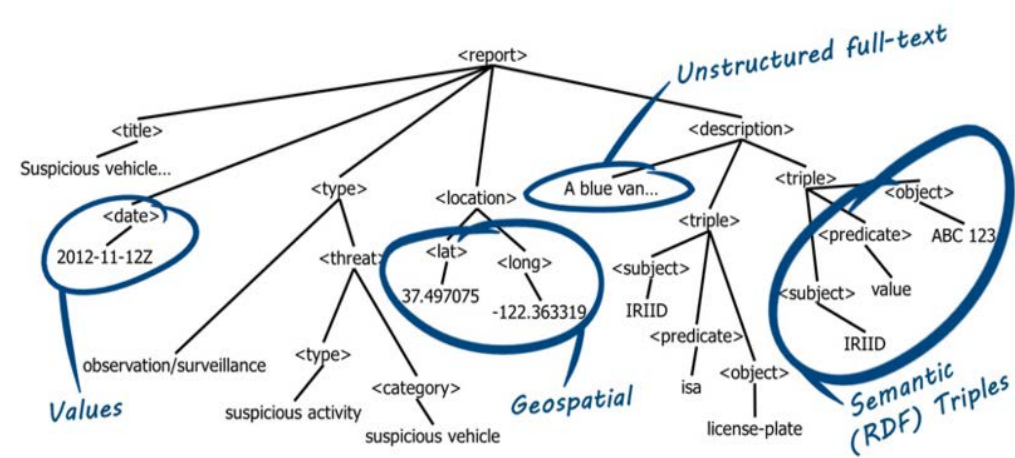
Multi-model database is an only designed to support multiple data models against a single application. Marklogic DB is one of the NoSQL database that uses multi-model database design.

Marklogic is the only Enterprise NoSQL Database. It is optimized for structured and unstructured data that allows you to store, manage, query and search across JSON, XML, RDF (Triplestore), Geospatial data, text, and large binaries. With Marklogic one can handle data in a schema-agnostic fashion or built in application server and leads to faster time-to-results.

It provides capabilities like ACID Transaction, high availability and disaster recovery, Security. Marklogic is designed to run on Hadoop and help you to use their technology in better way. It is also easily deployed on cloud to maintain hardware and provide all benefits of elasticity. It also has built in application services and text search capabilities. It allows to discover new facts by act.

Below are the segregate characteristics which describes the Mark Logic so precisely:

- 1)Flexible Data Model
- 2)Clear Semantics
- 3)Scalability and Elasticity
- 4)ACID Transaction
- 5)High availability and Disaster Recovery
- 6)Hadoop Integration
- 7)Bitemporal
- 8)Certified Security



### How Marklogic Works?

It uses XML document as its data model, and stores the documents within a transactional repository. It indexes the words and values from each of the loaded documents, as well as the document structure. And, because of its unique Universal Index, Marklogic doesn't require advance knowledge of the document structure nor complete adherence to a particular schema. Marklogic Server clusters on commodity hardware using a shared-nothing architecture and differentiates itself in the market by supporting massive scale and fantastic performance-customer deployments have scaled to hundreds of terabytes of source data while maintaining sub-second query response time.

In addition to XML, Marklogic can store JSON, text, and binary documents. JSON documents are internally transformed to XML for purposes of indexing. Text documents are indexed as if each was an XML text node without a parent. Binary documents are by default unindexed, with the option to index their metadata and extracted contents.

#### The Universal Index

- ✓ Words
- ✓ Phrases
- ✓ Stemmed words and phrases
- ✓ Structure
- ✓ Words and phrases in the context of structure
- ✓ Values
- ✓ Collections
- ✓ Security Permissions

| Term                              | Term List                         |
|-----------------------------------|-----------------------------------|
| "blue"                            | 123, 125, 126, 130, 132, 344, ... |
| "van"                             | 123, 125, 126, 129, 130, 132, ... |
| "observed"                        | 125, 126, 516, 522, 765, 890, ... |
| "blue van"                        | 123, 125, 129, 130, 132, 496, ... |
| STEM "observe"                    | 125, 126, 516, 522, 765, 890, ... |
| <report>                          | ...                               |
| <report>/<location>               | ...                               |
| <report>/<category>               | ...                               |
| <type>=suspicious activity/<type> | ...                               |
| <date>=2012-11-12Z/<date>         | ...                               |
| Collection/Method                 | ...                               |
| Role/Analyst + Action/Read        | ...                               |
| ...                               | ...                               |
| ...                               | ...                               |



### Implementation

### Conclusion

Marklogic is designed to handle the volume, variety, and velocity of Big Data like other NoSQL solutions, and has the enterprise features that made last-generation relational databases so reliable. And Marklogic gives a way to manage the hierarchical content, distributed graph data, and also XML and RDF in the same database. MarkLogic combines database, search, and application services in one component to greatly simplify the delivery platform, and brings the power of Enterprise NoSQL with schema flexibility to easily add new content and data; horizontal scale out to grow as your business and products grow; and proven enterprise features to operate your product. Eight of the top 10 information providers in the world use MarkLogic to power their information delivery platforms – and we can help you deliver better information to your users, too.

“ Standardized search technology not only provides a better experience for our customers, but also gives us a superior platform for developing new content applications and products across the company.”

