MarkLogic Enterprise NoSQL

Do more with less!

Mats Stellwall – Principal Sales Engineer
Rickard Fernström – Nordic Sales Manager

MarkLogic

- **Key-Value**
- **Graph/Triple Store**
- **Document Oriented**
- **Column**

MarkLogic

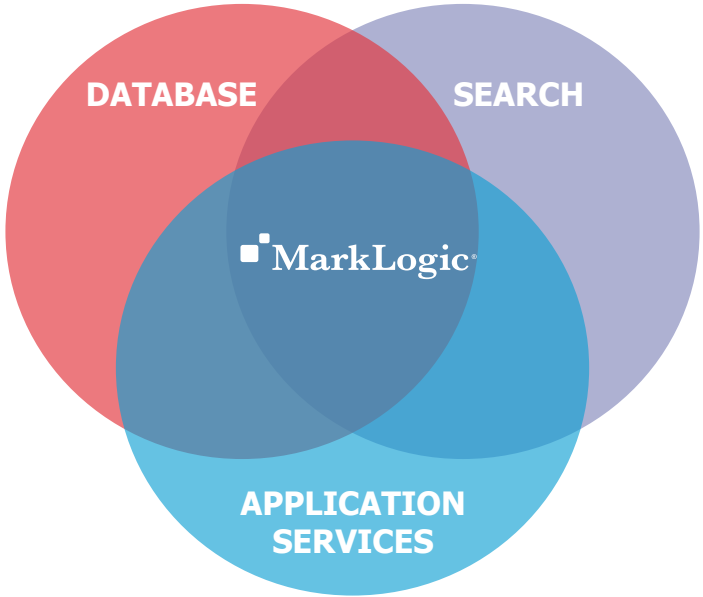
NoSQL Databases

- **Schema Flexibility**: Ability to query multiple schema at run
- **Free of Complex Joins**: Generally-designed so joins aren't necessary or recommended
- **Horizontally Scalable**: Easy to incorporate new machines (of comparable power) into cluster
- **Compatible with Commodity Hardware**: Most NoSQL database systems run on commodity hardware
- **Self-contained**: Nothing is shared – each node is a stand-alone with respect to storage and processing power.
- **Rapid Application Development:** Data is always ready – whenever you are

MarkLogic

The Only Enterprise NoSQL Database

- Search & Query
- ACID Transactions
- High Availability / Disaster Recovery
- Replication
- Government-grade Security
- Scalability & Elasticity
- On-premise or Cloud Deployment
- Hadoop for Storage & Compute
- Semantics/triple store
- Faster Time-to-Results



MarkLogic / Enterprise NoSQL Database Platform

POWERFUL

- Native JSON Store
- Native XML Store
- Native RDF Triple Store
- Geospatial Support
- Full-text Search
- Flexible Indexes
- Bitemporal
- Real-time Alerting
- Semantic Inference
- Tiered Storage
- Server-side JavaScript
- Fully Transactional

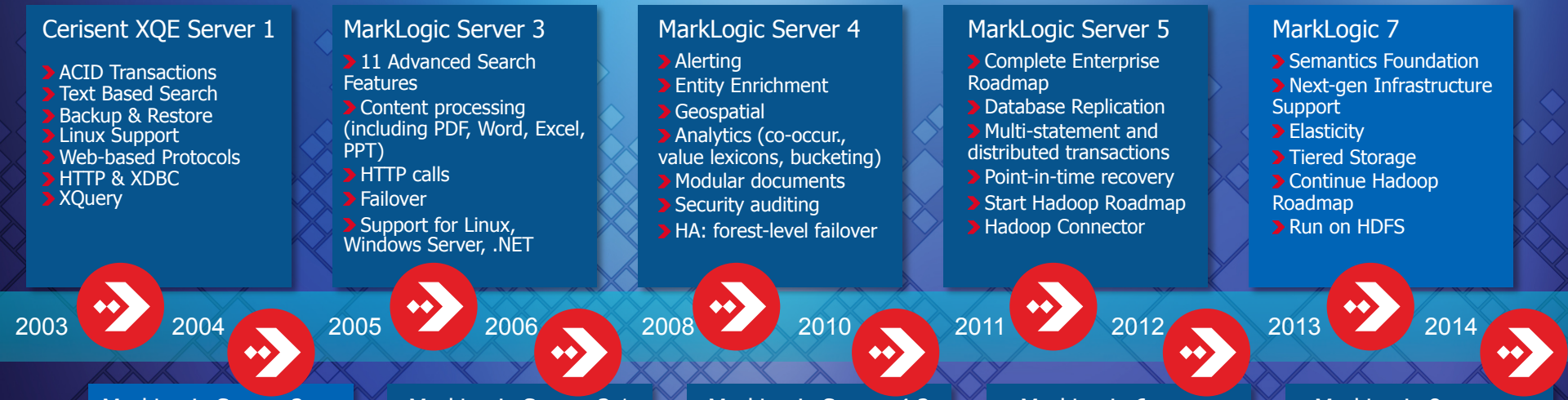
AGILE

- Scalable and Elastic
- Cloud Ready (AWS)
- Hadoop and HDFS
- REST API
- SQL Support
- Multi-OS Support
- Schema Agnostic
- Samplestack
- MarkLogic Content Pump
- XA Transactions
- Ad-hoc Queries
- Index Across Data Types

TRUSTED

- Performance at scale
- LDAP and Kerberos Security
- Security Certifications
- Monitoring and Management
- Configuration Management
- 24/7 Engineering Support
- ACID Transactions
- Flexible Replication
- Customizable Backup
- Customizable Failover
- Point-in-time Recovery
- Atomic Forests

MarkLogic Product evolution for over a decade



MarkLogic

How Does it Work?

Schema-agnostic design

Real-time indexing for search and query

Reverse queries for event processing and alerting

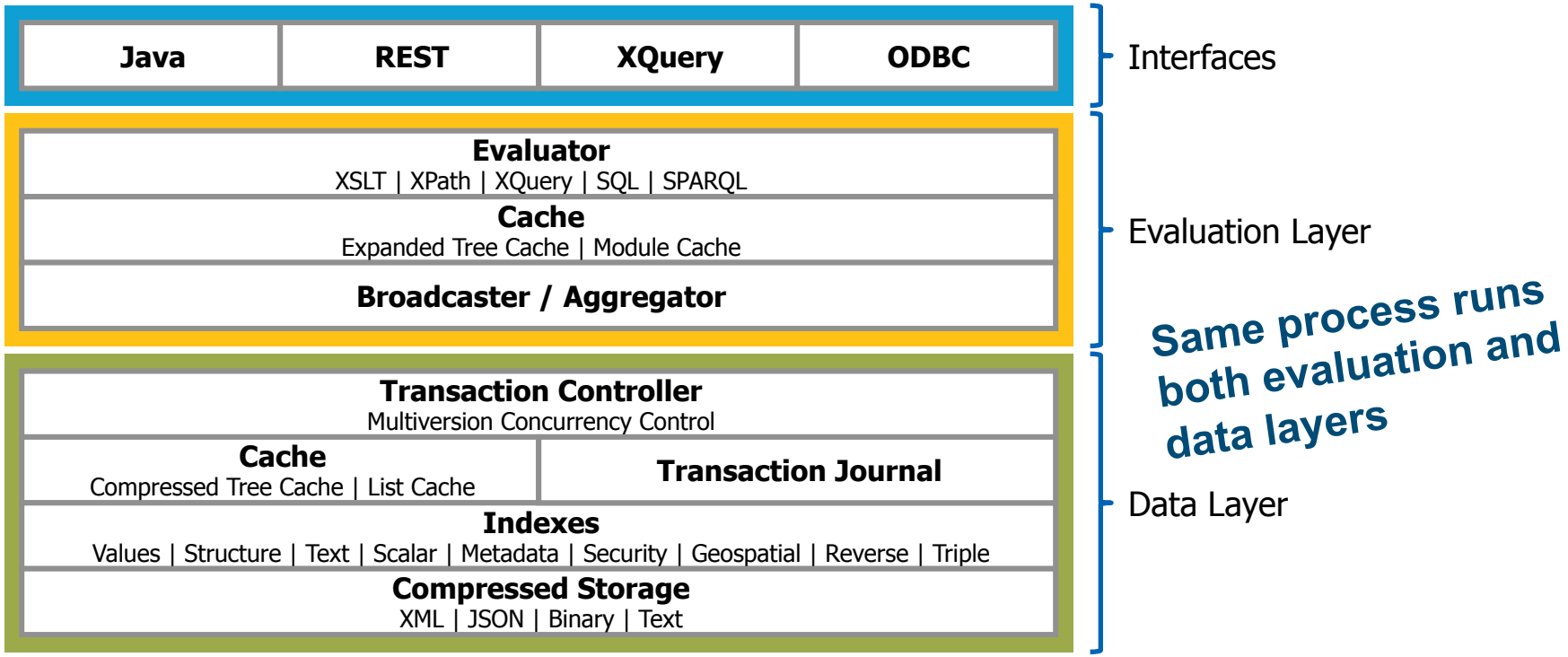
Scale-out shared-nothing cluster topology

Analytics and Visualization

Transactions

MarkLogic

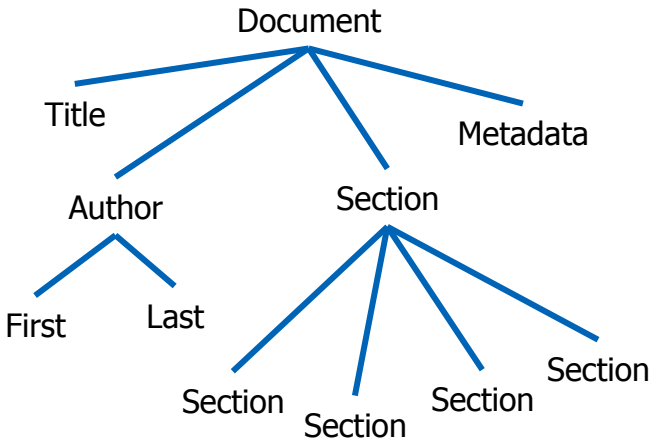
MarkLogic Architecture

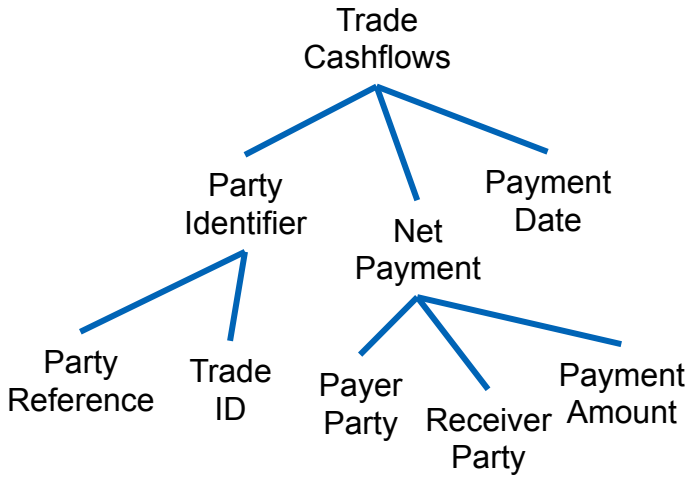


MarkLogic

Hierarchical Data Model

- MarkLogic Server is a document-centric database
  - Supports any-structured data via hierarchical data model
  - Stores compressed binary trees





MARKLOGIC IN ACTION

MarkLogic

What we will do

- Create a database
- Loading data
- Modifying data
- Searching and querying
- Fun with Indexes
- Building an application

LET’S START!

MarkLogic

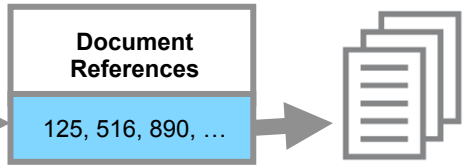
SOME WORDS ABOUT INDEXES

MarkLogic

Universal Index

*Which vetted SAR reports contain the phrase **blue van**?*

Term	Term List
"blue"	123, 125, 129, 130, 152, 344, …
"van"	123, 125, 126, 129, 130, 152, …
"observed"	125, 152, 516, 522, 765, 890, …
"blue van"	123, 125, 129, 130, 152, 486, …
STEM "observe"	125, 152, 516, 522, 765, 890, …
<SAR>	…
<SAR>/<location>	…
<threat>/<category>	…
<type>suspicious activity</type>	…
<date>2012-11-12Z</date>	…
Collection:Vetted	…
Role:Analyst + Action:Read	…
…	…
…	…
…	…



MarkLogic indexes…

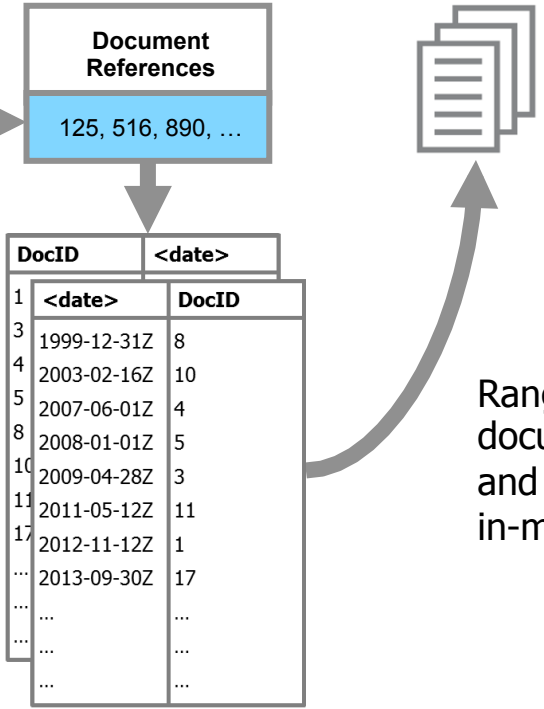
- Words
- Phrases
- Stemmed words and phrases
- Structure
- Words and phrases in the context of structure
- Values
- Collections
- Security Permissions

MarkLogic

Range Index

*Which vetted SAR reports containing the phrase **blue van** were submitted before 2013?*

Term	Term List
"blue"	123, 125, 129, 130, 152, 344, …
"van"	123, 125, 126, 129, 130, 152, …
"observed"	125, 152, 516, 522, 765, 890, …
"blue van"	123, 125, 129, 130, 152, 486, …
STEM "observe"	125, 152, 516, 522, 765, 890, …
<SAR>	…
<SAR>/<location>	…
<threat>/<category>	…
<type>suspicious activity</type>	…
<date>2012-11-12Z</date>	…
Collection:vetted	…
Role:Analyst + Action:Read	…
…	…
…	…
…	…



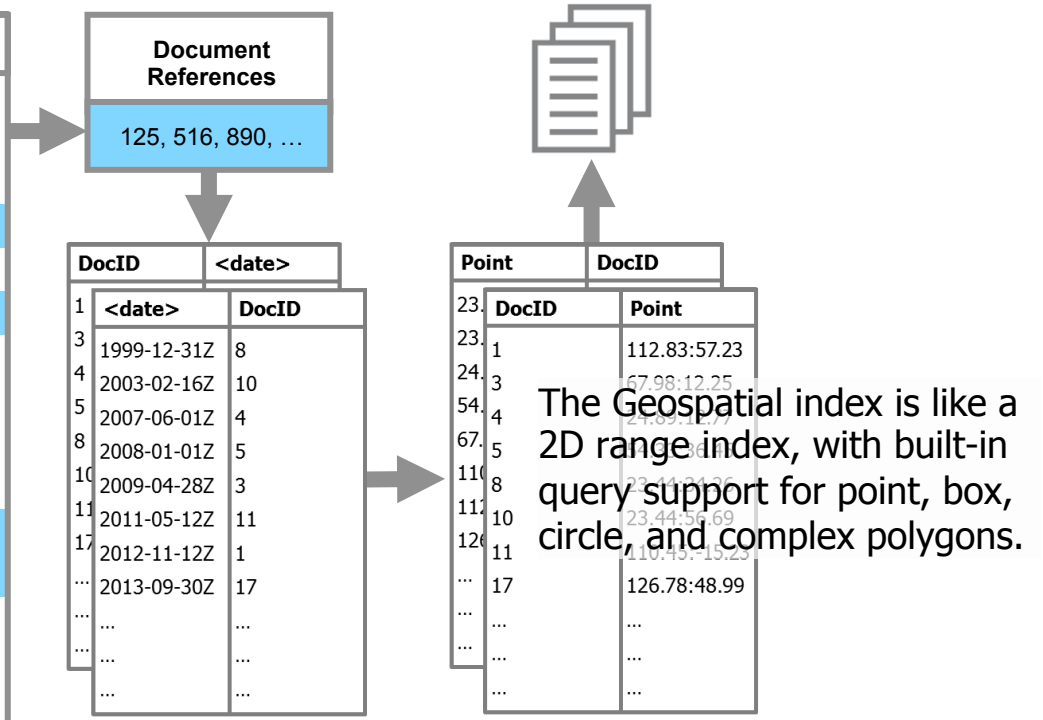
Range indexes map document IDs to values, and vice-versa in a compact in-memory representation.

MarkLogic

Geospatial Index

Which vetted SAR reports about a blue van from before 2013 refer to a location near the airport?

Term	Term List
"blue"	123, 125, 129, 130, 152, 344, …
"van"	123, 125, 126, 129, 130, 152, …
"observed"	125, 152, 516, 522, 765, 890, …
"blue van"	123, 125, 129, 130, 152, 486, …
STEM "observe"	125, 152, 516, 522, 765, 890, …
<SAR>	…
<SAR>/<location>	…
<threat>/<category>	…
<type>suspicious activity</type>	…
<date>2012-11-12Z</date>	…
Collection:Vetted	…
Role:Analyst + Action:Read	…
…	…
…	…
…	…

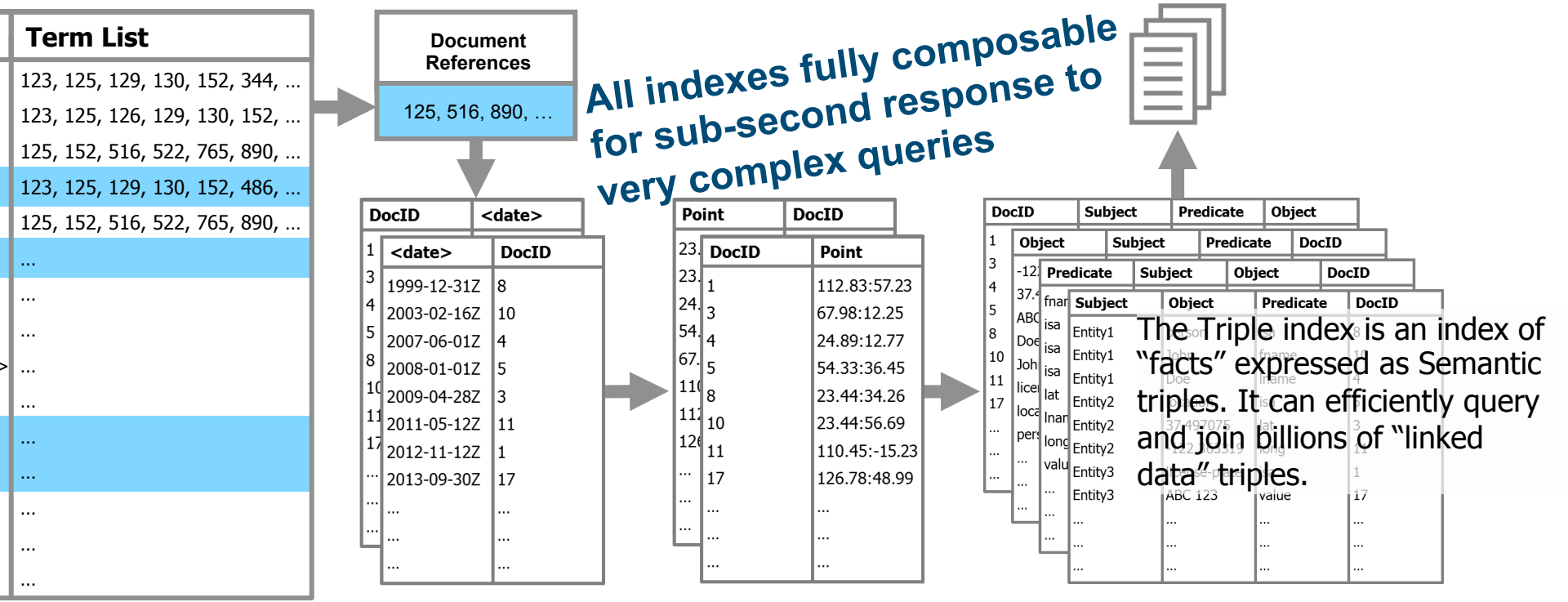


The Geospatial index is like a 2D range index, with built-in query support for point, box, circle, and complex polygons.

MarkLogic

Triple Index

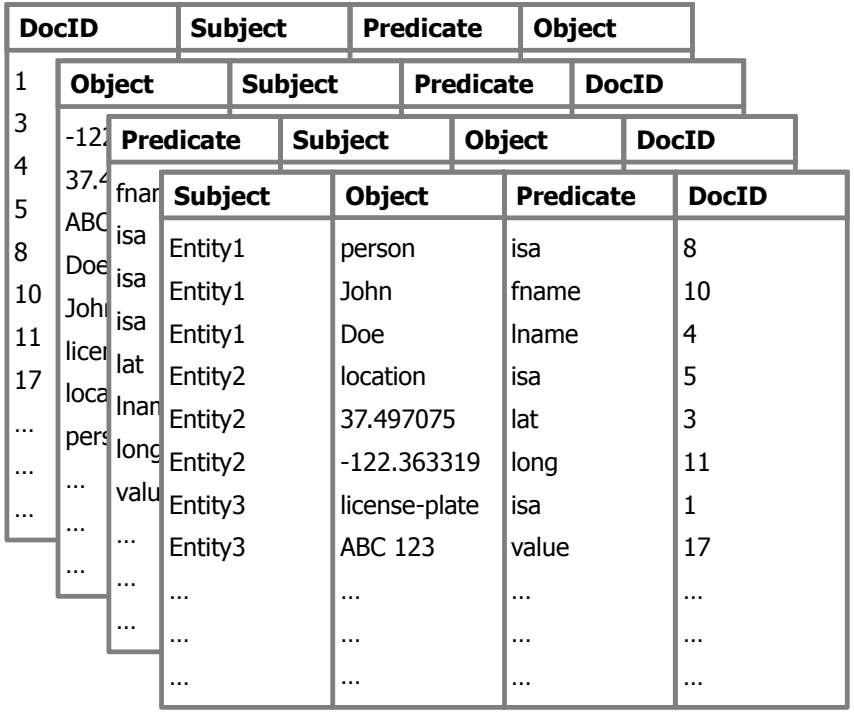
*Which vetted SARs about a blue van from before 2013 with this location refer to partial plate **ABC**?*



The Triple index is an index of "facts" expressed as Semantic triples. It can efficiently query and join billions of "linked data" triples.

MarkLogic

Triple Index



With the new Triple index, you can:

- Ingest triples from public linked data sources
- Combine triples in your documents with triples from other sources to answer questions like
  - What drivers' licenses (and names, addresses, etc) are linked to this license plate?
  - Show me everything you know about this license plate (traffic cam info, registration info, etc)

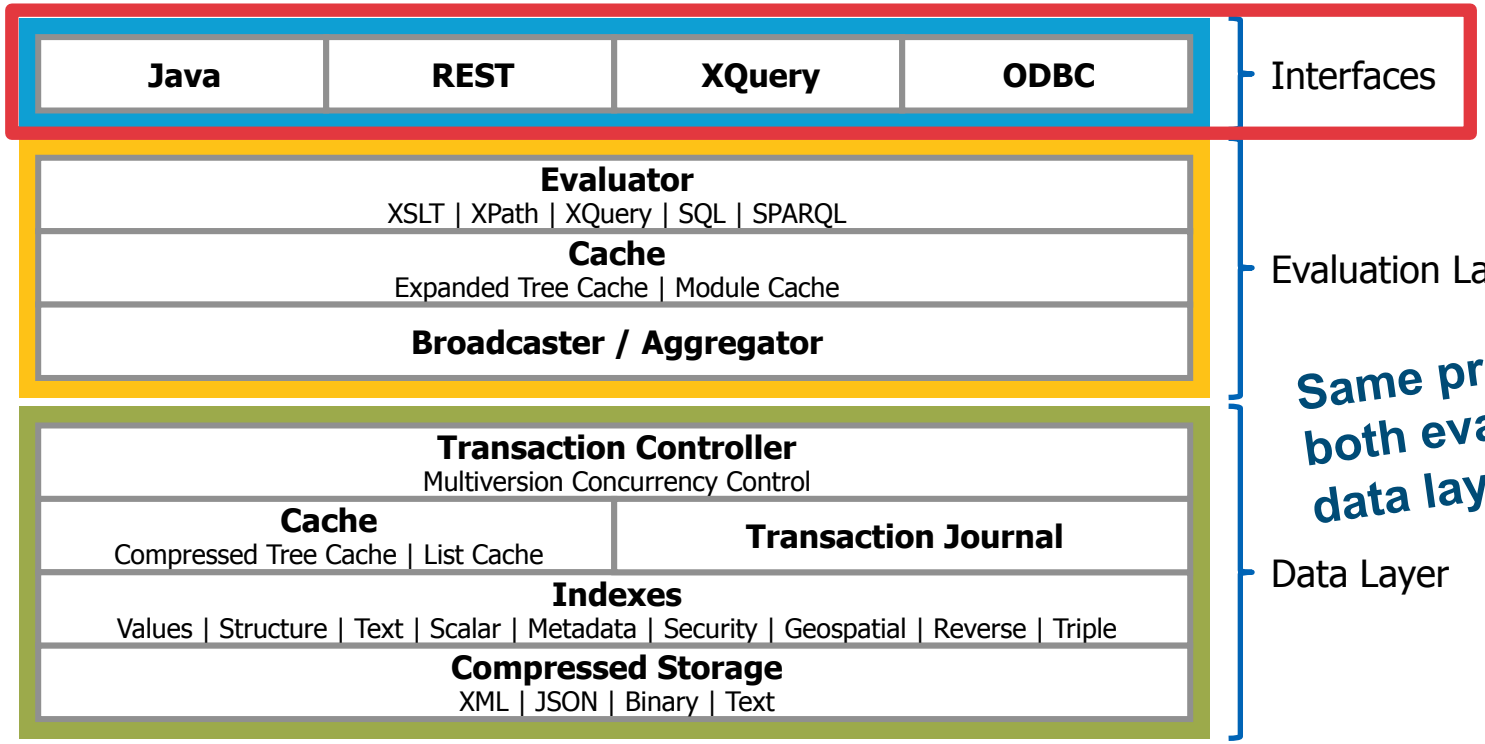
LET’S CONTINUE!

MarkLogic

DEVELOPMENT

MarkLogic

Remember this?



MarkLogic

REST

MarkLogic is REST Server

REST Interface to MarkLogic Server

Foundation for language-specific API's

- Full-text and facet search
- Access to value indexes and aggregates
- Document management for XML, text, binary, and JSON
- Separated read, write and administrative access
- Administrative services for managing API configuration
- Extensibility with custom URL endpoints
- Scriptable and UI bootsraping

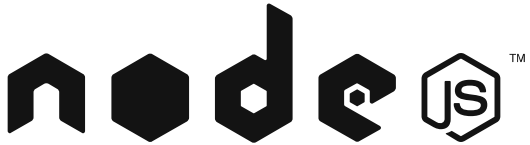
MarkLogic

Node.js Client API

New in version 8

A fluent JavaScript interface for key MarkLogic capabilities such as search, transactions, aggregates, alerting, and geospatial—the Node.js Client API allows developers to move faster to production with proven reliability at scale.

The Node client is developed in the open on GitHub, https://github.com/marklogic/node-client-api, and is licensed under the open-source Apache 2.0 license.



```
var marklogic = require('marklogic');
var conn = require('./env.js').connection; // Host and auth details

var db = marklogic.createDatabaseClient(conn);
var q = marklogic.queryBuilder;
db.documents.query(
  q.where(
    q.collection('countries'),
    q.value('region', 'Africa'),
    q.or(
      q.word('background',   'France'),
      q.word('Legal system', 'French')
    )
  )
)
  .result(function(documents) {
    documents.forEach(function(document) {
      console.log(JSON.stringify(document))
    })
  });
```

MarkLogic

Java Client API

New in version 8

The MarkLogic Java Client API is an open-source library that allows developers to quickly, easily, and reliably access MarkLogic from their Java applications.

- Faster development and less custom code with out-of-the-box data management, search, and alerting
- Pure Java query builder and conveniences for POJOs, JSON, XML, and binary I/O
- Built-in extensibility for moving performance-critical code to the database
- Always open-source and developed on GitHub, https://github.com/marklogic/java-client-api

```
package com.marklogic.examples;

import com.marklogic.client.DatabaseClient;
import com.marklogic.client.DatabaseClientFactory;

public class Configuration {
  private static DatabaseClient client =
DatabaseClientFactory.newClient(
      "localhost", 8000, // Every instance comes with a REST Client API
instance
                         // pre-installed on port 8000
      "Documents", // Each connection can specify its database at runtime
      "admin", "********", DatabaseClientFactory.Authentication.DIGEST);

  public static DatabaseClient exampleClient() {
    return client;
  }
}
```

BUILDING AN APPLICATION!

MarkLogic

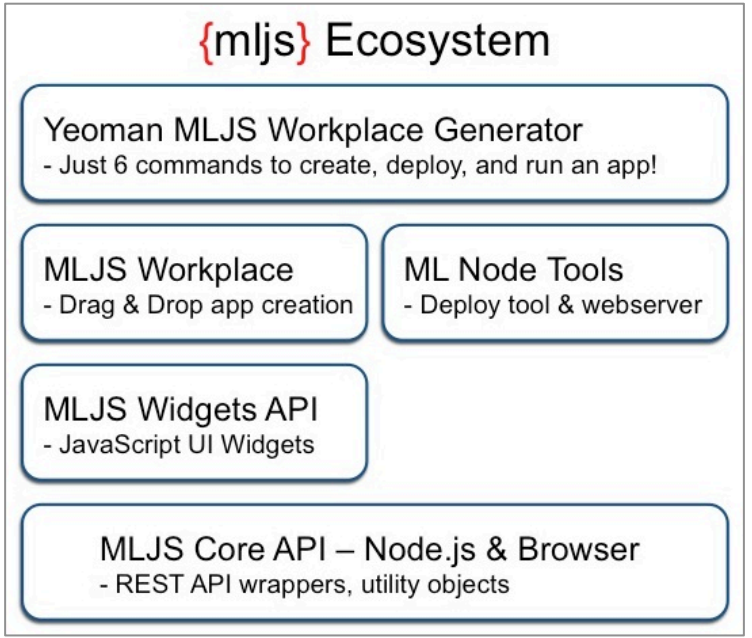
MLJS

Developed by an MarkLogic employee

MLJS is a MarkLogic JavaScript wrapper API. It comprises several components:-

- Core API that wraps the MarkLogic REST API and provides utility objects. Works in Node.js and the Browser
- Browser Widgets API – provides over 30 JavaScript widgets linked to MarkLogic
- MLJS Workplace – A drag/drop application that enables visual creation and configuration of a MarkLogic web application

More information at
https://adamfowlerml.wordpress.com/mljs/



MarkLogic

Getting started!

- Download your free developer license
  - http://developer.marklogic.com/products
  - Or test the latest release (early access) http://ea.marklogic.com
- Learn
  - By tutorials/examples http://developer.marklogic.com/learn
  - By free instructor lead online training http://mlu.marklogic.com/registration/
- Get involved
  - http://stackoverflow.com/questions/tagged/marklogic
  - MarkLogic Mailing list http://developer.marklogic.com/discuss

Contact me on mats.stellwall@marklogic.com
Follow me on twitter @MatsStellwall

ANY QUESTIONS?