

Tamara Nicole Rodríguez Luna.

Jeffrey Daniel Leiva Cascante.

Jonathan Andrey Porras Sandí.

Javier Díaz Jiménez.

Instituto Tecnológico de Costa Rica

IC-4302 Databases II

Investigation about ArangoDB.

27th April 2023.

Teacher: Erick Hernández Bonilla.

Table of contents

Cover	1
Table of contents.....	2
Index of tables.....	2
Index of Figures.....	2
Research objectives	4
Background	5
Detailed description.....	6
Architecture and Distribution	7
Consistency.....	7
Construction	8
Security	8
Use cases	9
Current Development Experiences	11
Conclusions.....	12
Prototype.....	13
Conclusions of the Prototype.....	14
References.....	15

Index of Figures

Figure 1 – Scalable Graph Technology.	5
Figure 2 – ArangoDB Components.....	7

a. Research objectives:

1. Analyze the behavior and documentation of the ArangoDB database.
2. Carry out and propose a strategy to put into practice good programming techniques, self-taught and teamwork.
3. Establish, specify, and generate a prototype that implements this database.
4. Select and distinguish the most important data in the history of the database for a correct implementation of the same.

b. Background:

“ArangoDB was first released in the year 2011 as AvocadoDB and later changed its name to ArangoDB in 2012, developed by ArangoDB GmbH. (Celler, 2012)

He came up with the ArangoDB 3 test stable version. It was used by Docker, DC/OS, Liaison.” (Kumar, 2022)

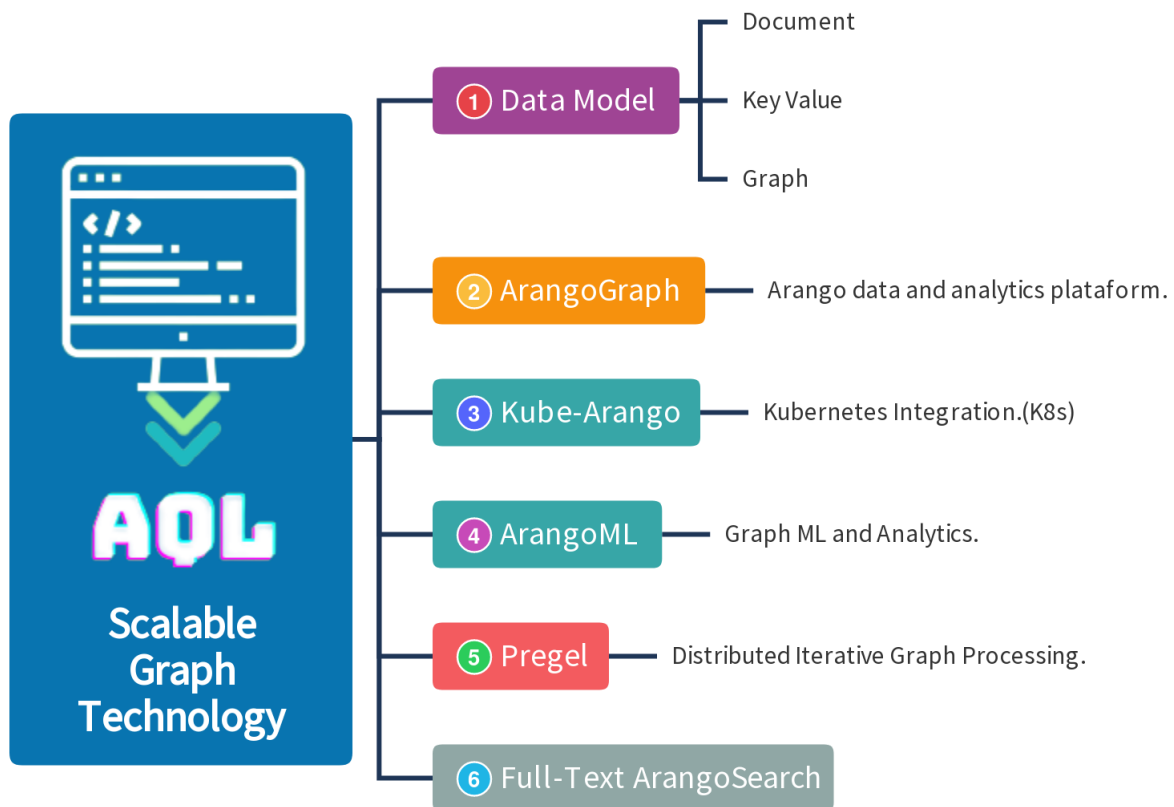
The datababase was founded in 2015 in Germany, ArangoDB is a company behind ArangoGraph Insights Platform: that is a next-generation graph data and analytics platform that accelerates application innovation and performance to drive faster value creation. (ArangoDB, s.f.a)

c. Detailed description:

Currently the database is in its ArangoDB 3.6 version, which highlights an optimization of the performance of subqueries, as well as UPDATE and REPLACE operations. (Darkcritz, 2020)

Users can store, retrieve, and manage data in a flexible, document-oriented, graph, and key-value format using the ArangoDB NoSQL database, which has a multi-model database management system. It has extensive capabilities like indexing, querying, and transactions for effective data management and is built to handle large-scale data storage and retrieval operations. Users can run complex operations and queries on their data using the ArangoDB Query Language (AQL), which is supported by ArangoDB. It is frequently utilized in many different application settings, such as IoT (Internet of Things) devices, analytics, and web and mobile apps. (Asensio, 2017)

Figure 1 – Scalable Graph Technology.



(ArangoDB, s.f.b)

Reference source: own elaboration.

Declarative consultative language enables the fusion of several data access consumers into a single consult. The data base may be managed using either the ArangoSH console client or the web-based user interface. The Apache 2 license governs the distribution of ArangoDB's source code. The project was created using JavaScript and C. (Darkcrist, 2020)

This method, which is very similar to the one employed by, for example, Couchbase and which produces such fantastic results, makes use of the need that each document has a unique identifier to store the remaining information as a value associated with that identification, much like the documental model does. (Asensio, 2017)

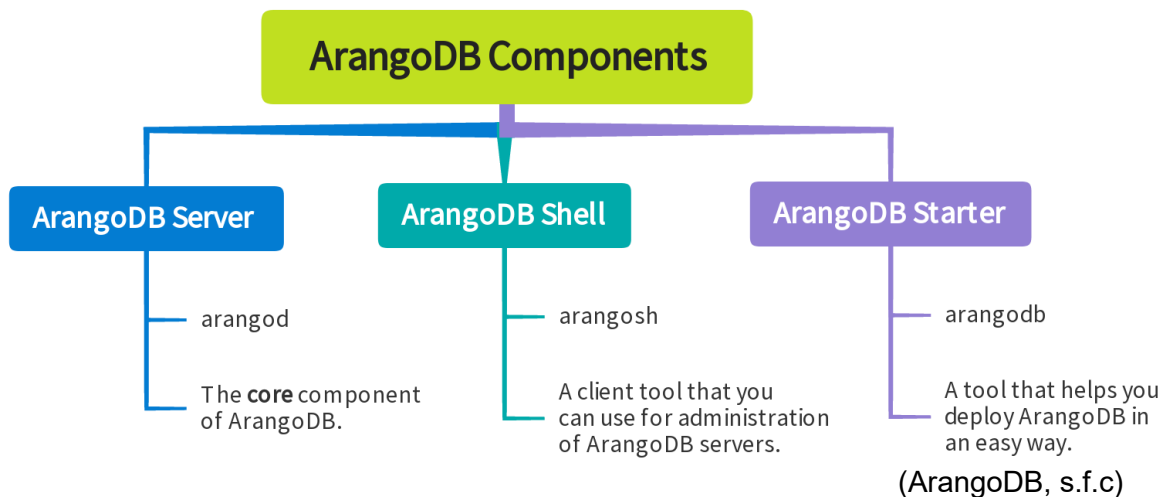
d. Architecture and distribution:

1. Available deployment modes:

You may install it as a single server, an optional robust pair with asynchronous replication and automatic failover, or a cluster made up of many nodes with synchronous replication and failover on-premises for high availability and resilience. For the highest level of data protection, you may arrange off-site replication for your whole cluster. (ArangoDB, s.f.c)

2. ArangoDB Components:

Figure 2 – ArangoDB Components.



Reference source: own elaboration.

3. Deploying by technology:

An environment can be implemented using a variety of techniques. You may manually launch all necessary procedures on your machine or in Docker containers. Alternately, use ArangoDB Starter, the arangodb binary, for local installs that make use of Docker processes or containers. To deploy to your Kubernetes cluster, use the ArangoDB Kubernetes operator (kube-arangodb).

The fastest way to start using ArangoDB is to run it on the cloud. A fully managed cloud service is offered by ArangoGraph Platform and is available on AWS, Microsoft Azure, and Google Cloud Platform. (ArangoDB, s.f.c)

e. Consistency:

Scaling ArangoDB:

He suggests a Master/Master CP (CAP Theorem) design with no single points of failure to provide horizontal scalability. This indicates that, like nearly other NoSQL systems, it prioritizes information consistency above availability when a network issue arises, and

that its architecture ensures that, in the case of a failure, the entire system does not crash. (Asensio, 2017)

f. Construction:

ArangoDB works on a distributed cluster and is certified for the Data Center Operating System (DC/OS). ArangoDB may be deployed using DC/OS in most current ecosystems, including Amazon Web Services (AWS), Google Compute Engine, and Microsoft Azure. (A, 2021)

g. Security:

According to asensio in 2017:

ArangoDB works on a distributed cluster and is certified for the Data Center Operating System (DC/OS). ArangoDB may be deployed using DC/OS in the majority of current ecosystems, including Amazon Web Services (AWS), Google Compute Engine, and Microsoft Azure. It is another one of those features that must be improved with each release. By including RocksDB, data can be stored encrypted on disk, taking one more step towards HIPAA compliance.

LDAP is integrated as an authentication mechanism, thus facilitating user management if we already have an active directory in our organization. (Asensio, 2017)

h. Use cases:

ArangoDB as a Graph Database:

1. Fraud Detection:

Find hard-to-detect patterns to uncover unlawful activity. ArangoDB enables you to see beyond individual data points in diverse data sources, allowing you to integrate and harmonize data to evaluate actions and relationships collectively. This gives you a more comprehensive understanding of connection patterns and helps you spot sophisticated fraudulent conduct, such as fraud rings.

(ArangoDB,s,f.d)

2. Recommendation Engine:

Provide users with recommendations for goods, services, and information based on data relationships. To create a movie recommendation system, for instance, you might combine ArangoDB and PyTorch Geometric. You would analyze the movies users have viewed, and then use a graph neural network (GNN) to predict links between the two.

(ArangoDB,s,f.d)

3. Network Management:

By integrating and visualizing the network, infrastructure, and code, downtime may be decreased. It makes sense to model network devices and their connections as graphs. When exploring routes between nodes, traversal algorithms give you the choice of stopping at subnet borders or accounting for factors like connection bandwidth when determining the best course of action.

(ArangoDB,s,f.d)

4. Customer 360:

Integrating several data sources and code can let you fully comprehend your consumers. With the additional capability of connecting related records and tracking data origins using graph features, ArangoDB may serve as the platform for merging and consolidating data in any shape. (ArangoDB,s,f.d)

5. Identity and Access Management:

By controlling data access based on function and position, you can increase security and compliance. An organizational chart may be represented as a graph, and ArangoDB can be used to establish who has access to what data. Use ArangoDB's graph capabilities to construct permission inheritance and access control lists. (ArangoDB,s,f.d)

6. Supply Chain:

By observing and improving the movement of commodities via a supply chain, you may hasten shipments. To identify the potential causes of snags and delays, you may see your inventory, supplier, and delivery information as a graph. (ArangoDB,s,f.d)

ArangoDB as a Document Database:

1. Content Management:

Store data of any kind without a prior specification of the schema. You can easily handle heterogeneous material with ArangoDB since it lacks a schema and stores each data entry as a separate document. Lay the foundation for the next (headless) content management system with ArangoDB. (ArangoDB,s,f.d)

2. E-Commerce Systems:

ArangoDB uses high consistency and resilience capabilities along with data modeling freedom to power online stores and ordering platforms. Process checkouts with the appropriate transactional assurances and handle product catalog data with ease using any combination of free text and structured data. (ArangoDB,s,f.d)

3. Internet of Things:

For a comprehensive perspective, gather sensor readings and other IoT data in ArangoDB. Store all data points in a single system that enables sliding window aggregation queries for effective data analysis. (ArangoDB,s,f.d)

ArangoDB as a Key-Value Database:

1. The most simple database systems are key-value stores. Each record is kept as a separate block of data and with a unique key. The system just saves the opaque data and may use the IDs to access it for you without knowing anything about the information it contains. Despite being able to store binary data, ArangoDB is best used with small to medium-sized JSON objects and is not intended to handle binary big items (BLOBs). (ArangoDB,s,f.d)

i. Current Development Experiences:

1. NOHO and ArangoDB:

“ArangoDB is very stable and by the great documentation and support of the community we got quickly up to speed.” – NOHO software.

(ArangoDB, d,f,e)

2. Juniper Networks:

“ArangoDB lets me enrich JSON documents over time as more network information is discovered.” - David Gee, Juniper Networks.

(ArangoDB, d,f,e)

3. Orange:

“ArangoDB’s support for a diverse, broad, and versatile data model is a key advantage to simplifying our complex architecture.” - Thomas Hassan, Orange

(ArangoDB, d,f,e)

4. Altana Technologies:

“We found ArangoGraph to be the best solution for combining geospatial, graph, and free-text search all into one coherent platform that enables us to empower the overall capability set that is Altana.” - Ian Cadieu, Altana

(ArangoDB, d,f,e)

5. Plural Technology:

“When our customers saw ArangoSearch, they were amazed and thrilled that they could retrieve correct data even if the spelling was wrong. Also, the response time was stunning.” - Deenu Gengiti, Plural Technology

(ArangoDB, d,f,e)

j. Conclusions:

1. Investigating this database was a fun exercise; we've looked at the many facets of making, keeping, and using an ArangoDB. The success of any project, in our case a prototype, will be significantly determined by how effectively it is developed and maintained.
2. We have concluded that every firm trying to remain competitive in the digital era needs a well-planned and maintained database. For something that is regarded as Universal, ArangoDB is simply unique due to the efficiency it offers, its tools, and its components.
3. From our perspective, its service is good when compared to other NoSQL databases we have dealt with. Because small, medium, and big businesses utilize it in the workplace, it is crucial to study and put into practice because it greatly enhances our understanding of software development.

k. Prototype:

We created a system for medical appointments, where we have administrators and users in the system. The scope is for nationals of the country of Costa Rica, with a national identification card made up of 9 digits, that is, XXXXXXXXX where X is each digit that confirms their identification card without spaces or special characters. Part of it is that you can only schedule one appointment per day for any of the specialties, in other words, on the same day there can be 1 appointment in odontology, 1 in medicine and another in ophthalmology, but never more than one on the same day and specialty, appointments are originally available.

In the system we will find a LogIn that is for system administrators, where they can see the available appointments, the busy ones, schedule an appointment or delete it. The system can be empty if there is no scheduled appointment.

Repository link:

[Medikare](#)

I. Conclusions of the prototype:

1. The fact that the language used for its AQL queries was so like SQL made it much easier for us to understand, learn and implement it in the system.
2. Our prototype is limited in scope for Costa Rican nationals or naturalized people with a 9-digit identification card, it is very efficient when it comes to making an appointment and it is a medical system that works perfectly according to the functionalities implemented. It is very important that we implemented security in the part of the administrator where he could access everything in the system, in addition to allowing easy access for normal users to schedule an appointment in the specialty they want.

References

ArangoDB. (s,f.a). Accelerating innovation with next-generation graph.

<https://www.arangodb.com/about-arangodb-2/>

ArangoDB. (s,f.b). What is ArangoDB? <https://www.arangodb.com/docs/stable/index.html>

ArangoDB. (s,f.c). Architecture. <https://www.arangodb.com/docs/stable/architecture.html>

ArangoDB. (s,f.d). ArangoDB Use Cases. <https://www.arangodb.com/docs/stable/use-cases.html>

ArangoDB. (d,f.e). Case Studies. <https://www.arangodb.com/solutions/case-studies/>

Asensio, A. (2017, September 14). ¡Hola! ¿Conoces ArangoDB? Paradigma Digital.

<https://www.paradigmadigital.com/dev/hola-conoces-arangodb/>

A, D. (s.f). ArangoDB, un sistema de base de datos NoSQL gratuito. Ubunlog.

<https://ubunlog.com/arangodb-sistema-base-de-datos-nosql-gratuito/>

Celler, F. (2012). “AvocadoDB” becomes “ArangoDB.” ArangoDB.

<https://www.arangodb.com/2012/05/avacadodb-become-arangodb/>

Darkcrist. (2020, January 7). ArangoDB un sistema de base de datos multimodelo de código abierto. Linux Adictos. <https://www.linuxadictos.com/arangodb-un-sistema-de-base-de-datos-multimodelo-de-codigo-abierto.html>

Kumar, R. (2022, March 26). What is ArangoDB? DevOpsSchool.com.

<https://www.devopsschool.com/blog/what-is-arangodb/>