

Developers

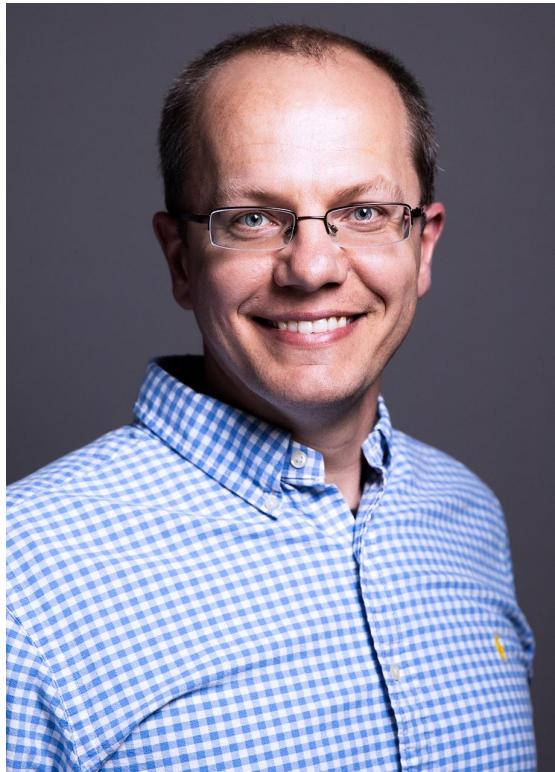
Introduction to NoSQL Databases



Developer Advocate



@ArtemChebotko



- Data professional, computer scientist
- Data modeling, data quality, data warehousing, data analytics
- Author of the Cassandra Data Modeling Methodology
- Google Cloud Certified Data Engineer



Artem Chebotko

Developer Advocate



@RyanWelford



- Apache Cassandra™ expert
- Front End Developer



Ryan Welford



01

Workshop Organization



02

Introduction to NoSQL

03

Document Databases

04

Key-Value Databases

05

Tabular Databases

06

Graph Databases

07

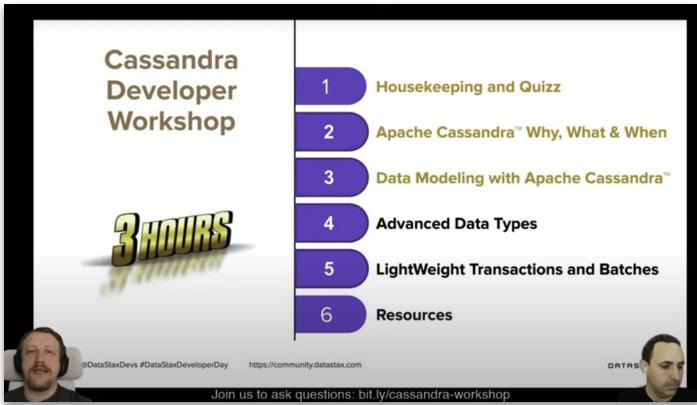
What's next?



Agenda

Live and interactive

Livestream: youtube.com/DataStaxDevs

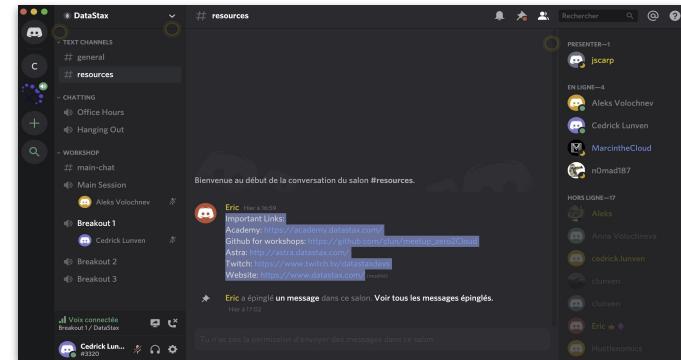


YouTube



Twitch

Questions: <https://dtsx.io/discord>



Discord

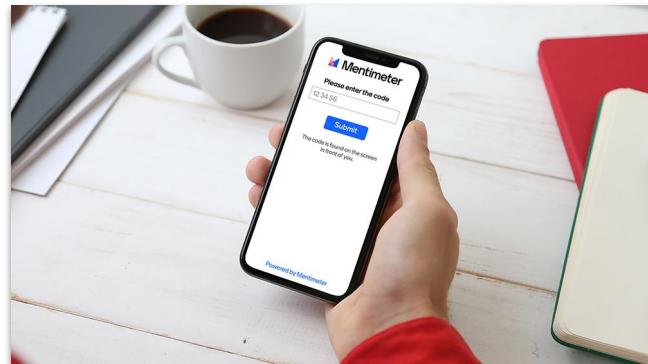


YouTube



Available on the iPhone App Store

Games menti.com



Mentimeter

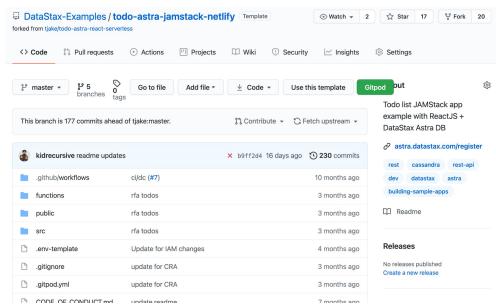


GET IT ON
Google play

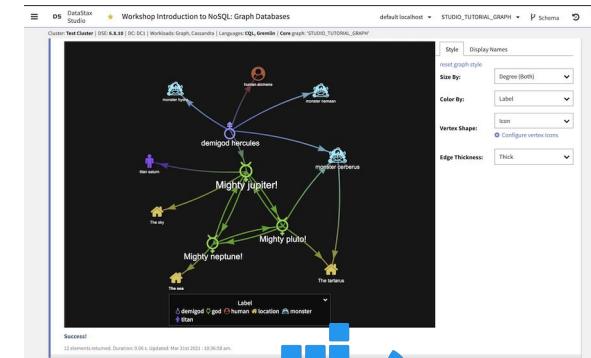
Attend the live sessions

Nothing to install !

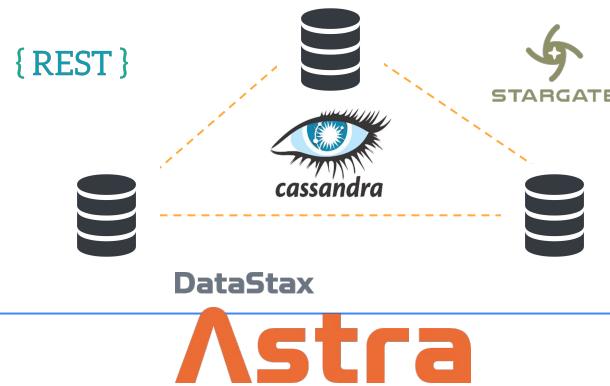
Source code + exercises + slides



DataStax Studio



Database + CQL + APIs



Do the hands-on

550+ and counting

Intro to NoSQL Homework

cedrick.lunven@datastax.com [Switch account](#)

 Draft restored

The name and photo associated with your Google account will be recorded when you upload files and submit this form. Only the email you enter is part of your response.

* Required

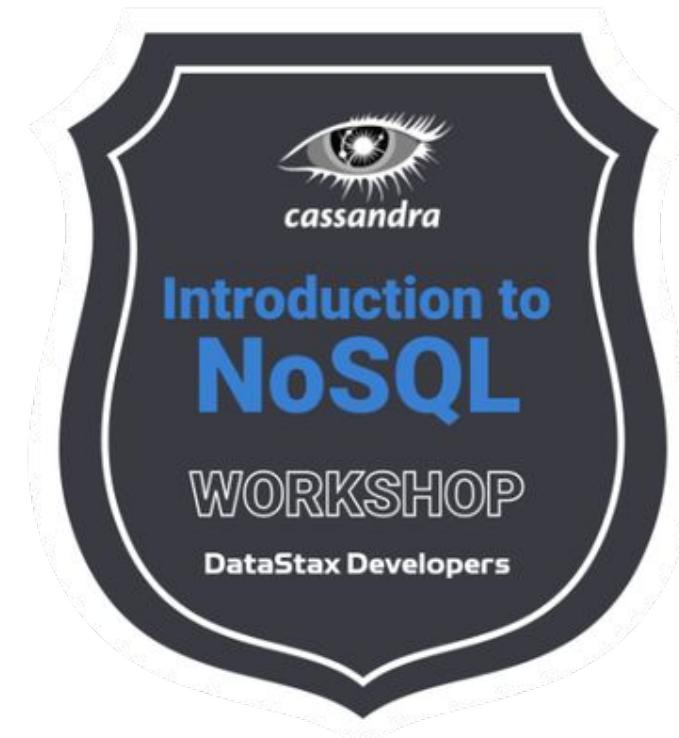
Email *

cedrick.lunven@datastax.com

Full Name *

Your full name (to be displayed on the badge)

Cedrick Lunven

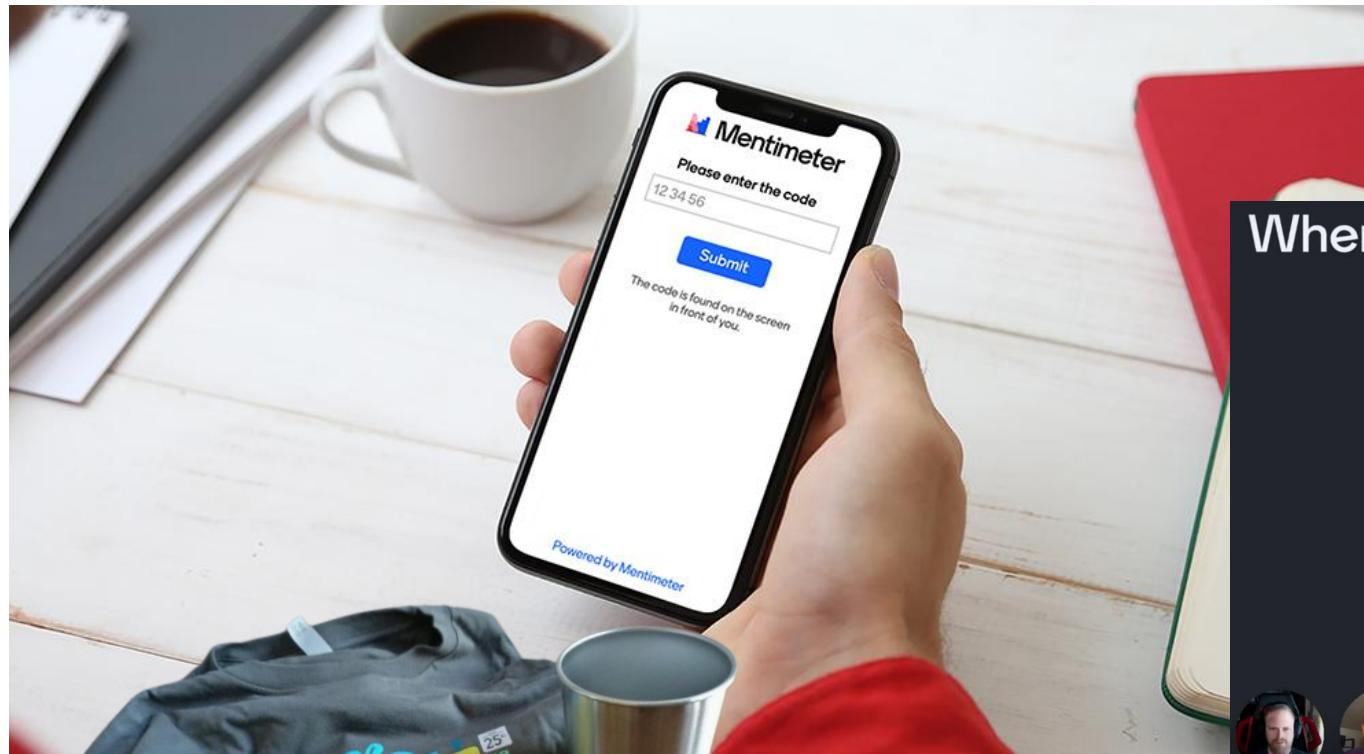


Do homework and earn a badge

datastax.com/workshops



Become a Jedi Master of Astra



Where are you from?

Mentimeter



menti.com ⇒ enter code
Don't answer in YT chat
Look at phone (not at YT)
Keep it open for later

Menti.com for surveys and quizzes

01

Workshop Organization

02

Introduction to NoSQL



04

Document Databases

05

Key-Value Databases

03

Tabular Databases

06

Graph Databases

07

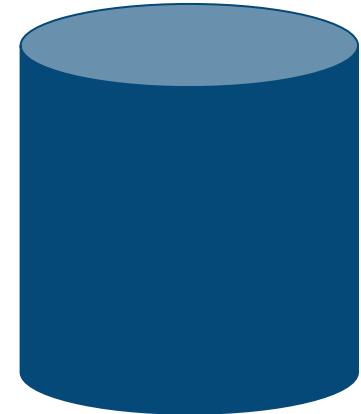
What's next?



Agenda

What is a Database?

- An organized collection of related data items
- Database Management System (DBMS)



Databases by Data Model

- Navigational, hierarchical, network databases
- Relational databases
 - Row-oriented for online transaction processing or OLTP
 - Column-oriented or columnar for online analytical processing or OLAP
- Object databases
- ...
- NoSQL databases
- ...
- Multi-model databases



Tabular or wide-column



Document



Key-value



Graph

Databases by Deployment Model

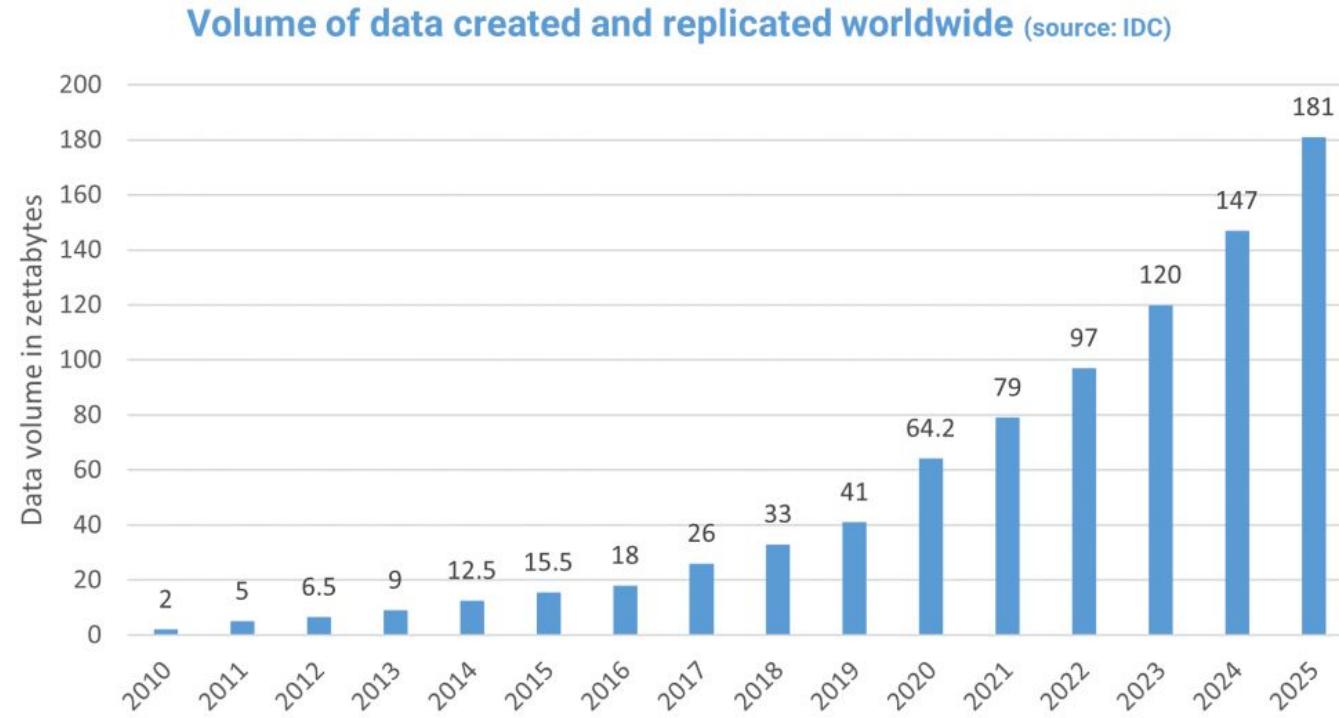
- Single-machine database
- Cluster, distributed database
- On-prem databases
- Cloud databases
 - Virtual machines
 - Database-as-a service or DBaaS
 - Managed service
 - Serverless, auto-scalable

Why Not Relational?

- Relational database pros
 - Standardization and theoretical foundation
 - relational data model
 - relational algebra and calculus
 - normalization theory
 - SQL
 - ACID transactions
 - Integration databases shared by many apps
- Relational database cons
 - Designed to run on a single machine and scale vertically
 - Impedance mismatch between the relational data model and app in-memory data structures

Why NoSQL?

- New high-growth challenges
 - Big data and 3Vs: volume, velocity, variety
 - Performance and scalability
 - Clusters



Origins of term “NoSQL”

- Meetup name on June 11, 2009 in San Francisco
 - Catchy hashtag intended to refer to databases like BigTable and DynamoDB
 - Meetup presentations: Cassandra, MongoDB, CouchDB, HBase, Voldemort, Dynomite, and Hypertable
- Sometimes we can find the “Not only SQL” interpretation today

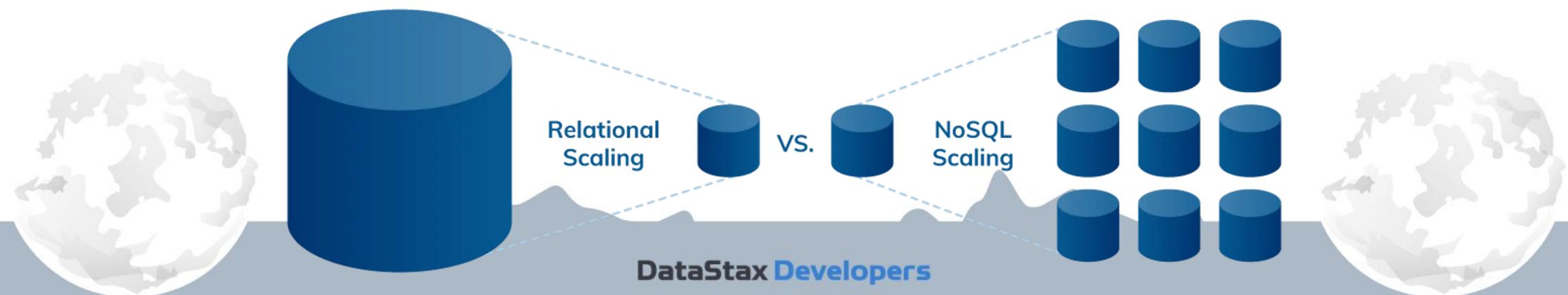


What is NoSQL?

- A database that can meet high-growth challenges
- Common characteristics
 - Non-relational data model
 - Runs on clusters
 - Globally, geographically distributed
 - High throughput, real-time transactions
 - Horizontal scalability, preferably linear scalability
 - Open source or based on open source
 - Bound by the CAP theorem

Relational vs. NoSQL

- Relational
 - Standard relational data model and language SQL
 - ACID transactions
 - Integration database
 - Designed for a single machine
 - Hard to scale
 - Impedance mismatch
- NoSQL
 - Variety of data models and languages
 - Lower-guarantee transactions
 - Application database
 - Designed for a cluster
 - Easy to scale
 - Better database-app compatibility



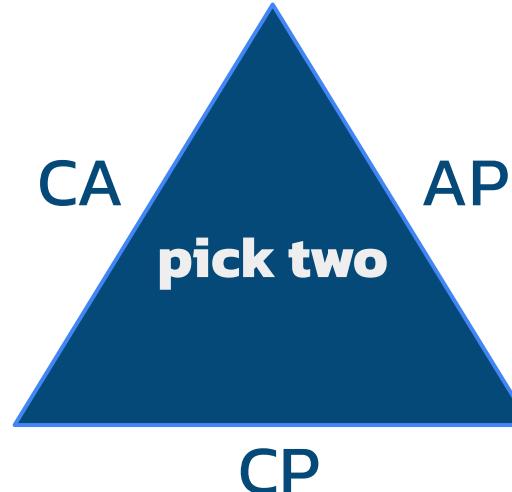
The CAP Theorem

Always responds,
may not always return
the most recent write

Availability

Consistency

Every read receives
the most recent write
or an error

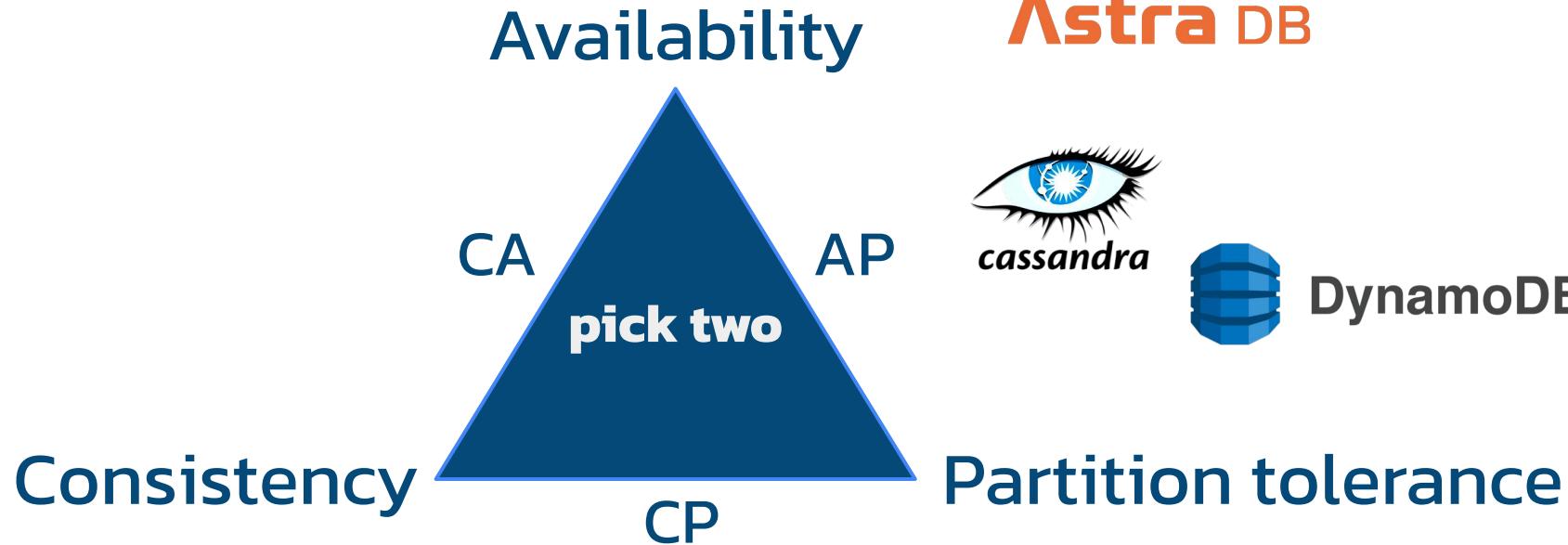


Partition tolerance

Operates in the
presence of network
partition failures



The CAP Theorem



Astra DB



cassandra



DynamoDB



Consistency

Partition tolerance



redis



BigTable



Apache HBase



Couchbase

DataStax Developers



State-of-the-art NoSQL Database YOU will be Using Today

Astra DB

- DBaaS, serverless, auto-scalable
- Multi-cloud, distributed, multi-node cluster
- NoSQL, multi-model
- Tabular, document, key-value
- Based on open-source *Apache cassandra*

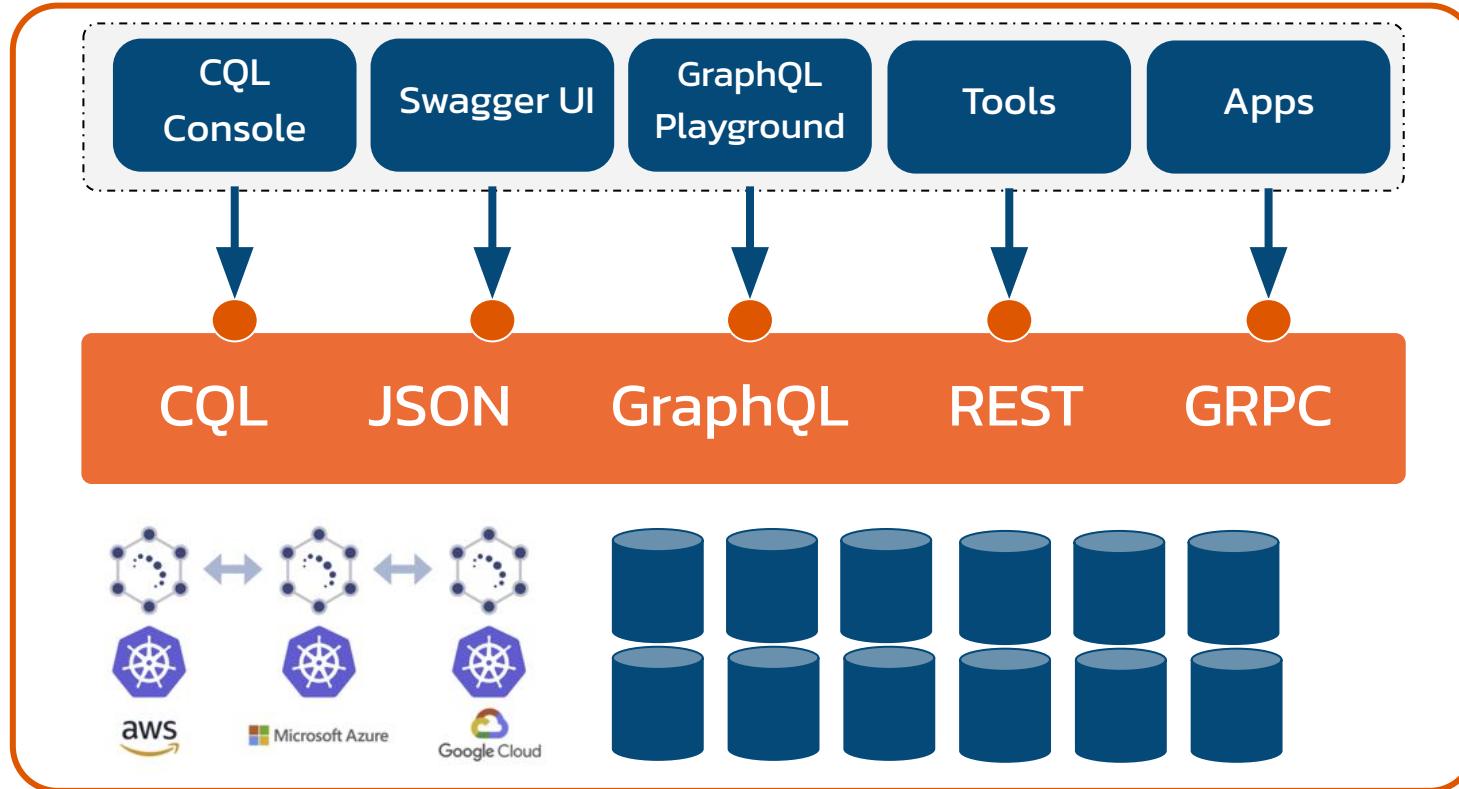


DataStax

Astra DB

\$25/month credit

Launch a database in the cloud with a few clicks, no credit card required.



User Interface
Web-based developer tools and apps

OSS Stargate.io
A data gateway to allow multiple usages

OSS Apache Cassandra
A tabular NoSQL database

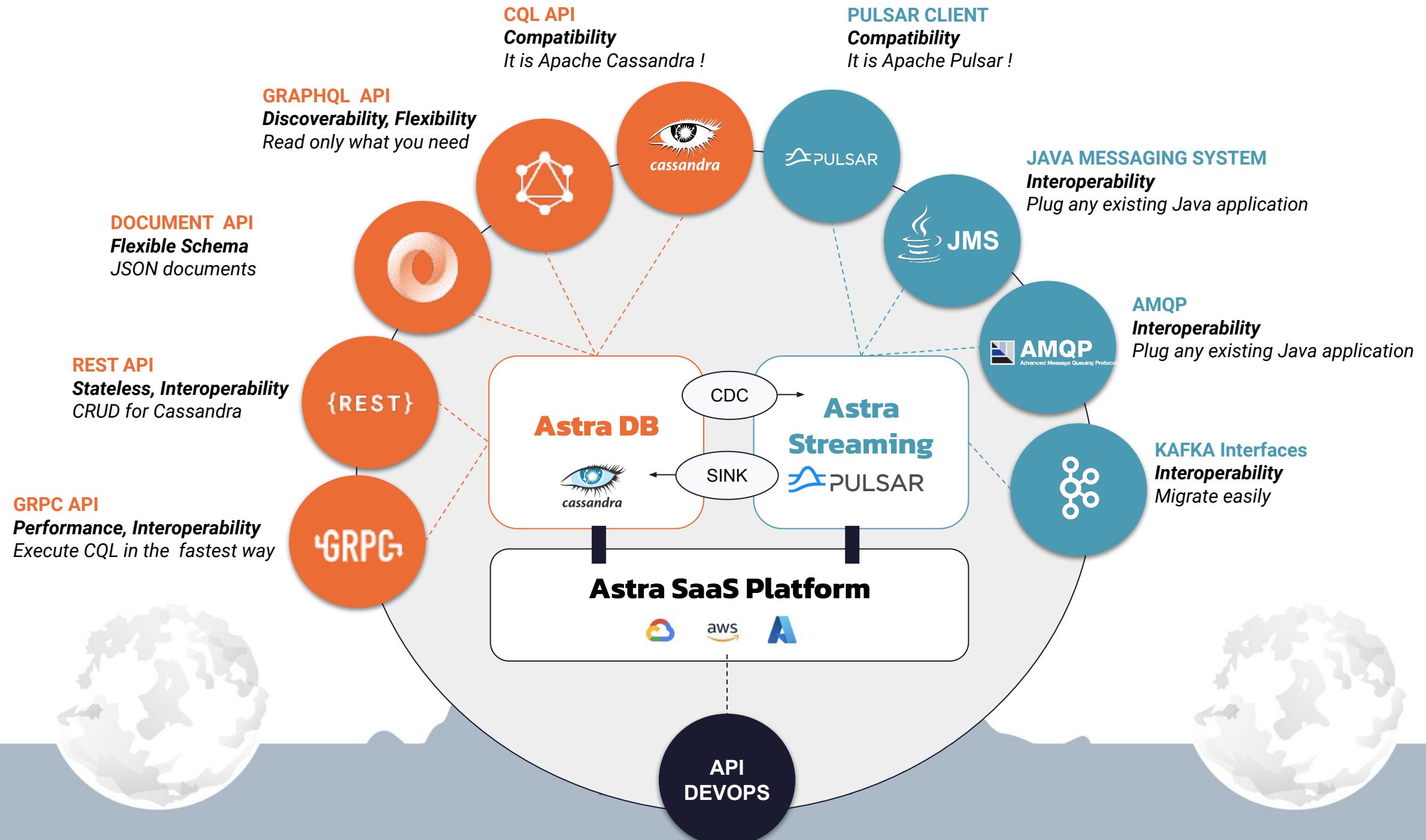
DataStax Developers

Hands-on (!github)

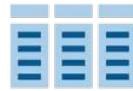
#1 Database Setup

github.com/datastaxdevs/workshop-introduction-to-nosql

- ✓ Create a new database
- ✓ Wake up an existing hibernated database



4 Types of NoSQL Databases



Tabular or wide-column database

Astra DB



Document database

Astra DB



Key-value database

Astra DB



Graph database



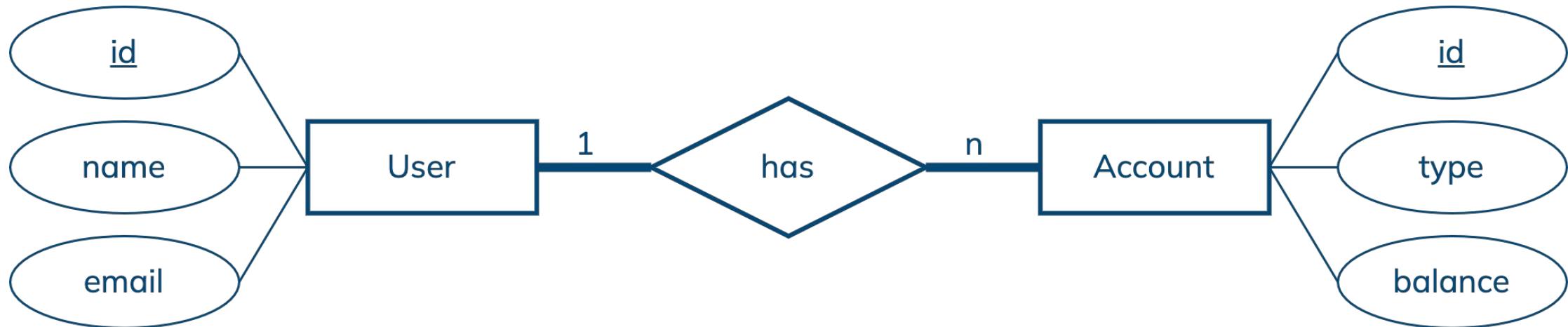
DataStax Graph



DataStax Developers



Running Example: Entity-Relationship Model

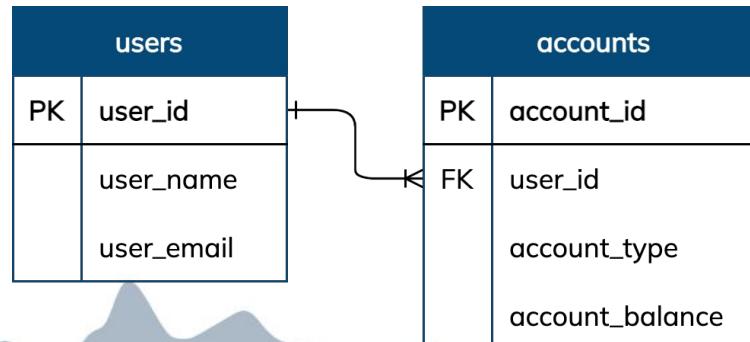


- Meet Alice  `alice@example.org` and Bob  `bob@example.org`
-  **has Checking and Savings accounts with balances 2500 and 1500**
-  **has one Checking account with balance 1000**

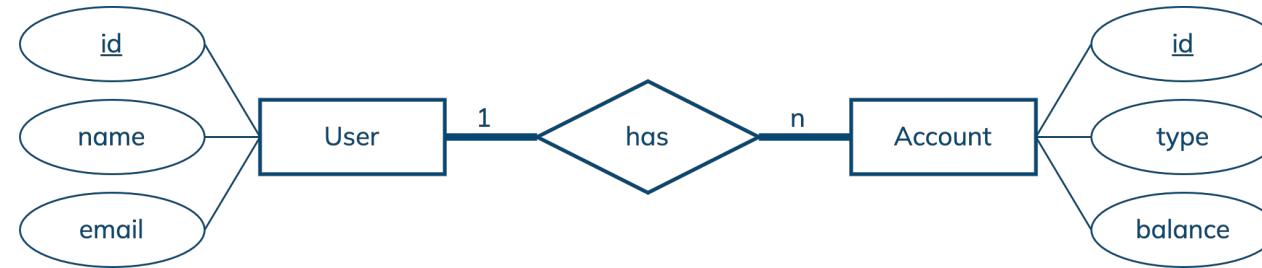
Running Example: Relational Data Model

users		
user_id	user_name	user_email
1cafb6a4-396c-4da1-8180-83531b6a41e3	Alice	alice@example.org
0d2b2319-9c0b-4ecb-8953-98687f6a99ce	Bob	bob@example.org

accounts			
account_id	account_type	account_balance	user_id
83428a85-5c8f-4398-8019-918d6e1d3a93	Checking	2500	1cafb6a4-396c-4da1-8180-83531b6a41e3
811b56c3-cead-40d9-9a3d-e230dc64f2f	Savings	1500	1cafb6a4-396c-4da1-8180-83531b6a41e3
81def5e2-84f4-4885-a920-1c14d2be3c20	Checking	1000	0d2b2319-9c0b-4ecb-8953-98687f6a99ce



Running Example: What about NoSQL?



?

?

?

?



01

Workshop Organization

02

Introduction to NoSQL

03

Tabular Databases



04

Document Databases

05

Key-Value Databases

06

Graph Databases

07

What's next?



Agenda

Tabular Databases



Astra DB



Data

- Tables with columns and rows; denormalized tables
- Primary, partition, clustering (sorting) keys
- Secondary indexes, materialized views

Use Cases

- General-purpose; many similarities to relational databases
- Transaction Processing, IoT, Time Series, Messaging, Activity Tracking, Content Management, eCommerce, Retail, Finance

Query

- Retrieving one, some or all rows from a partition
- SQL-like, no joins, indexed columns, lightweight transactions



Astra DB

```
SELECT account_type,  
       account_balance  
  FROM accounts_by_users  
 WHERE user_id = ?;
```

user_id	account_id	user_name	account_balance	account_type
0d2b2319-9c0b-4ecb-8953-98687f6a99ce	81def5e2-84f4-4885-a920-1c14d2be3c20	Bob	1000	Checking
1cafb6a4-396c-4da1-8180-83531b6a41e3	811b56c3-cead-40d9-9a3d-e230dc64f2f	Alice	1500	Savings
1cafb6a4-396c-4da1-8180-83531b6a41e3	83428a85-5c8f-4398-8019-918d6e1d3a93	Alice	2500	Checking

Running Example: Tabular / Wide-Column Data Model

accounts_by_user	
user_id	K
account_id	C↑
account_type	S
account_balance	S
use_name	S
use_email	S

accounts_by_user					
user_id	account_id	account_type	account_balance	user_name	user_email
 1cafb6a4-396c-4da1-8180-83531b6a41e3	811b56c3-cead-40d9-9a3d-e230dcd64f2f	Savings	1500	Alice	alice@example.org
	83428a85-5c8f-4398-8019-918d6e1d3a93	Checking	2500		
0d2b2319-9c0b-4ecb-8953-98687f6a99ce	81def5e2-84f4-4885-a920-1c14d2be3c20	Checking	1000	Bob	bob@example.org

Hands-on (!github)

#2 Tabular Databases

github.com/datastaxdevs/workshop-introduction-to-nosql



01

Workshop Organization

02

Introduction to NoSQL

03

Tabular Databases

04



Document Databases

05

Key-Value Databases

06

Graph Databases

07

What's next?

Document Databases



Data

- JSON, semi-structured documents
- Schemaless is a marketing term
- Schema-on-read, schema embedded into the document

Query

- Search using document IDs and JSON internal field values

Use Cases

- Flexible schema
- Excellent mapping between database and app data models
- Content Management, Catalogs, Mobile Apps

users

{ JSON_A }

{ JSON_B }

mongoDB db.users.find({ name: "Alice" })

Astra DB

GET namespaces/nosql1/collections/users?
where={"name": { "\$eq": "Alice" } }

Running Example: Document Data Model

```
{  
  "id": "0d2b2319-9c0b-4ecb-8953-98687f6a99ce",  
  "name": "Bob",  
  "email": "bob@example.org",  
  "accounts": [  
    {  
      "id": "81def5e2-84f4-4885-a920-1c14d2be3c20",  
      "type": "Checking",  
      "balance": "1000"  
    }  
  ]  
}
```



```
{  
  "id": "1cafb6a4-396c-4da1-8180-83531b6a41e3",  
  "name": "Alice",  
  "email": "alice@example.org",  
  "accounts": [  
    {  
      "id": "83428a85-5c8f-4398-8019-918d6e1d3a93",  
      "type": "Checking",  
      "balance": "2500"  
    },  
    {  
      "id": "811b56c3-cead-40d9-9a3d-e230dc64f2f",  
      "type": "Savings",  
      "balance": "1500"  
    }  
  ]  
}
```



Hands-on (!github)

#3 Document Databases



github.com/datastaxdevs/workshop-introduction-to-nosql

01

Workshop Organization

02

Introduction to NoSQL

03

Tabular Databases

04

Document Databases

05

Key-Value Databases



06

Graph Databases

07

What's next?

Key-Value Databases



DynamoDB

Astra DB

Data

- Key-value pairs
- Values can be blobs, strings, sets, maps, JSON docs, etc

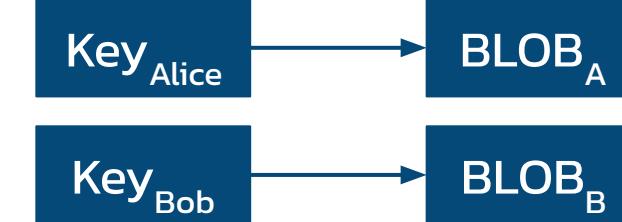
Query

- GET/PUT/DELETE key lookups, limited SQL-like support

Use Cases

- Best for simple primary key lookups
- Caching, User Sessions, Serving Content by Key

users



GET /buckets/users_kv/keys/Alice



Astra DB

SELECT value FROM users_kv
WHERE key = 'Alice';



Running Example: Key-Value Data Model

KV	
key	value
user:1cafb6a4-396c-4da1-8180-83531b6a41e3:name	Alice
user:1cafb6a4-396c-4da1-8180-83531b6a41e3:email	alice@example.org
user:1cafb6a4-396c-4da1-8180-83531b6a41e3:accounts	{ 83428a85-5c8f-4398-8019-918d6e1d3a93, 811b56c3-cead-40d9-9a3d-e230dcd64f2f }
user:0d2b2319-9c0b-4ecb-8953-98687f6a99ce:name	Bob
user:0d2b2319-9c0b-4ecb-8953-98687f6a99ce:email	bob@example.org
user:0d2b2319-9c0b-4ecb-8953-98687f6a99ce:accounts	{ 81def5e2-84f4-4885-a920-1c14d2be3c20 }
account:83428a85-5c8f-4398-8019-918d6e1d3a93:type	Checking
account:83428a85-5c8f-4398-8019-918d6e1d3a93:balance	2500
account:811b56c3-cead-40d9-9a3d-e230dcd64f2f:type	Savings
account:811b56c3-cead-40d9-9a3d-e230dcd64f2f:balance	1500
account:81def5e2-84f4-4885-a920-1c14d2be3c20:type	Checking
account:81def5e2-84f4-4885-a920-1c14d2be3c20:balance	1000

Hands-on (!github)

#4 Key-Value Databases



github.com/datastaxdevs/workshop-introduction-to-nosql

01

Workshop Organization

02

Introduction to NoSQL

03

Tabular Databases

04

Document Databases

05



Key-Value Databases

06

Graph Databases

07

What's next?

Graph Databases



DataStax Graph



neo4j



AllegroGraph

Data

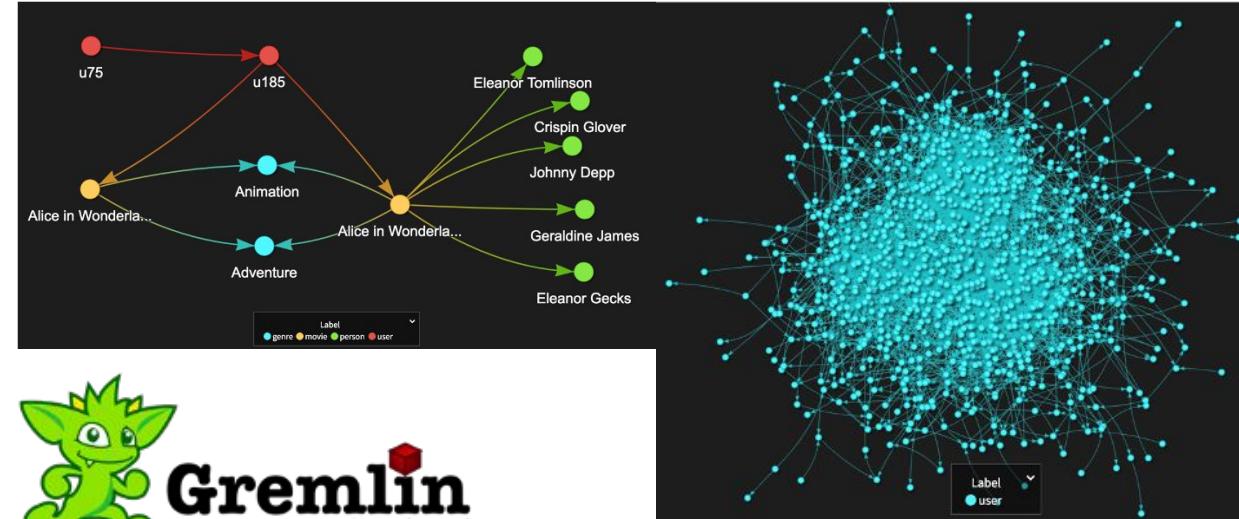
- Vertices, edges, properties
- Property graphs, RDF graphs, knowledge graphs

Query

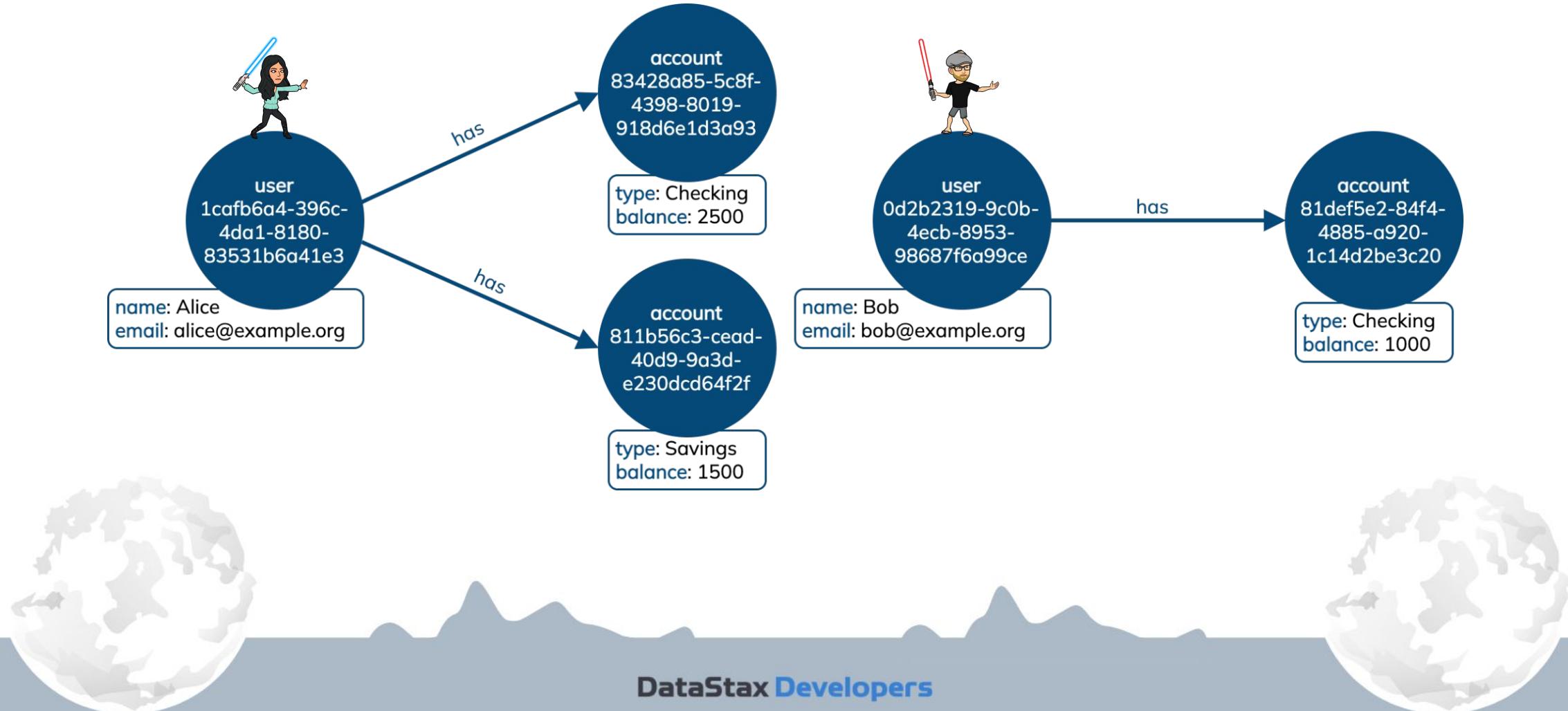
- Gremlin, Cypher, GSQL, SPARQL

Use Cases

- Focus on exploring connections, links and relationships
- Customer 360, Personalization, Fraud detection, Recommendations, Internet of Things, Asset management, Data integration



Running Example: Graph Data Model



Hands-on (!github)

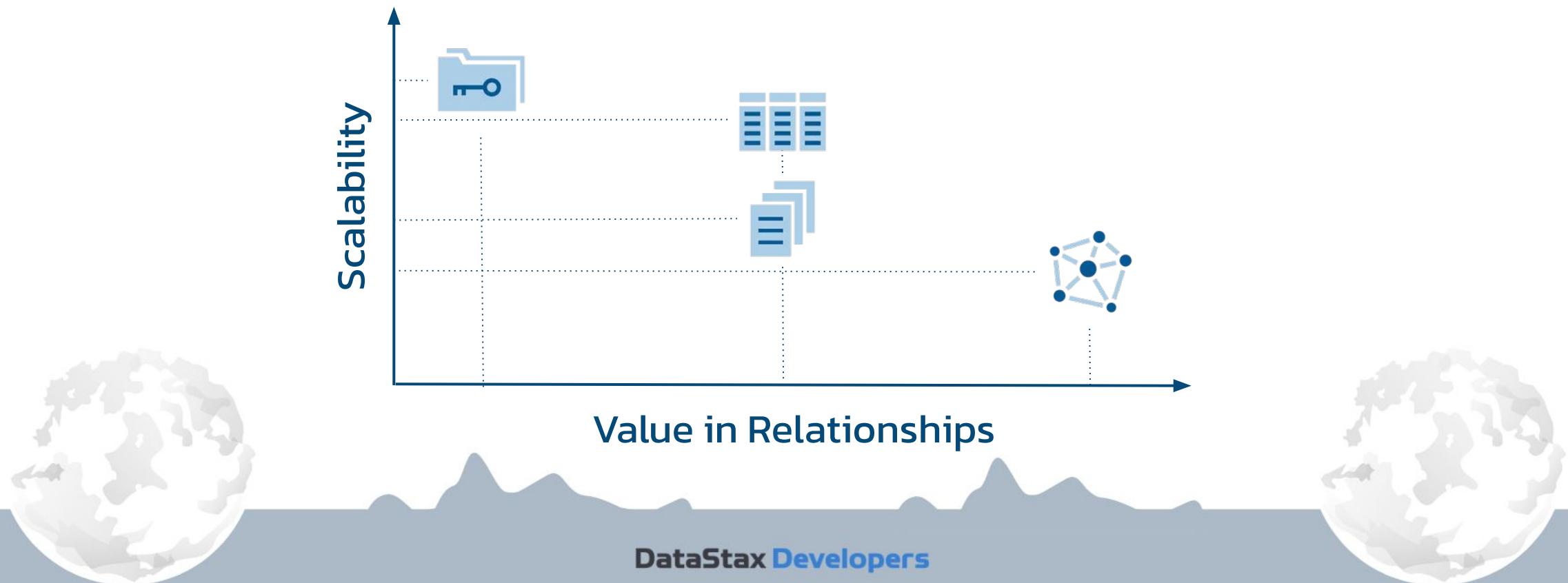
#5 Graph Databases

github.com/datastaxdevs/workshop-introduction-to-nosql



Should You Choose database?

- Consider a multi-model database, data model is only one aspect
- Use case, performance, scalability, availability, DBaaS, multi-cloud, ...



01

Workshop Organization

02

Introduction to NoSQL

03

Tabular Databases

04

Document Databases

05



Key-Value Databases

06

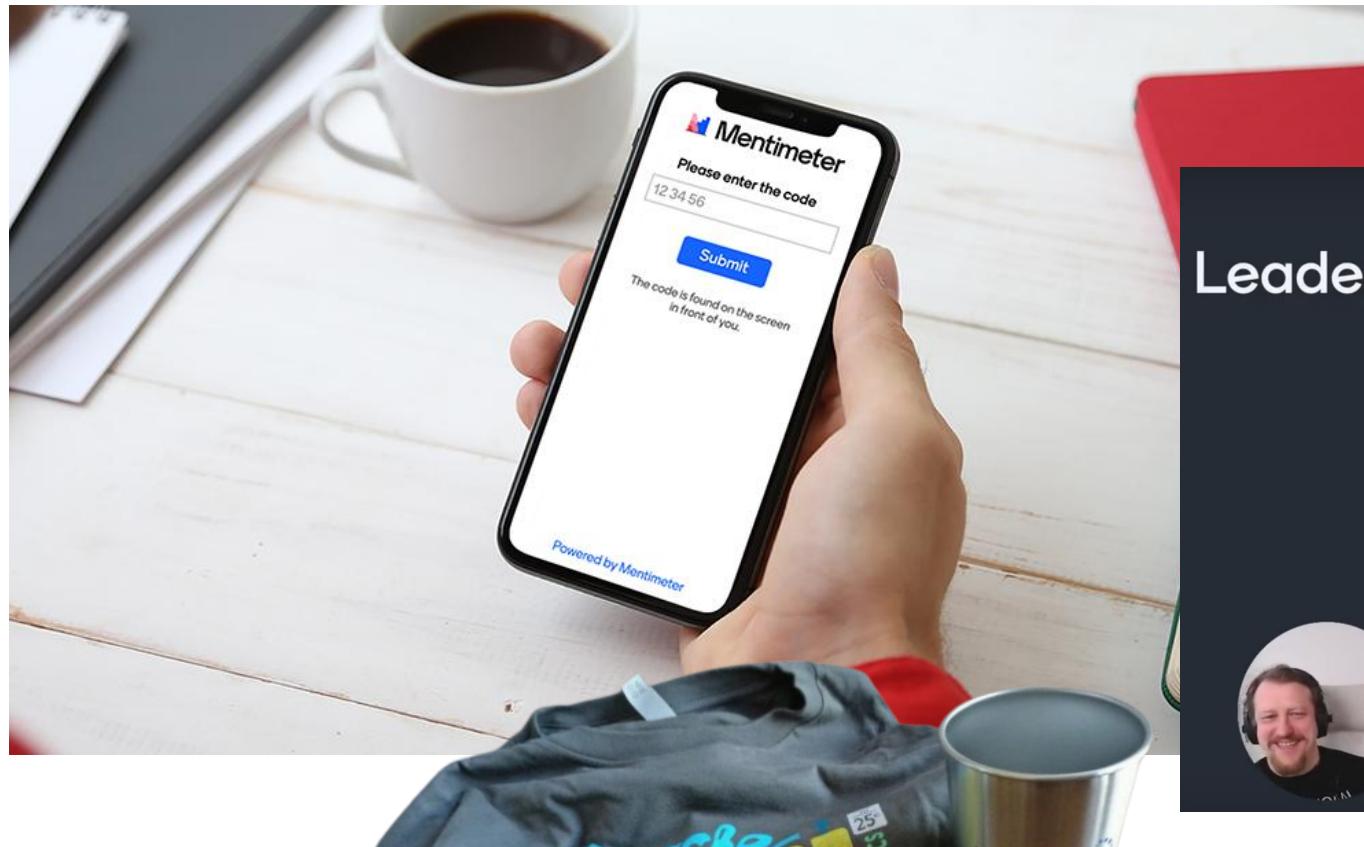
Graph Databases

07

What's next?



Agenda



menti . com ⇒ enter code
Don't answer in YT chat
Look at phone (not at YT)

Quiz on Menti.com !

SWAG WINNERS



Congratulations to 1st, 2nd and 3rd place on the Menti quiz!

To claim your prize, please send an email to:

gary.harvey@datastax.com

**** Include a screenshot of your Menti screen**



Swag Winners!



Intro to NoSQL Homework

cedrick.lunven@datastax.com [Switch account](#)

 Draft restored

The name and photo associated with your Google account will be recorded when you upload files and submit this form. Only the email you enter is part of your response.

* Required

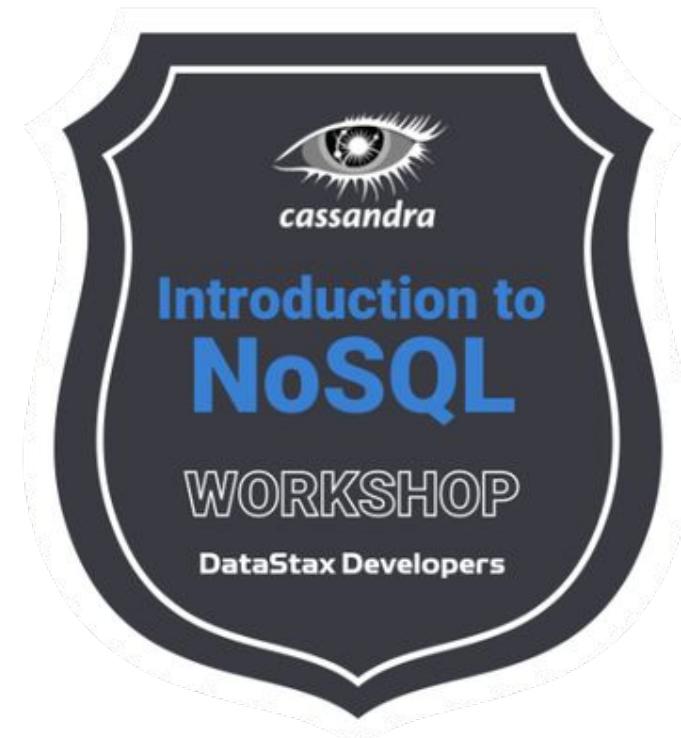
Email *

cedrick.lunven@datastax.com

Full Name *

Your full name (to be displayed on the badge)

Cedrick Lunven



Do homework and earn a badge

Homework (!github)

Instructions are on GitHub

github.com/datastaxdevs/workshop-introduction-to-nosql

- ✓ Complete hands-on exercises
- ✓ Submit a form with you results
- ✓ Get a badge



DataStax sponsors your education and certification!

Get your voucher and become
a certified NoSQL database developer
for FREE!

dtsx.io/workshop-voucher



Vouchers (normally \$145 each exam)

- valid for 3 months
- valid for 2 attempts



Official DataStax Certification Program

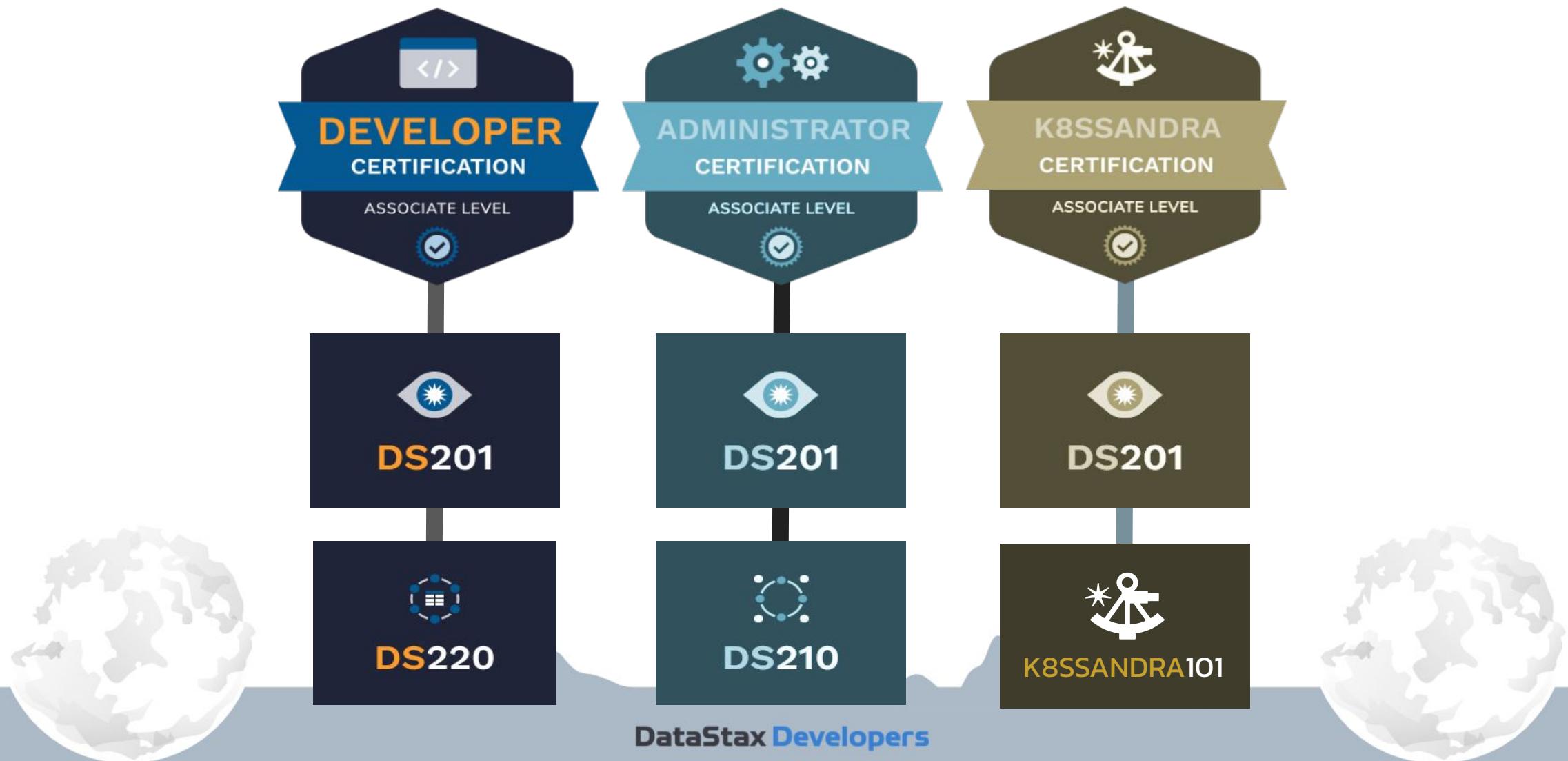
datastax.com/dev/certifications



DataStax Developers



Learning paths at the academy.datastax.com



ASTRA DB'S BUILD-A-THON

MAKING SIDE-HUSTLES A REALITY

21 February - 28 May 2022



ASTRA DB'S BUILD-A-THON

MAKING SIDE-HUSTLES A REALITY

21 February - 28 May 2022



JOIN OUR ASTRA DB BUILD-A-THON HACK!

📍 3 months, 3 rounds of challenges. 📍
Join 1 month, 2 months or all 3

Each month, we'll reveal a fresh new set of challenges you can partake in.

All you have to do is have Astra DB as your backend.

USD\$41,000 worth of prizes



REGISTER -

buildathonhack.com



!discord

dtsx.io/discord

DataStax Developers

workshop-chat

<https://www.youtube.com/watch?v=MuwT5xkFVWI> - Subscribe to mailing list: [http...](http://)

PRESENTER — 1
David Jones-Gilardi

HELPER — 7
012345, AaronP, B1nary, Chelsea Navo, Jeremy Hanna, John Sanda, Patrick_McFadin

EN LIGNE — 560
-samu-, 6304-42J8, Aahlya, Abdurahim, abhi3pathi, Abhiis.s, Abhineet, Abirsh

Événements
moderator-only
. WELCOME
start-here
code-of-conduct
introductions
upcoming-events
useful-resources
memes
your-ideas
the-stage

WORKSHOPS
workshop-chat
workshop-feedback
workshop-materials
upcoming-workshops

ASTRA DB
getting-started
astra-apis
astra-development
sample-applications

APACHE CASSANDRA

Cedrick Lun...

RIGGITYREKT Hier à 21:14
I have a 5 node datacenter, 4 nodes are on dse version 5.1.20, one is on dse5.0.15. I am doing some mixed version testing for a class and the one node that is 5.0.15 is coming up as an analytics workload. I dont have /etc/default/dse, instead I am using /etc/init.d/dse-cassandra. how do i make that node start in cassandra workload, not in analytics?

RIGGITYREKT Hier à 23:39
Okay I found out my issue, when I started DSE 5.0.15 it had endpointSnitch set to DseSimpleSnitch, the rest of my cluster is using PropertyFileSnitch, when I change it to PropertyFileSnitch, it still uses the simple snitch config. looking at the docs I see there is a way to go to GossipingPropertyFileSnitch, but I need the property file one. I can wipe this dbs, do anything with this node to get this done. how do I fix this?
@here

19 novembre 2021

@RIGGITYREKT Okay I found out my issue, when I started DSE 5.0.15 it had endpointSnitch set to DseSimpleSnitch, the rest...
Erick Ramirez Aujourd'hui à 02:19
mixed versions isn't supported and you're guaranteed to run into weird issues that will cause further problems down the track

@RIGGITYREKT I have a 5 node datacenter, 4 nodes are on dse version 5.1.20, one is on dse5.0.15. I am doing some mixed v...
Cedrick Lunven Aujourd'hui à 09:01
When you start a node you have parameters -k for analytics, -g for graph and -s for search. To remove analytics check and remove -k

Envoyer un message dans #workshop-chat

DataStax Developers Discord (18k+)

Subscribe

Introduction to NoSQL!
2:17:59

Crash Course | Introduction to Cassandra for Developers
1:09:34

Introduction to NoSQL Databases
2:18:01

Introduction to NoSQL Databases
2:19:57

#AppDev Learning Series
Week 3
Building your own NETFLIX Clone!
DataStax Developers
2:16:46

Build your own NETFLIX clone!
4K views • Streamed 2 weeks ago

Build your own NETFLIX clone!
7.4K views • Streamed 2 weeks ago

Astra Streaming Demo
177 views • 2 weeks ago

Kubernetes Ingress Management with Traefik Proxy
1:15:49

Build your own TikTok clone!
1.9K views • Streamed 3 weeks ago

Build your own TikTok clone!
4K views • Streamed 3 weeks ago

How to use the Connect Driver in Astra DB
113 views • 4 weeks ago

How to use the CQL Console in Astra DB
39 views • 4 weeks ago

How to create an Authentication Token in Astra DB
37 views • 4 weeks ago

How to use the Data Loader in Astra DB
62 views • 4 weeks ago

Astra DB Sample App Gallery
36 views • 4 weeks ago

How to use Secure Connect in Astra DB
42 views • 4 weeks ago

Cassandra Day India 2021
Room 2: Workshops
#CassandraDay
7:45:45

Cassandra Day India 2021
Room 1: Talks
#CassandraDay
5:07:03

Thank You!

