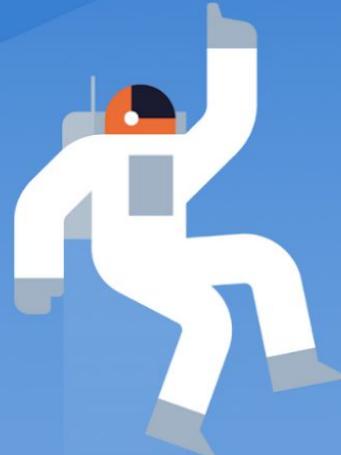




# Introduction to NoSQL Databases



LEVEL  
**UP**  
with the

DataStax  
**Developers**

# Your presenters



**Cedrick Lunven**

Director of Developer Advocacy

*Apache Cassandra™ expert*

*Open Source Developer*

*Java Geek*

 @clun

  @clunven

FF4J



**David Jones-Gilardi**

Developer Advocate

*Apache Cassandra™ expert*

*Experienced developer and educator*

*Still have an Oracle 8 cert somewhere from the mid 90's*



@SonicDMG



@david-gilardi

# Your presenters



**Ryan Welford**  
Developer Advocate

*Apache Cassandra™ expert*  
*Front End Developer*



**David Jones-Gilardi**  
Developer Advocate

*Apache Cassandra™ expert*  
*Experienced developer and educator*  
*Still have an Oracle 8 cert somewhere from the mid 90's*

 [@RyanWelford](#)

 [@ryanwelford](#)

  [@SonicDMG](#)

 [@david-gilardi](#)

# Your presenters



**Aleks Volochnev**

Developer Advocate at DataStax

- Apache Cassandra™ expert
- Experienced developer and educator
- Certified cloud architect



**Cedrick Lunven**

Director of Developer Advocacy at DataStax

- Apache Cassandra™ expert
- Kubernetes rookie
- Java Geek



# Your presenters



**Aleks Volochnev**

Developer Advocate at DataStax

- Apache Cassandra™ expert
- Experienced developer and educator
- Certified cloud architect

   @hadesarchitect



**Ryan Welford**

Developer Advocate

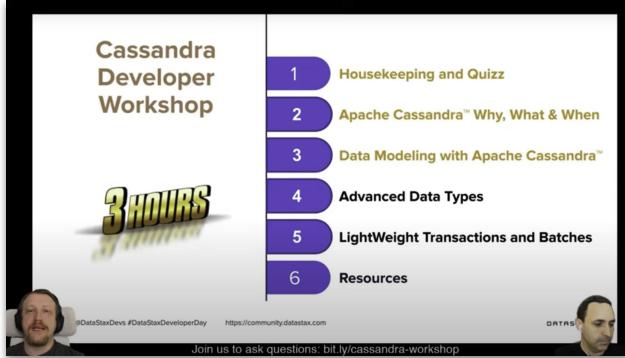
*Apache Cassandra™ expert*

*Front End Developer*

  @RyanWelford

# Housekeeping

**Livestream:** youtube.com/DataStaxDevs



**YouTube**

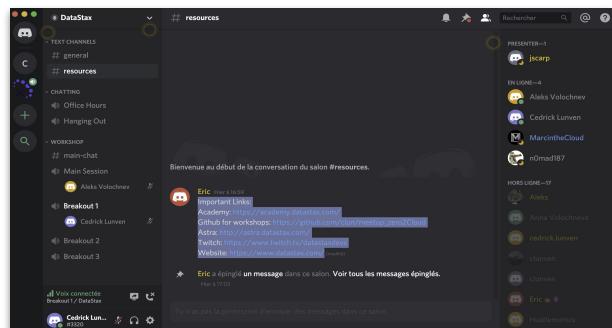


**Twitch**

**Runtime:** astra.new/intro-nosql

**DataStax**  
**Astra**

**Questions:** https://dtsx.io/discord



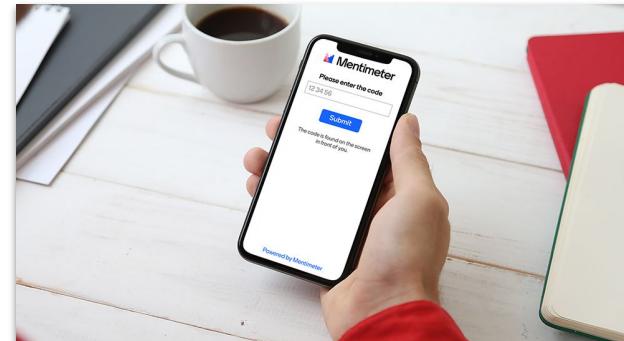
**Discord**



**YouTube**



**Quizz:** menti.com



**Mentimeter**



# Achievement Unlocked!

[dtsx.io/badges](https://dtsx.io/badges)



**K8ssandra Workshop**

Awarded to **Sylwester Lachiewicz** · slachiewicz@gmail.com  
Issued on **Mar 5, 2021**

Upgrade Complete! This badge is to certify successful completion of the DataStax K8ssandra Workshop: "Running Apache Cassandra on Kubernetes".

 **Verified**  
Last verified by Badgr on **Mar 31, 2021**

[Re-verify Badge](#)

**EARNING CRITERIA**  
Recipients must complete the earning criteria to earn this Badge

To earn this badge, individuals must complete the following steps during the **K8ssandra Workshop**:

- Attend the lecture
- Complete the practical steps by doing all required exercises

[View External Criteria](#)

Offered By  
**DataStax Developers**

**TAGS**  
kubernetes  
cassandra



Aron L.  
Marcin  
Brzozowski  
Demre Buyuk  
Muthu Krishnan  
Arvind V.  
Parth Trambadiya  
Jasbir Singh

Prateek Jain  
Roozbeh Dargahi  
Andrey Deryabin  
Akshay Wakhare  
Pranav Anant  
Joshi  
Haris  
Juan Alonso

Santosh Nepali  
Jorge Ortiz  
Ankit Bhavsar  
Aneliya Klevleeva  
Martin Coronel  
Govindasamy  
Ville Kerminen  
Priya Jakhar

Paul Robu  
Avinash Upadhyaya  
Włodzimierz Kozłowski  
Sharath Koushik  
Tom Rota  
Joel Reis  
Francesco Abbate

Sylwester Lachiewicz  
Ankit Bhavsar  
Jasbir Singh

AND MANY OTHERS!



## Anant

Business Platform Architects - We connect Customer Experience and Information Systems platforms with Data & Analytics  
Internet · Washington, District of Columbia · 409 followers



# Careers at Anant: [careers.anant.us](http://careers.anant.us)

"Anant has a few openings in our Data Platform team, and we are looking for candidates for these key roles. We are on a fast track within the Cassandra industry and are seeking talented DataStax Enterprise Certified Professionals to work on interesting and challenging projects."

# menti.com



Go to [www.menti.com](http://www.menti.com) and use the code 3491 9972

## Inequality predicates are allowed on ...

A bar chart titled "Inequality predicates are allowed on ...". The x-axis categories are "All table columns", "Partition key columns", "clustering key columns", and "No Inequality predicates are allowed". The y-axis ranges from 1 to 15. A green checkmark is above the "clustering key columns" bar at value 15. Red X marks are above the other three categories at values 4, 3, and 1 respectively. The bars are colored light blue, yellow, green, and pink.

4 All table columns  
3 Partition key columns  
15 clustering key columns  
1 No Inequality predicates are allowed

2:10:19 / 2:26:05

Go to [www.menti.com](http://www.menti.com) and use the code 3491 9972

## Leaderboard

4821 p	spanda
4820 p	Agent X9
4775 p	fastest
4711 p	Sam
4468 p	CCedrickThePresenter
4371 p	shubham
3895 p	aaa
3877 p	vignesh
3861 p	adry
3812 p	Millie
	Puggie

2:11:07 / 2:26:05

DataStax

# Hands-on exercise material



**Get your instance here:**

- <http://astra.new/intro-nosql>



**Repository:**

- <https://github.com/datastaxdevs/workshop-introduction-to-nosql>



# Agenda

**01**

Definitions and  
objectives of NoSQL

**02**

Tabular  
Databases



**03**

Document  
Databases



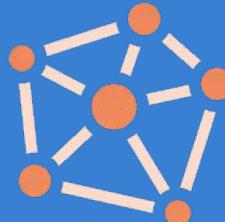
**04**



Key/values  
Databases

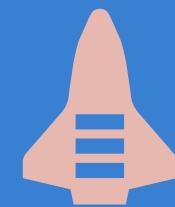
**05**

Graph  
Databases



**06**

Games  
TakeAways



# Agenda

**01**

Definitions and  
objectives of NoSQL

**02**

Tabular  
Databases



**03**

Document  
Databases



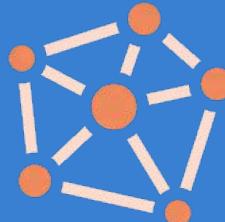
**04**

Key/values  
Databases



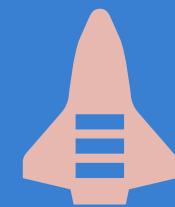
**05**

Graph  
Databases



**06**

Games  
TakeAways



# Get Ready = Hands-on #1



DataStax

Astra

**Get your instance here:**

- <http://astra.new/intro-nosql>



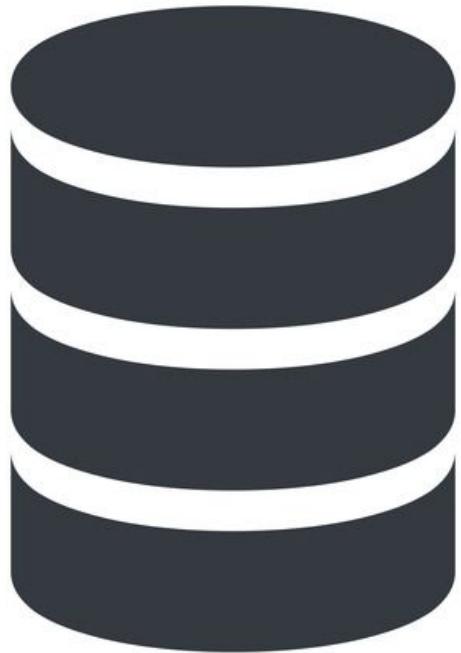
GitHub

**Repository:**

- <https://github.com/datastaxdevs/workshop-introduction-to-nosql>



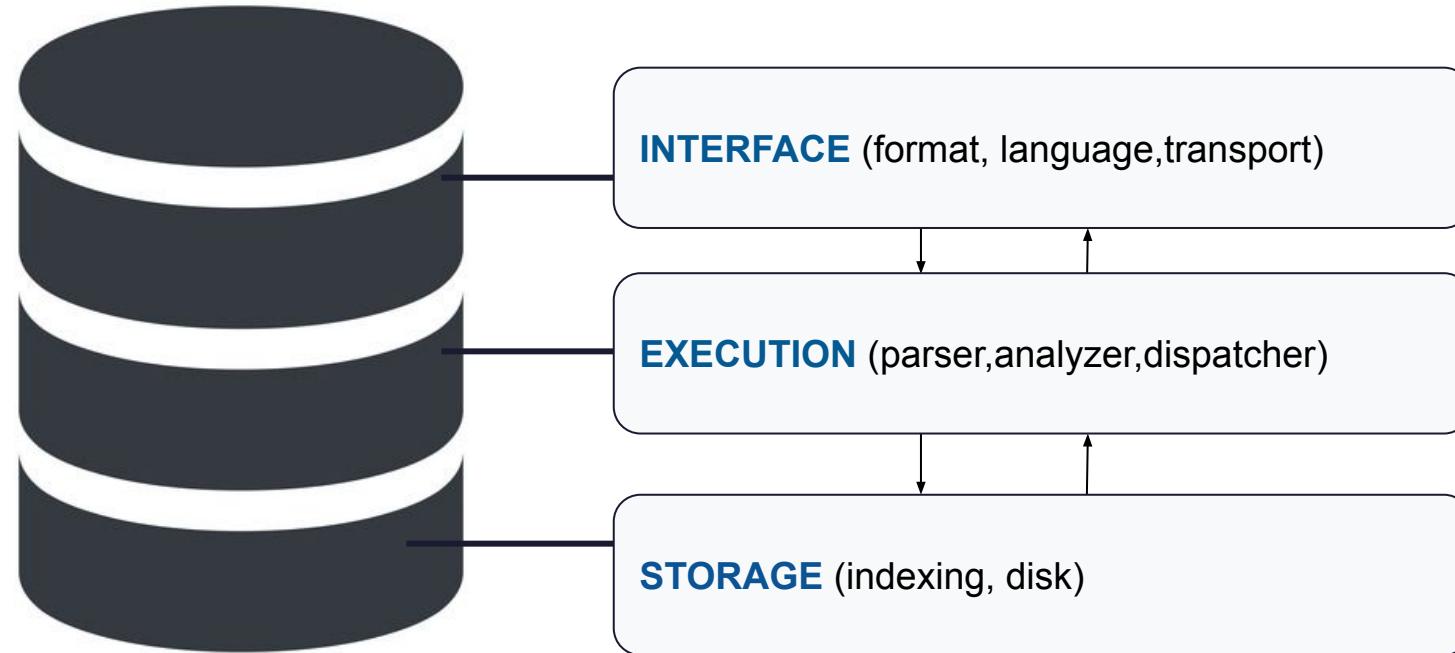
# Databases



**Software to save things  
and retrieve them later  
with queries**

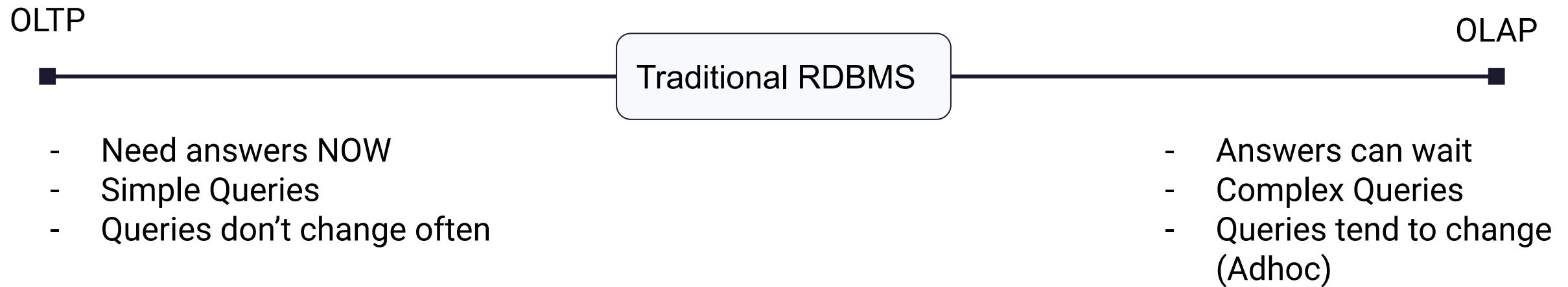
“That's all Folks!”

# Databases

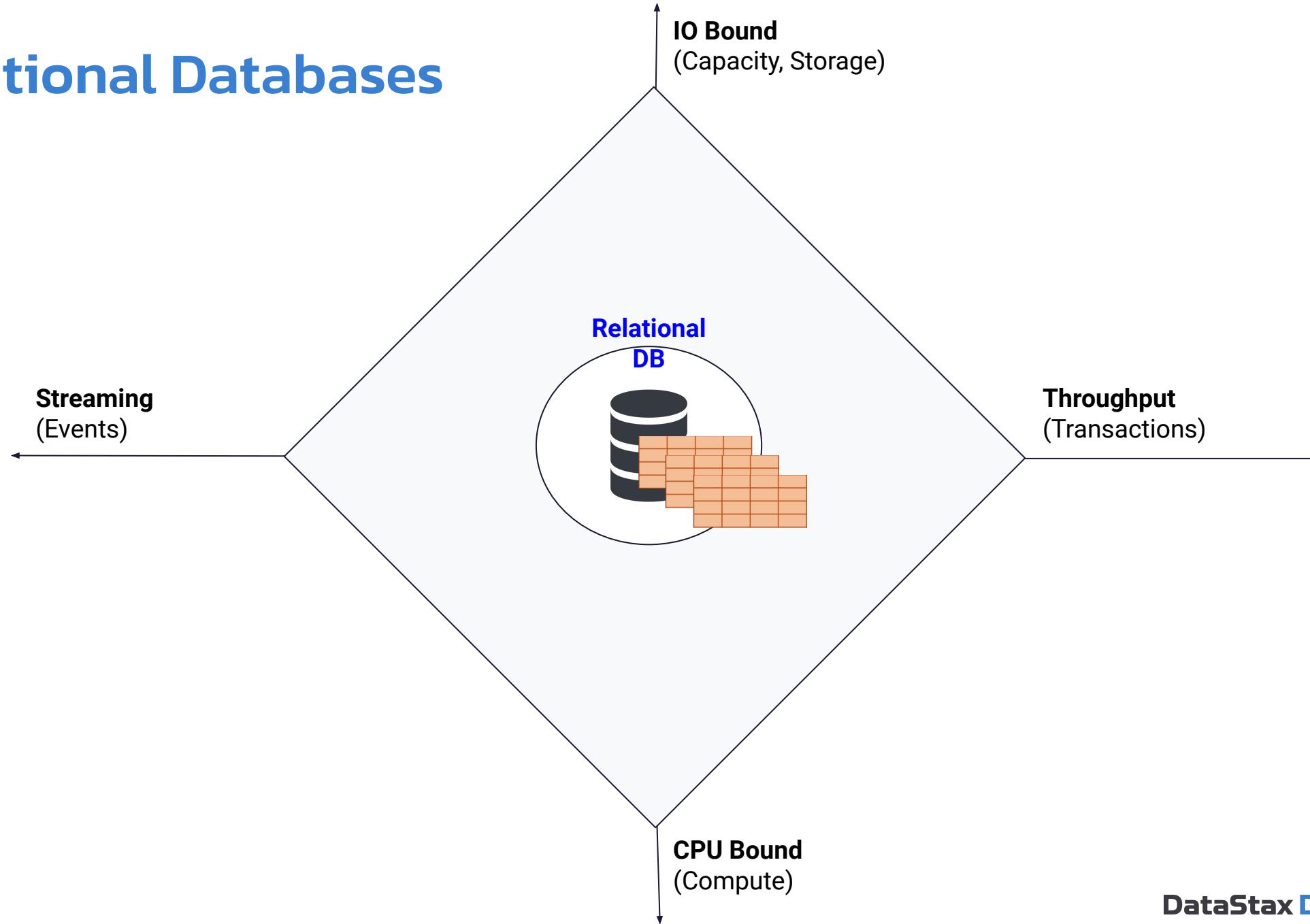


# OLTP / OLAP

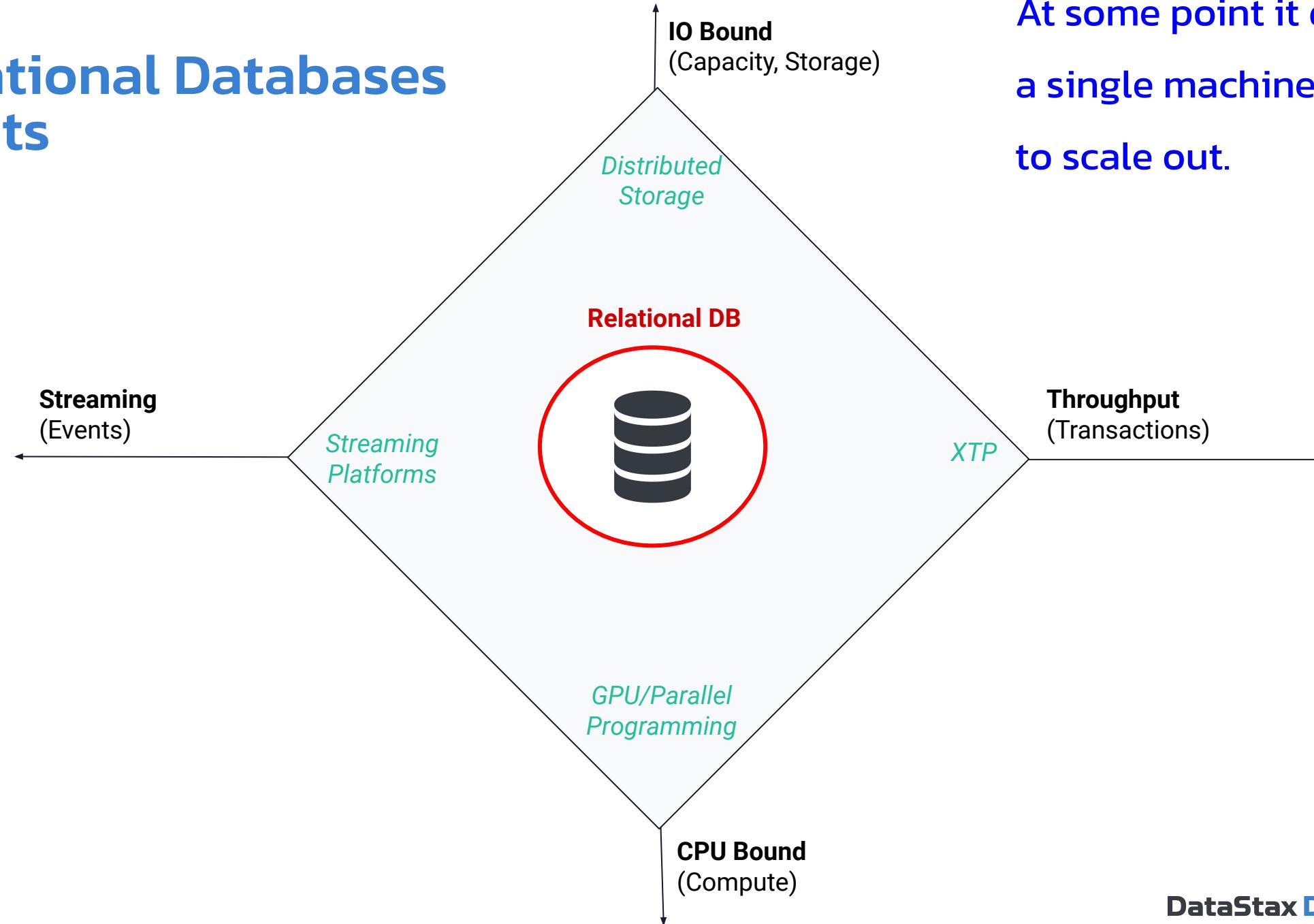
- OnLine Transaction Processing
- OnLine Analytical Processing



# Relational Databases



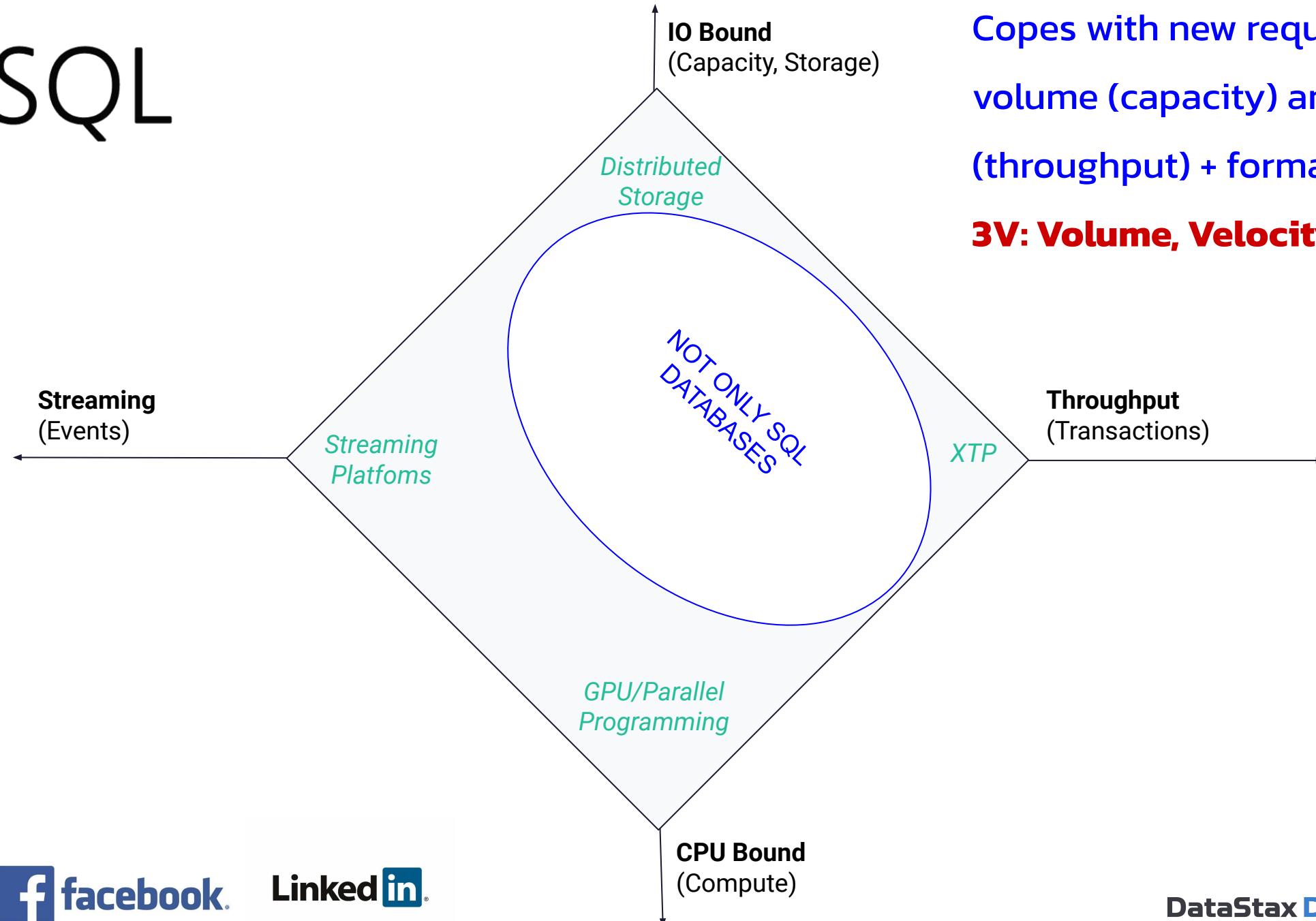
# Relational Databases Limits



At some point it does not fit  
a single machine you need  
to scale out.



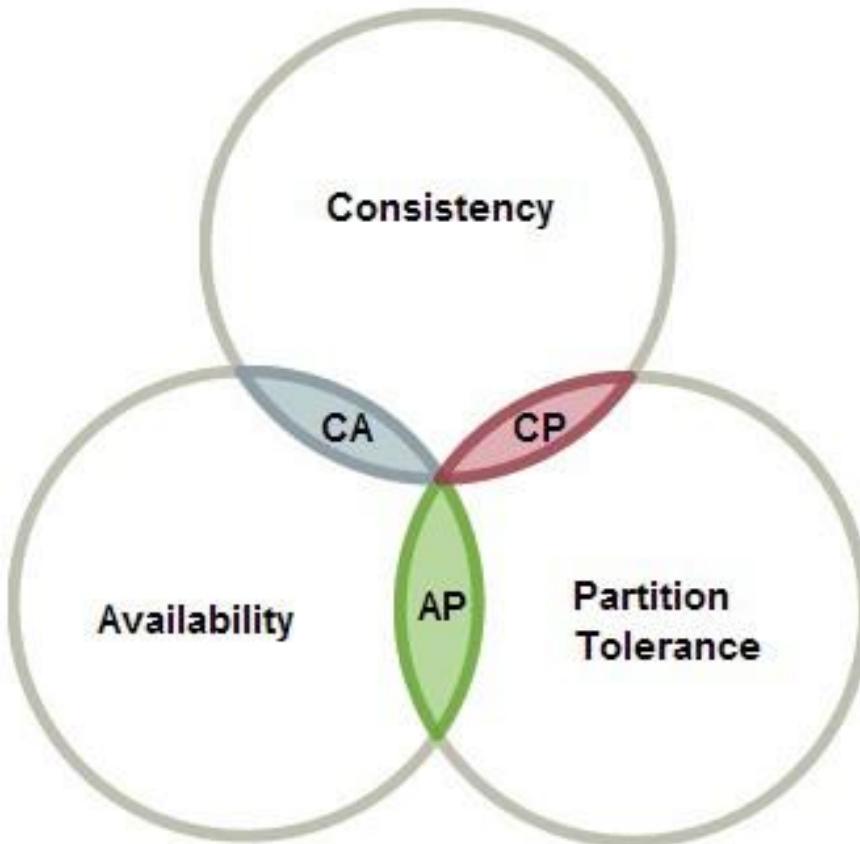
# Not Only SQL



Copes with new requirements in volume (capacity) and velocity (throughput) + format (variety)

**3V: Volume, Velocity, Variety**

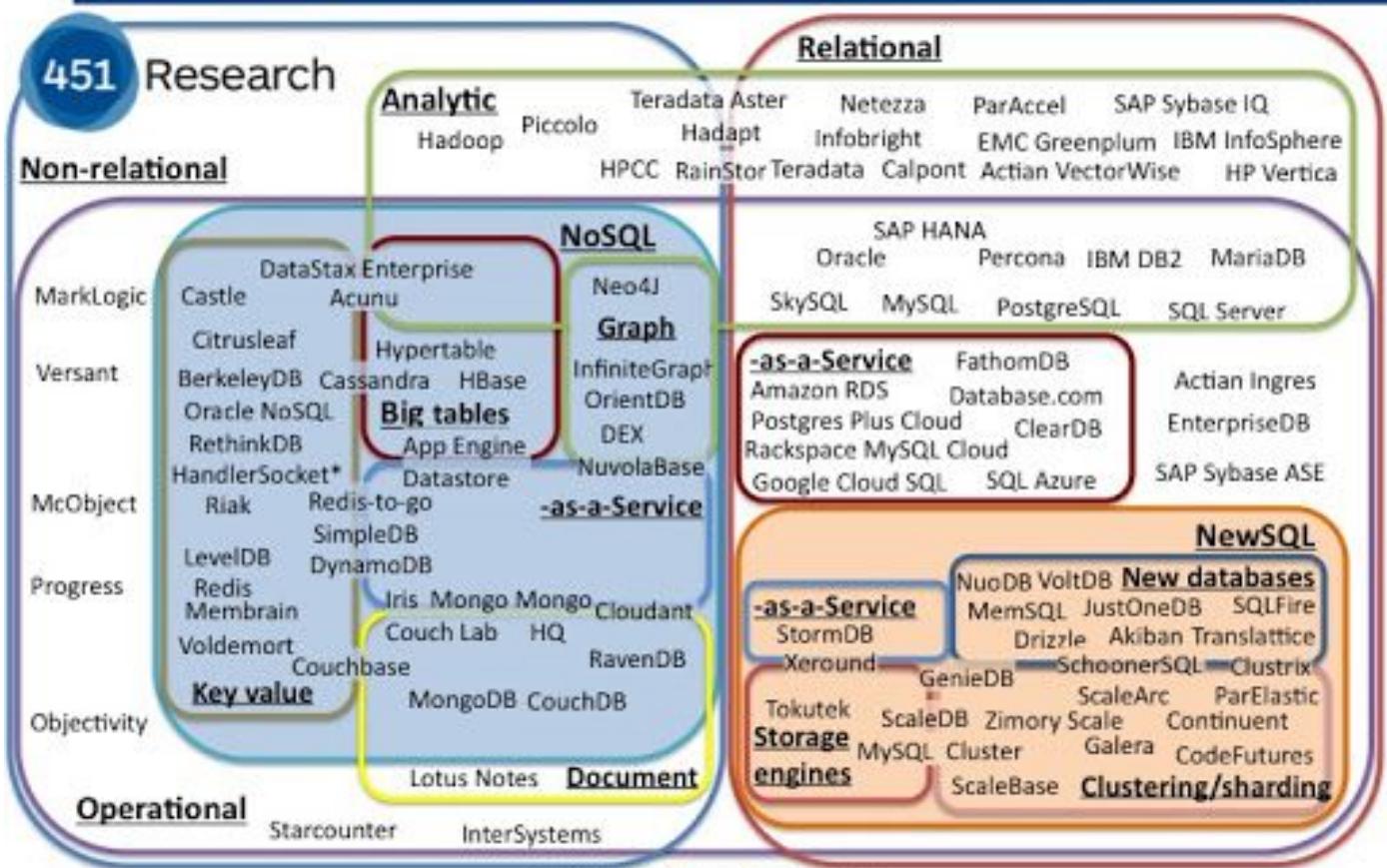
# Introduction to the C.A.P Theorem (Eric Brewer)



**NoSQL are  
Distributed Systems**

**Clouds like  
Distributed Systems**

## The evolving database landscape



# Main NoSQL Databases Types

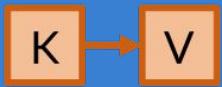
Column Oriented  
Tabular



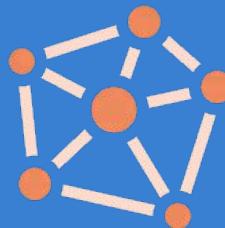
Document



Key/value



Graph



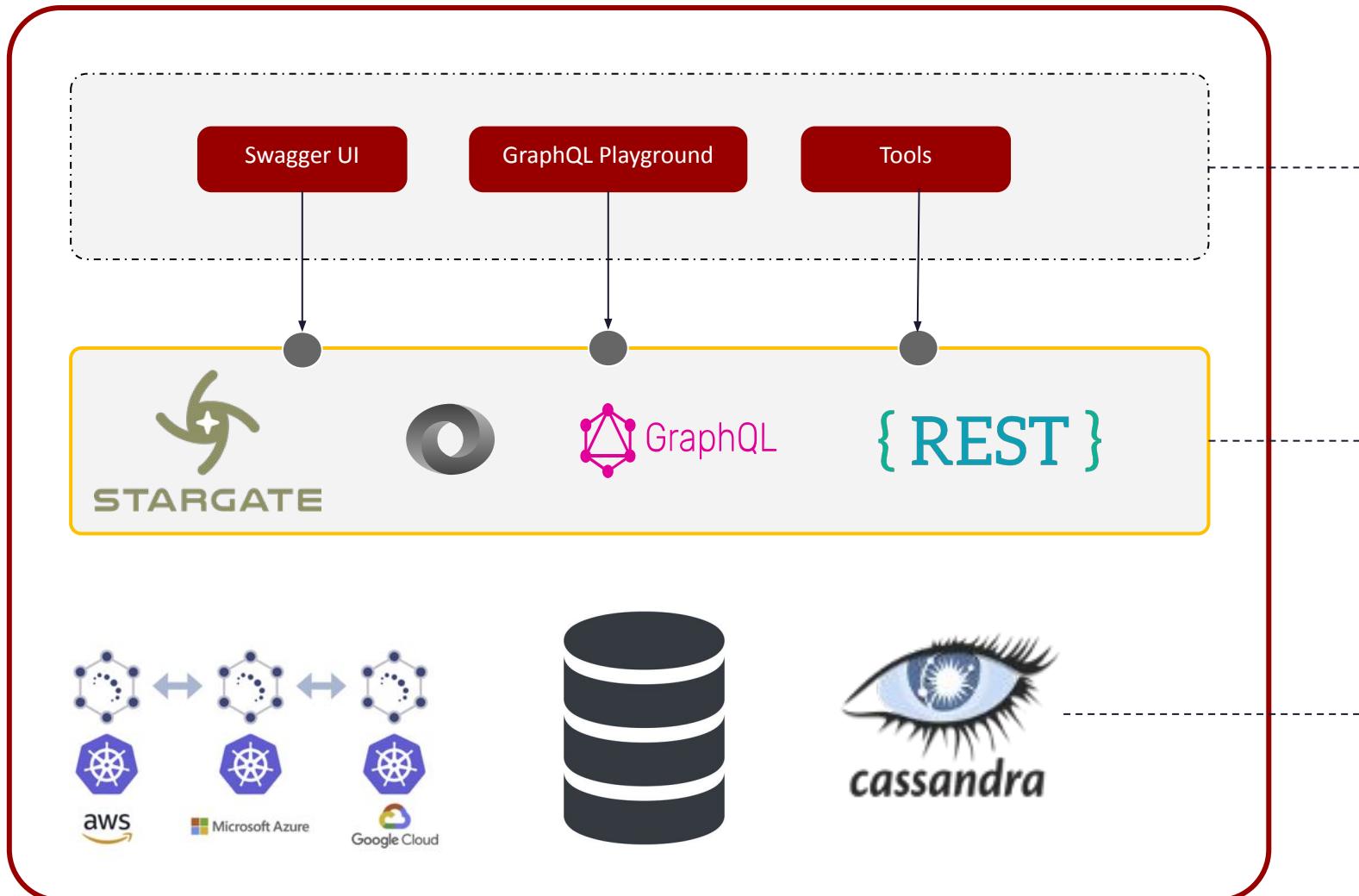
DataStax

# 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

**OSS Stargate.io**  
A data gateway to allow  
multiple usages



**OSS Apache Cassandra**  
A Column oriented NoSQL  
Database



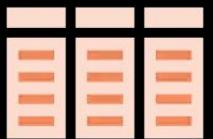
# Agenda

**01**

Definitions and  
objectives of NoSQL

**02**

Tabular  
Databases



**03**

Document  
Databases



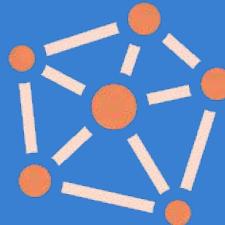
**04**



Key/values  
Databases

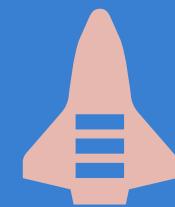
**05**

Graph  
Databases



**06**

Games  
TakeAways



# Tabular or Column Type



**Model:** Stored Tables sharded on keys to distribute on nodes

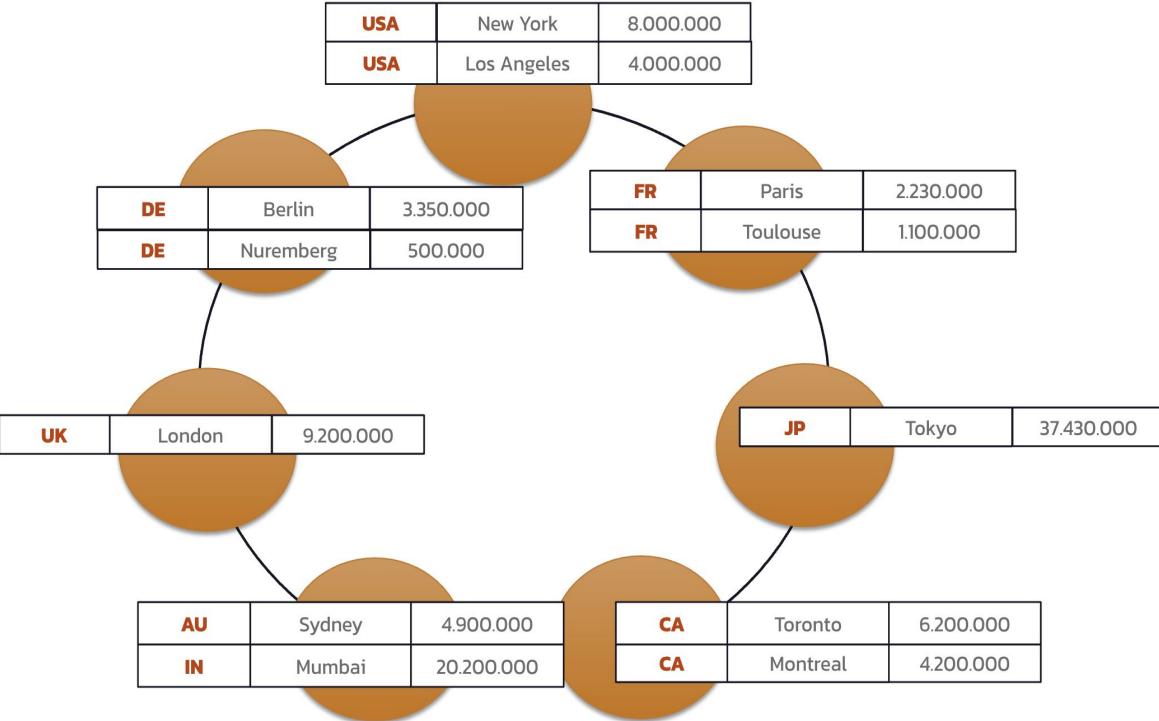
- Tables like relational (with a Schema)
- Distributes data based on a key
- Data is stored sorted on disk

## Query

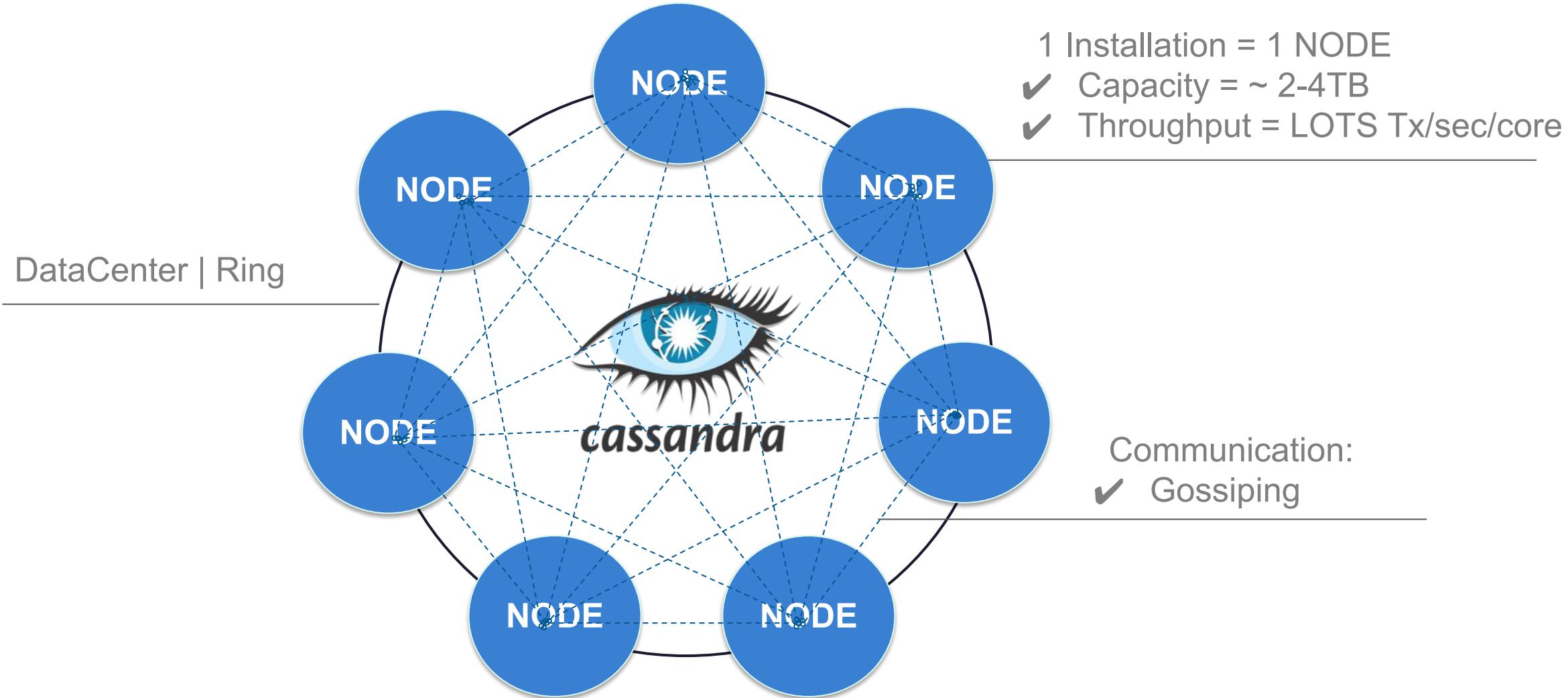
- Request with the partition key
- Secondary Indices are possible
- Select one or more columns for the record
- No joins but denormalization

## Use Cases

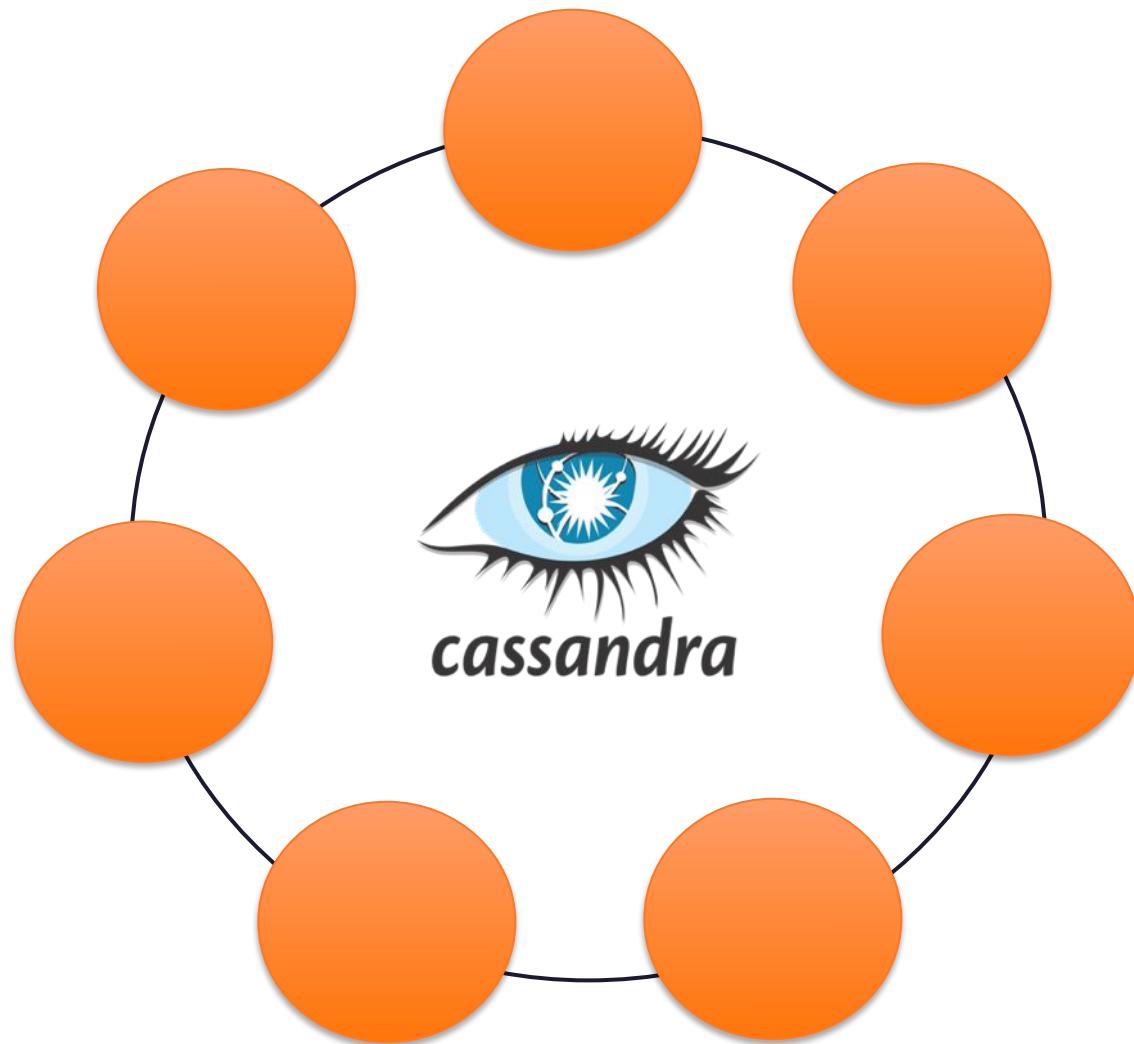
- CF CASSANDRA AFTER THAT



# Apache Cassandra™ = NoSQL Distributed Database



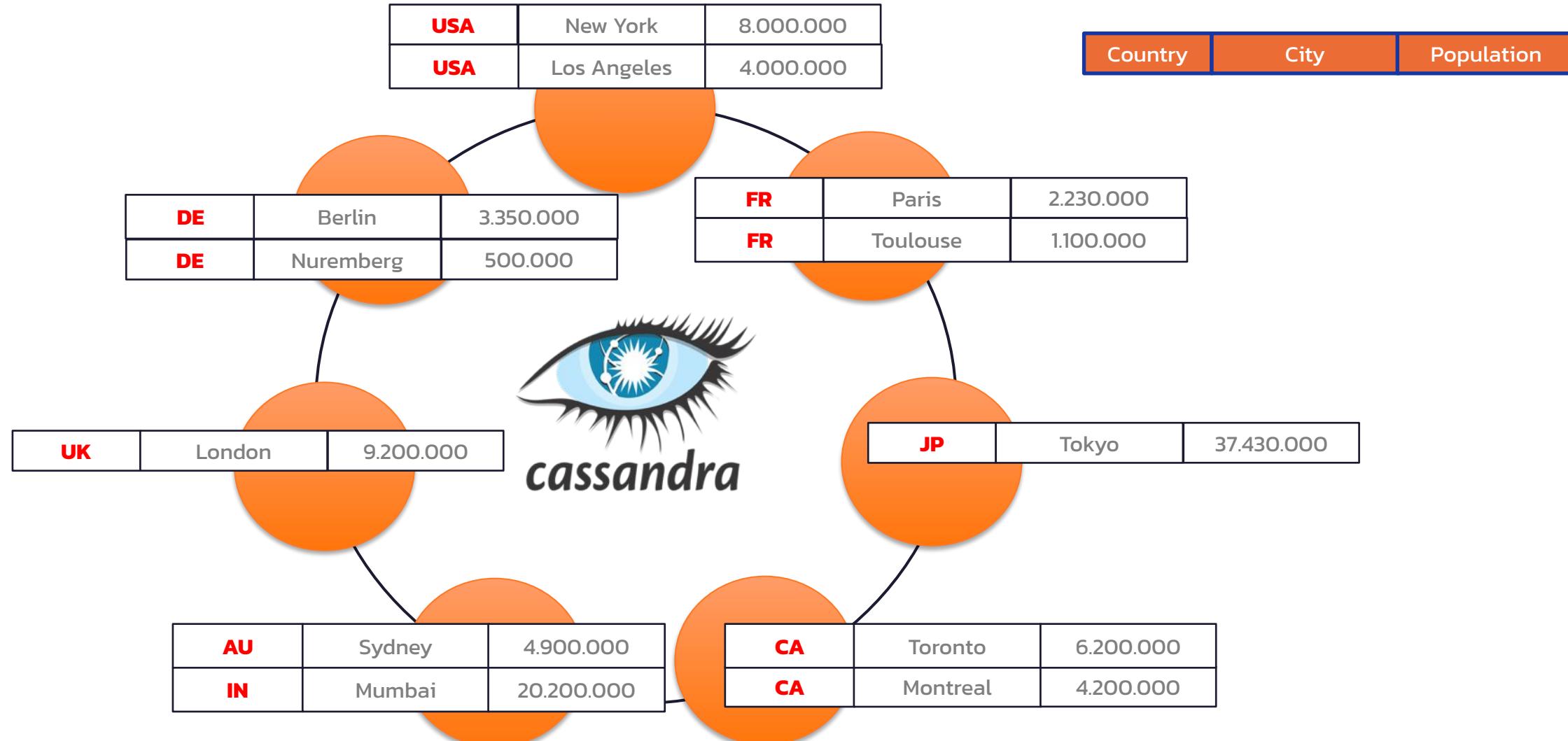
# Data is Distributed



Country	City	Population
USA	New York	8.000.000
USA	Los Angeles	4.000.000
FR	Paris	2.230.000
DE	Berlin	3.350.000
UK	London	9.200.000
AU	Sydney	4.900.000
DE	Nuremberg	500.000
CA	Toronto	6.200.000
CA	Montreal	4.200.000
FR	Toulouse	1.100.000
JP	Tokyo	37.430.000
IN	Mumbai	20.200.000

Partition Key

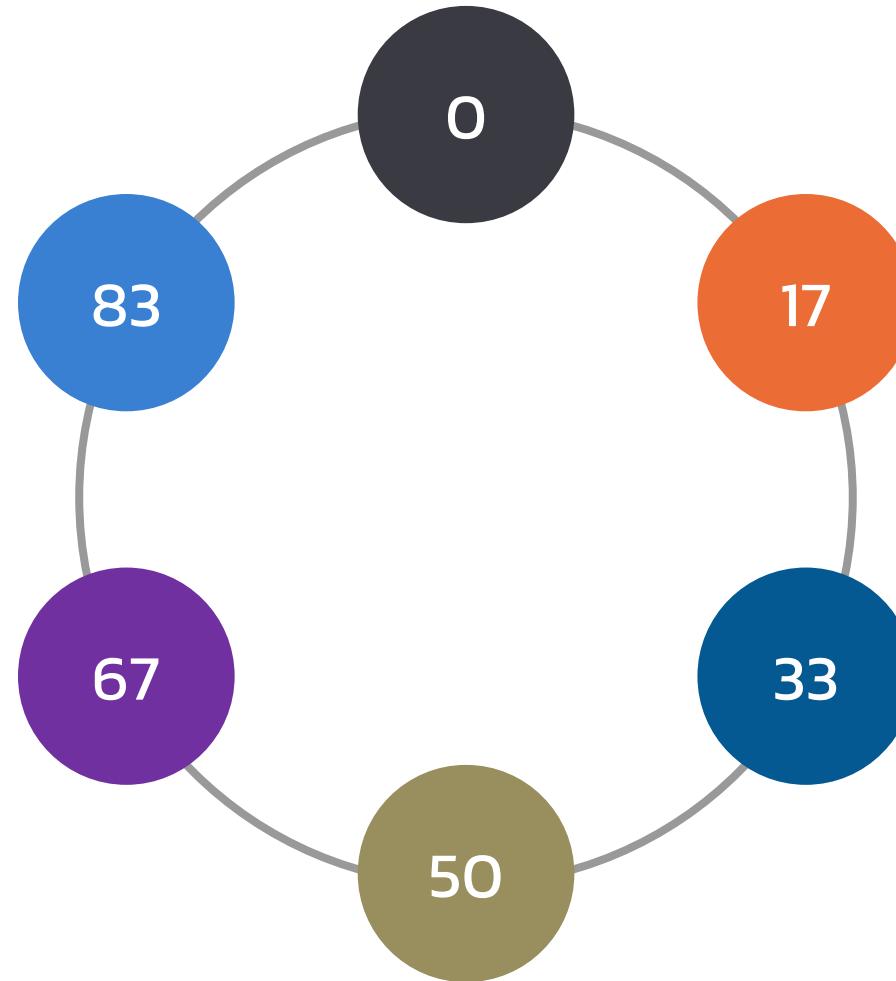
# Data is Distributed



# Data is Replicated

RF = ?

Replication Factor  
means the number  
of nodes used to  
store each partition

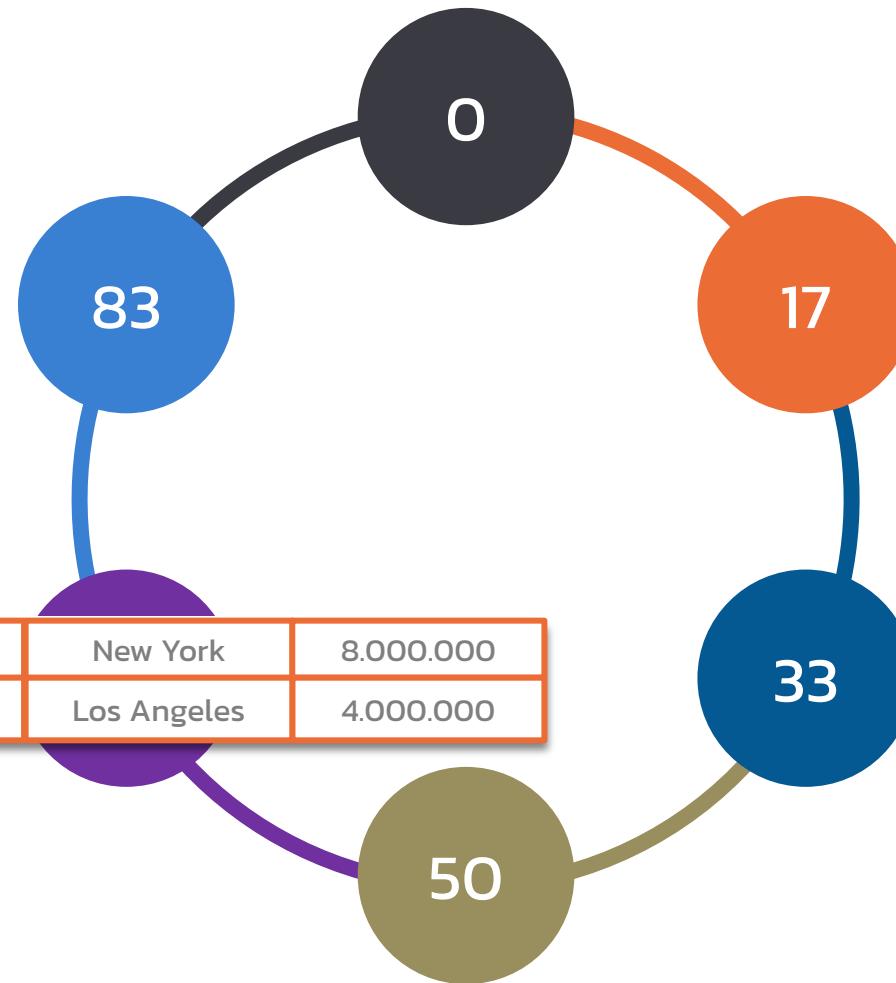


# Data is Replicated

RF = 1

Replication Factor 1  
means that every  
partition is stored  
on 1 node

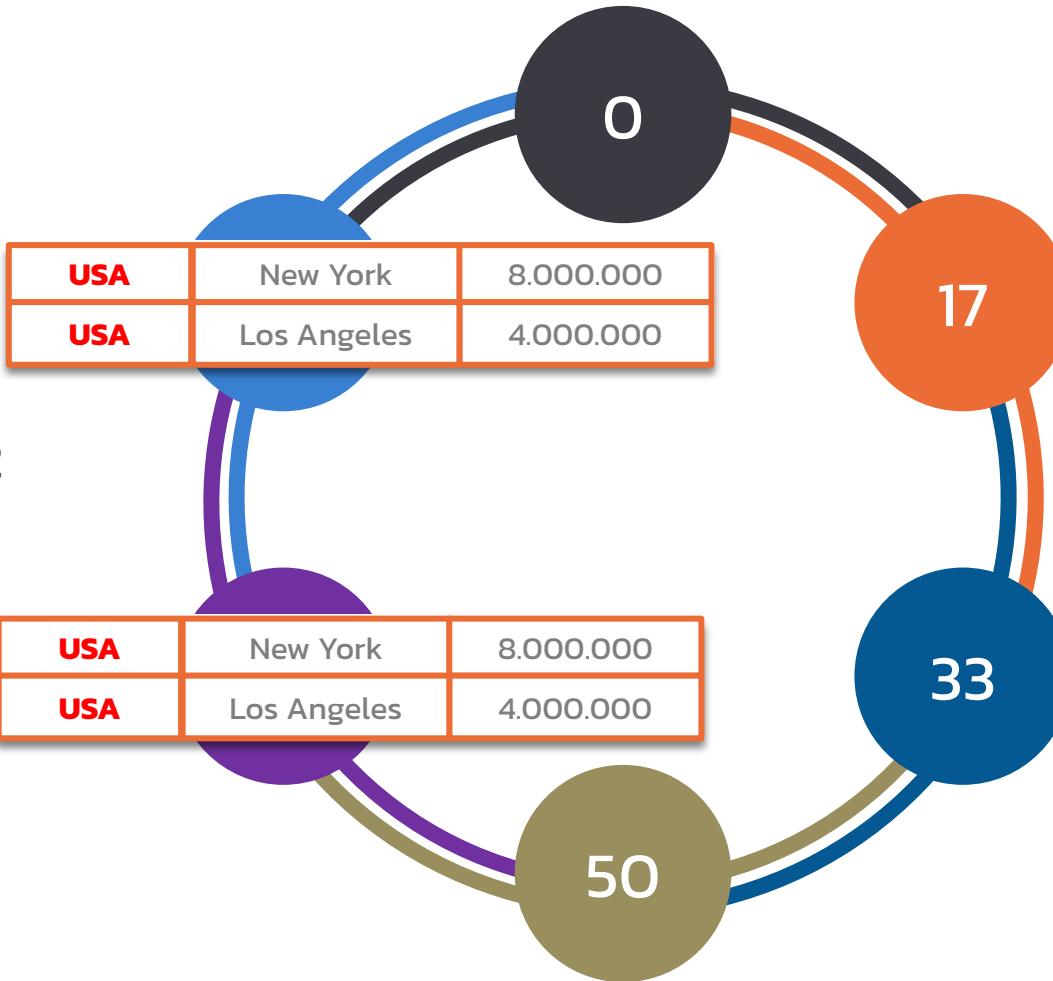
USA	New York	8.000.000
USA	Los Angeles	4.000.000



# Data is Replicated

RF = 2

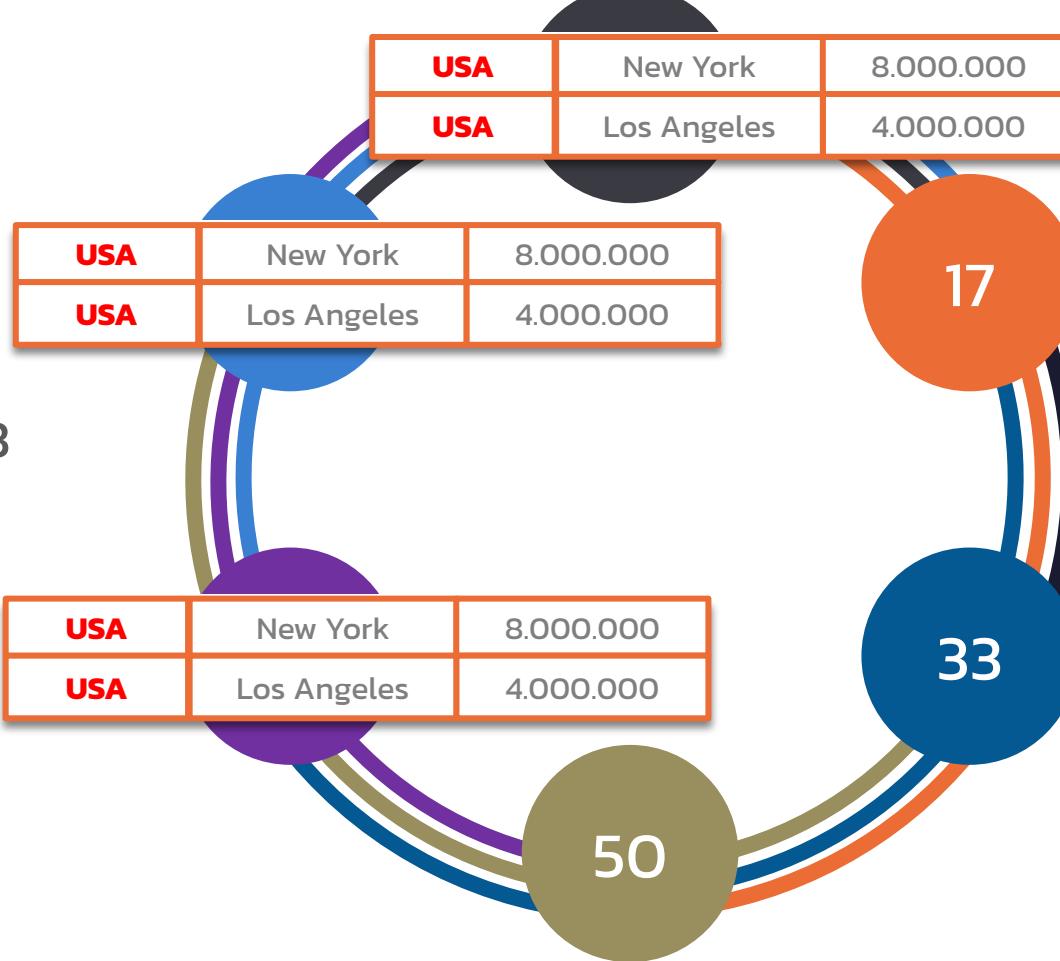
Replication Factor 2  
means that every  
partition is stored  
on 2 nodes



# Data is Replicated

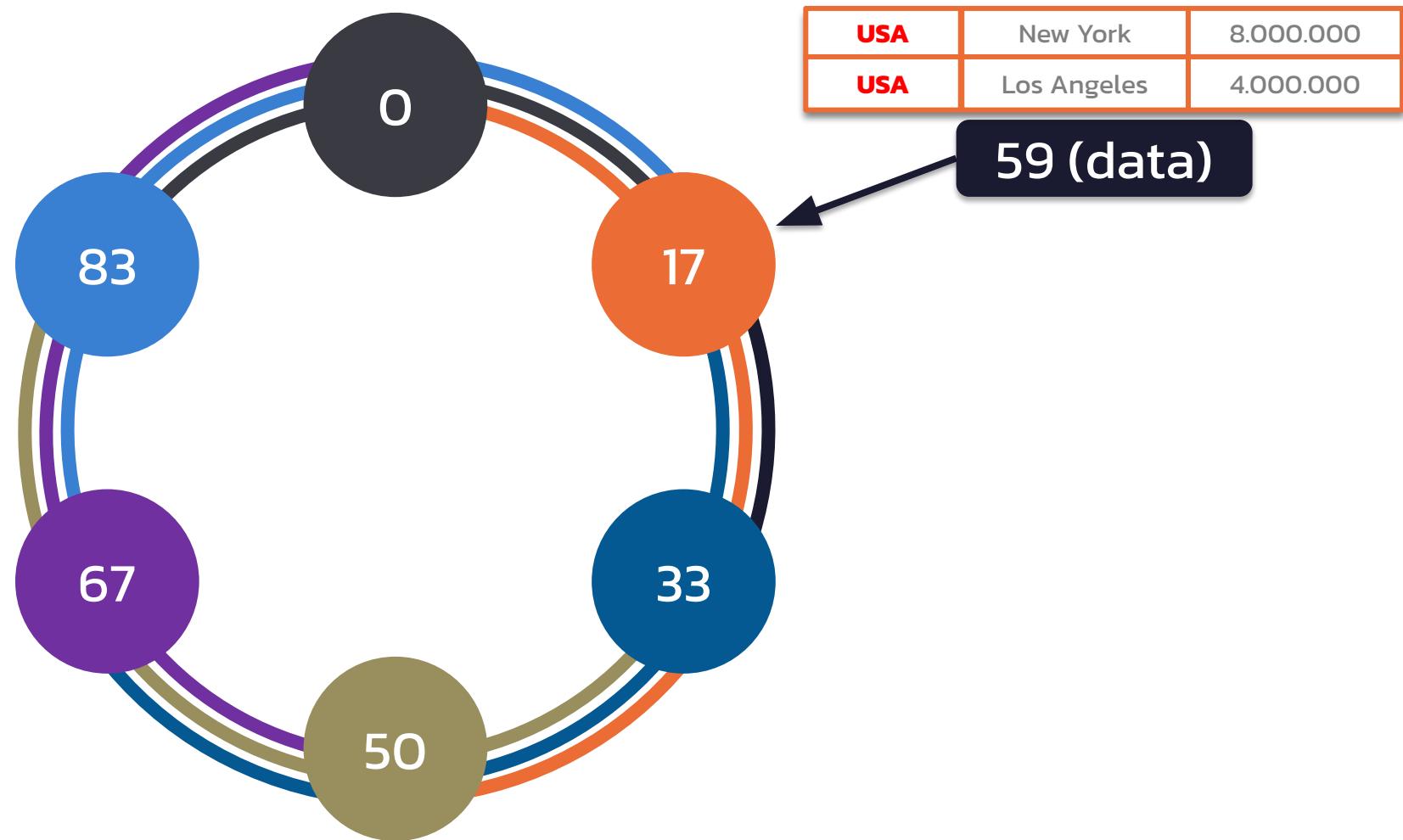
RF = 3

Replication Factor 3  
means that every  
partition is stored  
on 3 nodes



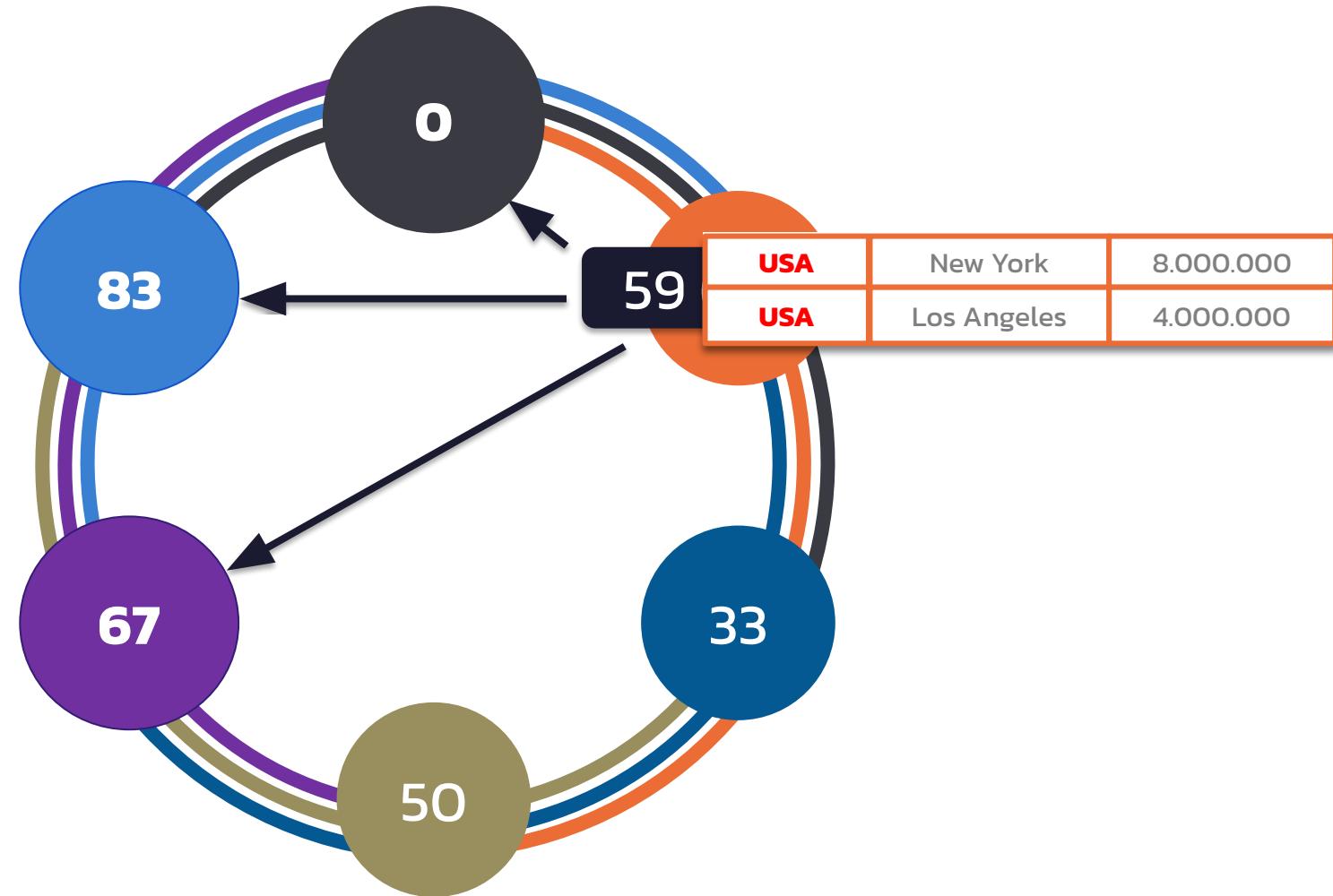
## Replication within the Ring

RF = 3



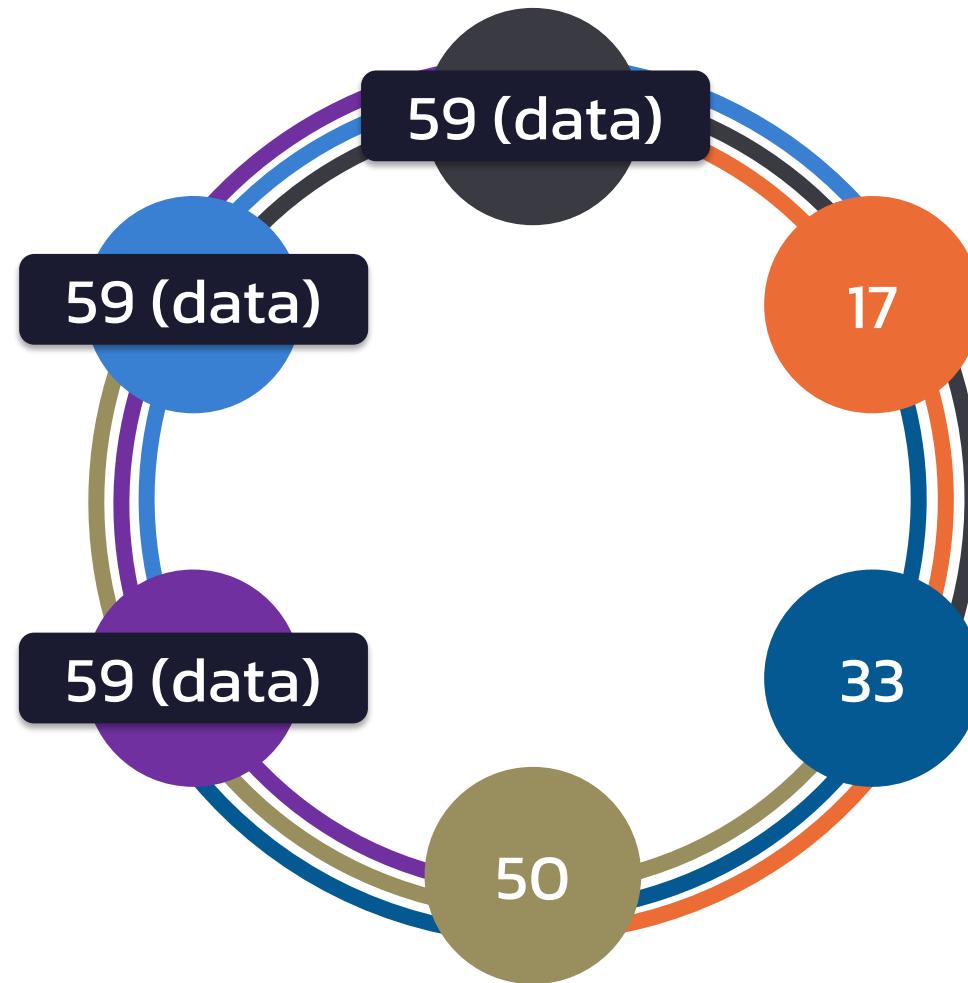
## Replication within the Ring

RF = 3



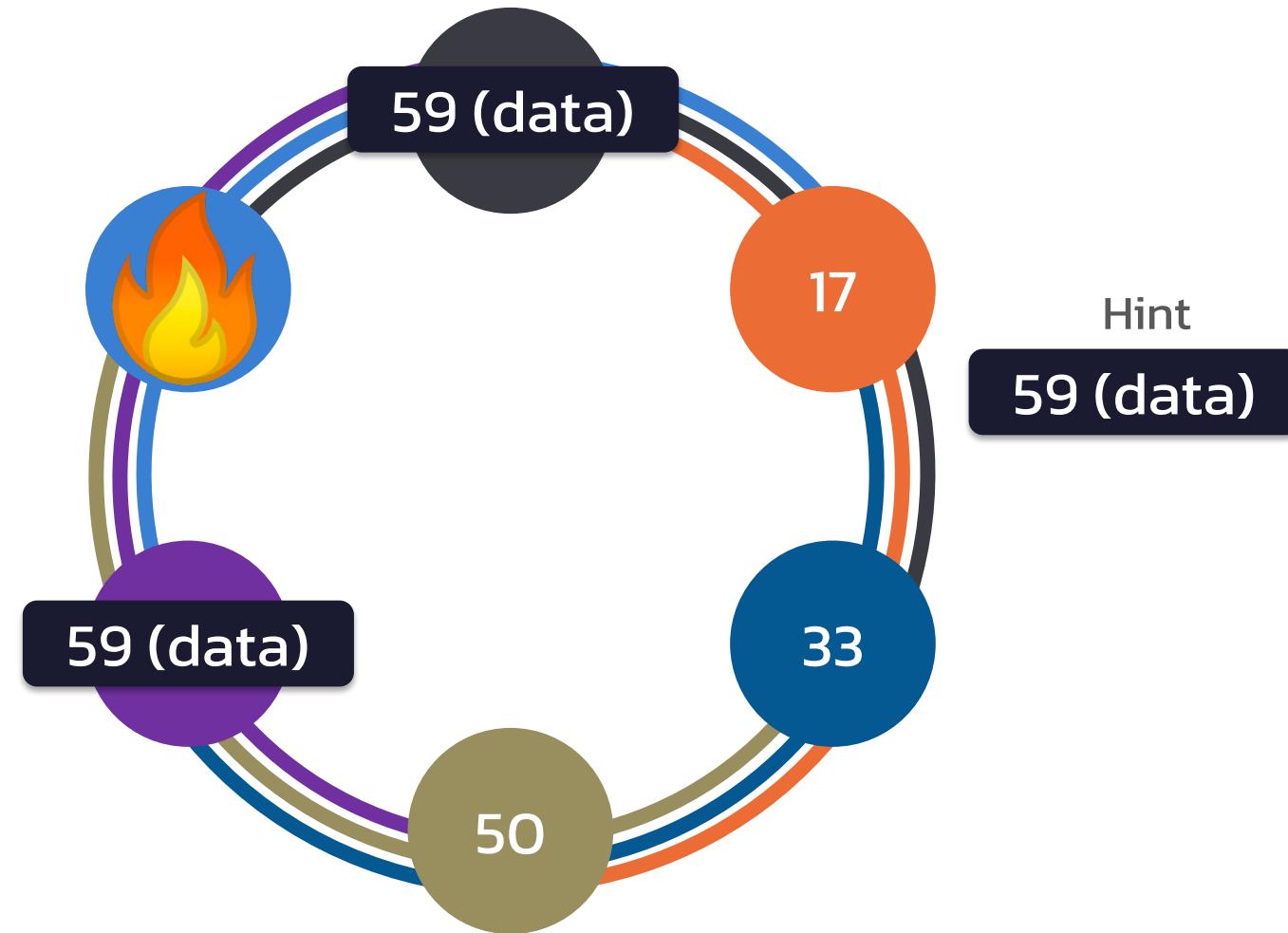
## Replication within the Ring

RF = 3



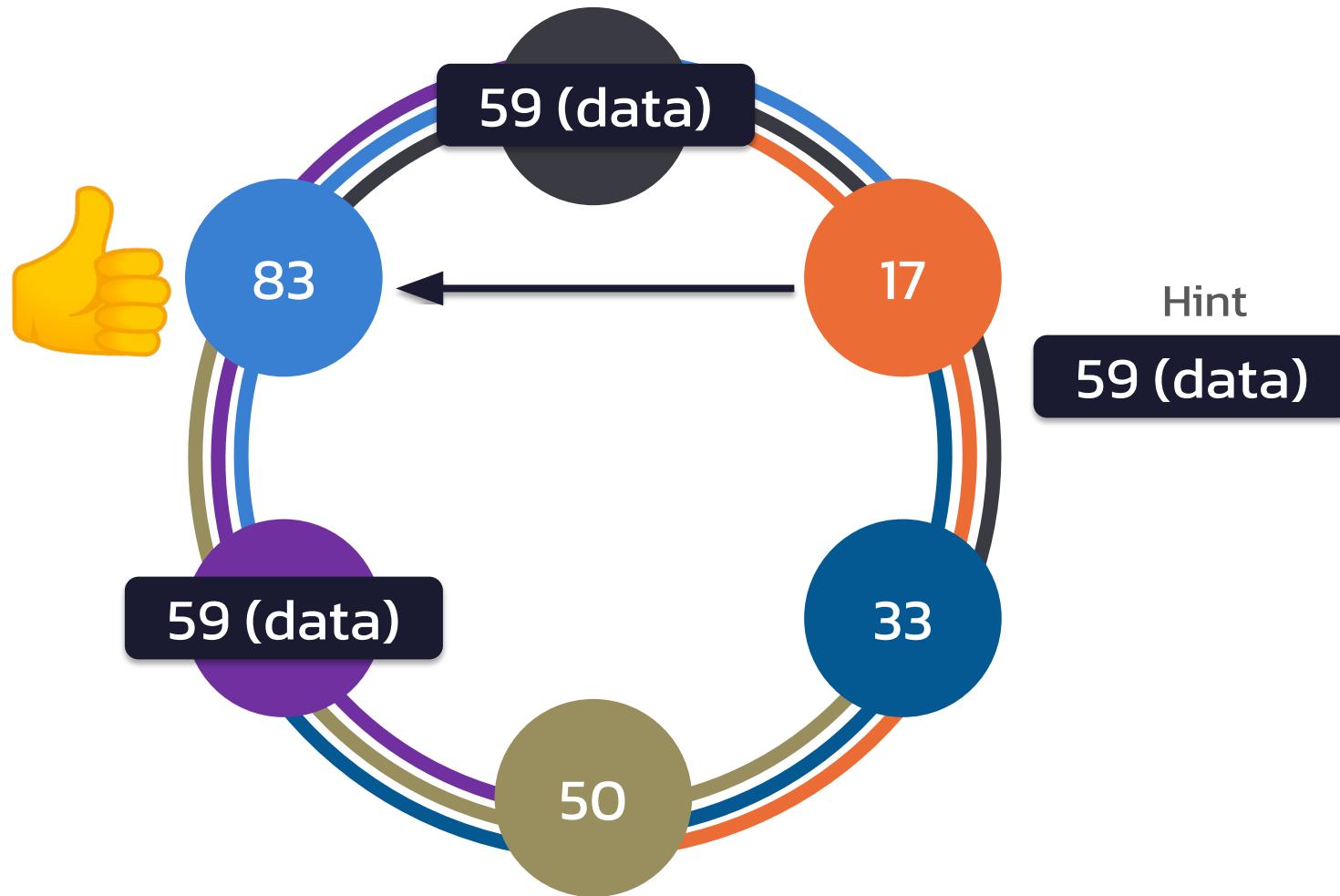
# Node Failure

RF = 3



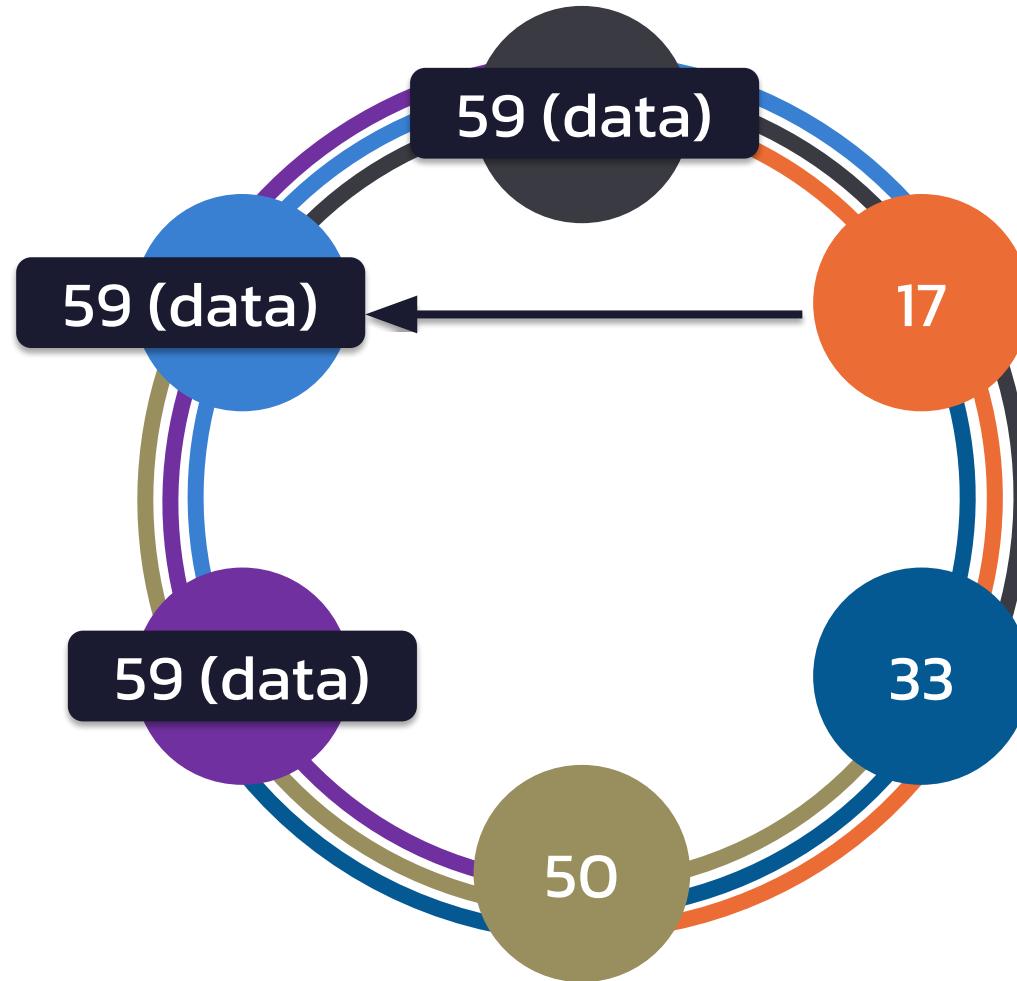
# Node Failure Recovered

RF = 3



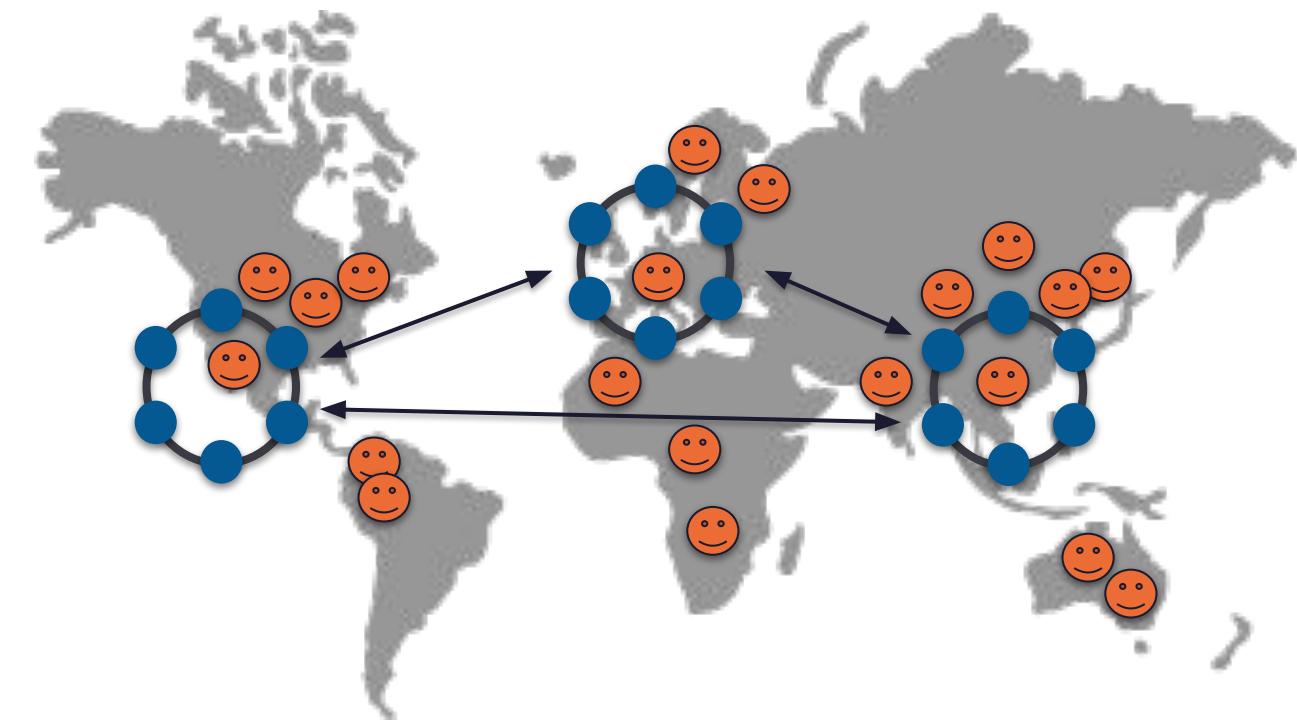
## Node Failure Recovered

RF = 3

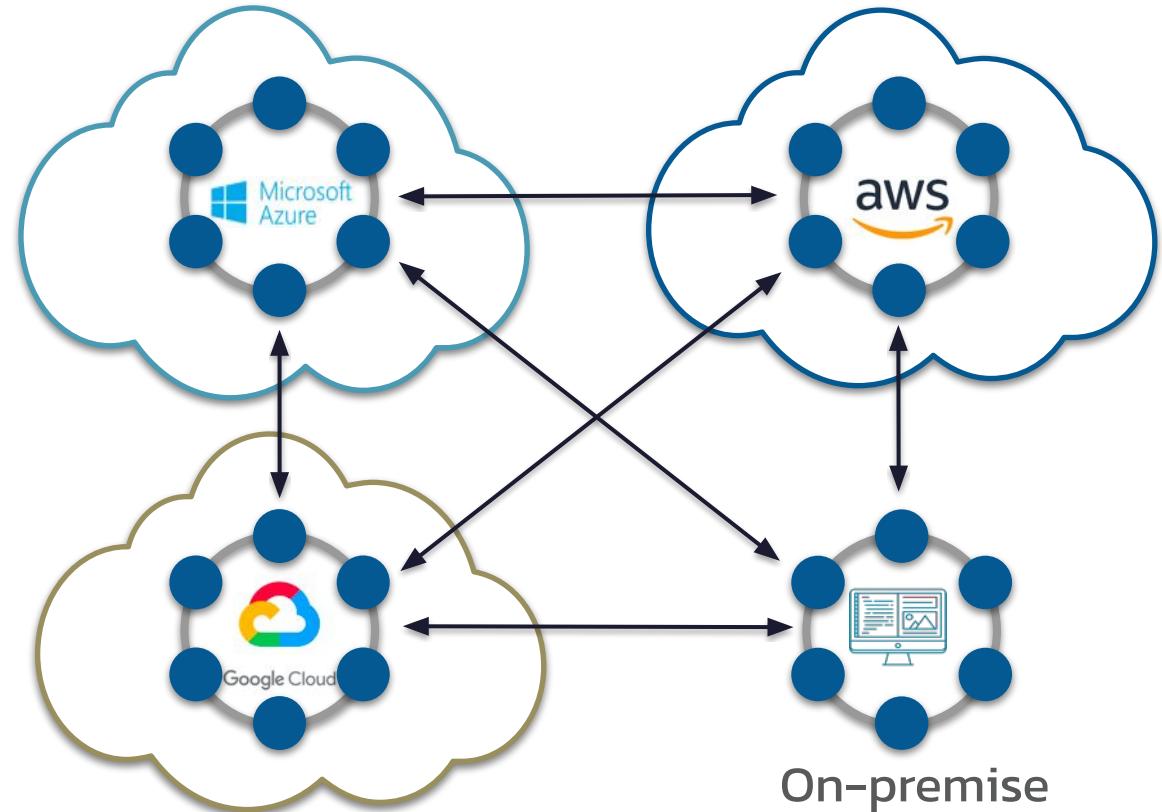


# Data Distributed Everywhere

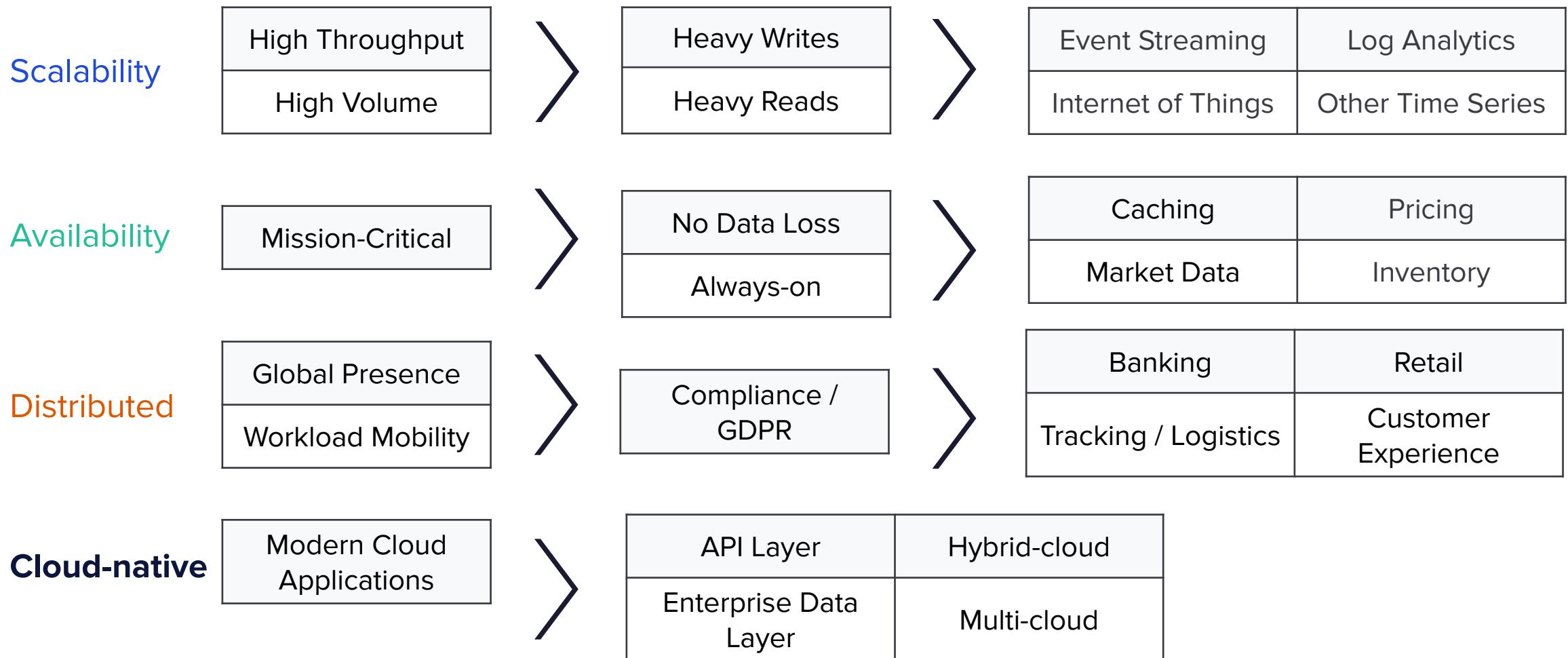
## Geographic Distribution



## Hybrid-Cloud and Multi-Cloud



# Understanding Use Cases



# HandsOn #2 Tabular Databases



DataStax

Astra

**Get your instance here:**

- <http://astra.new/intro-nosql>



GitHub

**Repository:**

- <https://github.com/datastaxdevs/workshop-introduction-to-nosql>



# Agenda

**01**

Definitions and  
objectives of NoSQL

**02**

Tabular  
Databases



**03**

Document  
Databases



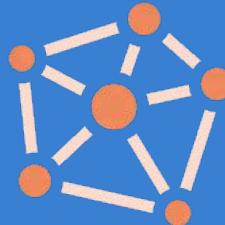
**04**



Key/values  
Databases

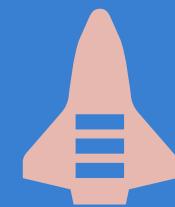
**05**

Graph  
Databases



**06**

Games  
TakeAways



# Document-Oriented Database



## Model: Structured Objects identified by a key

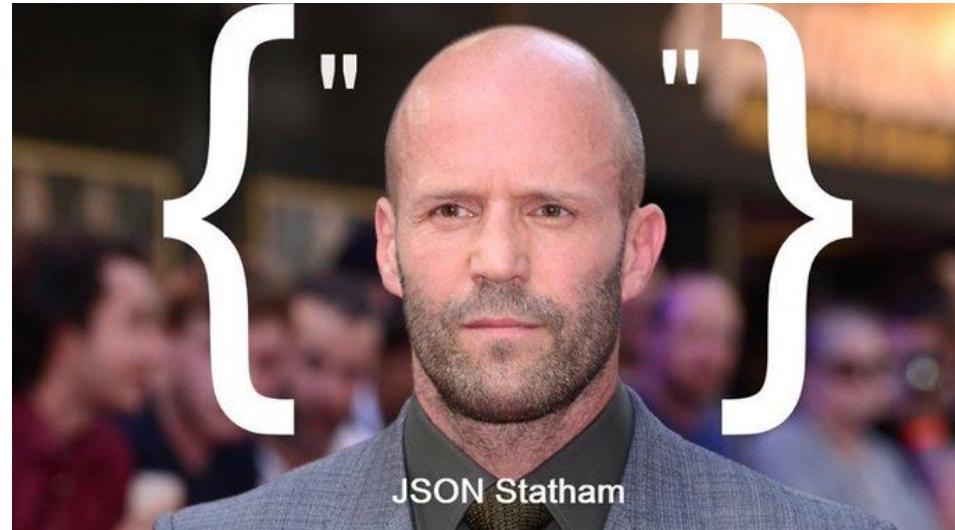
- Documents are structured data but with no schema
- Multiple format but mostly JSON
- Group of documents of same nature as “collections”

## Queries

- Request by the key
- Request on other fields tag/path in the document

## Use Cases

- Mainly reads, less writes
- Document storage with a structure but no schema
- Used in FrontEnd development matching the JSON used



mongoDB



Couchbase

# Document Shredding

```
create table <name> (
    key text,
    p0 text,
    ... p[N] text,
    bool_value boolean,
    txt_value text, d
    bl_value double, leaf text
)
```

# Document Shredding

```
{"a": { "b": 1 }, "c": 2}
```

The document would be “shredded” into rows looking like this:

key	p0	p1	dbl_value	
x	a	b	1	46
x	c	null	2	

# Document Shredding

For data with an array, such as:

```
{"a": { "b": 1 }, "c": [{"d": 2}]}
```

there would be two rows, like so:

key	p0	p1	p2	dbl_value	
x	a	b	null	1	<a href="#">47</a>
x	c	[0]	d	2	

# HandsOn #3 Documents DB



DataStax

Astra

**Get your instance here:**

- <http://astra.new/intro-nosql>



GitHub

**Repository:**

- <https://github.com/datastaxdevs/workshop-introduction-to-nosql>



# Agenda

**01**

Definitions and  
objectives of NoSQL

**02**

Tabular  
Databases



**03**

Document  
Databases



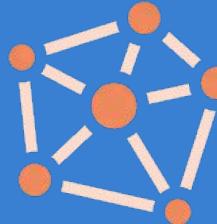
**04**



Key/values  
Databases

**05**

Graph  
Databases

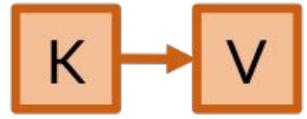


**06**

Games  
TakeAways



# Key Value Database



**Model:** Like a distributed HashTable

- One key, one value
- Keys are hashed into buckets (partitions)
- Similar to tabular but with a single value

**Queries**

- GET/PUT/DELETE/UPDATE direct CRUD only
- Value can be a single valued lists

**Use Cases**

- Distributed Cache !
- User cache Data, User Sessions
- data Deduplications

Key	Value
K1	AAA,BBB,CCC
K2	AAA,BBB
K3	AAA,DDD
K4	AAA,2,01/01/2015
K5	3,ZZZ,5623



# HandsOn #4 Key-Value DB



DataStax

Astra

**Get your instance here:**

- <http://astra.new/intro-nosql>



GitHub

**Repository:**

- <https://github.com/datastaxdevs/workshop-introduction-to-nosql>



# Agenda

**01**

Definitions and  
objectives of NoSQL

**02**

Tabular  
Databases



**03**

Document  
Databases



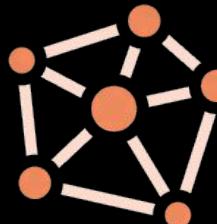
**04**



Key/values  
Databases

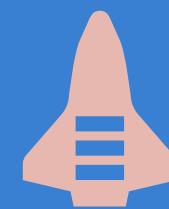
**05**

Graph  
Databases

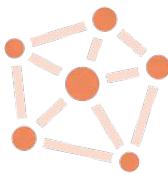


**06**

Games  
TakeAways



# Graph Database Database



## Model: Store Vertices and Edges data structured

- Data is represented as a Graph (Vertices & Edges)
- Dedicated to highly connected dataset (lot of “Joins”)
- Discovering simple and complex relationships between objects.

## Queries

- Find data based on filters on attributes for both nodes and edges
- Traversal following edges (cf gremlin)

## Use Cases

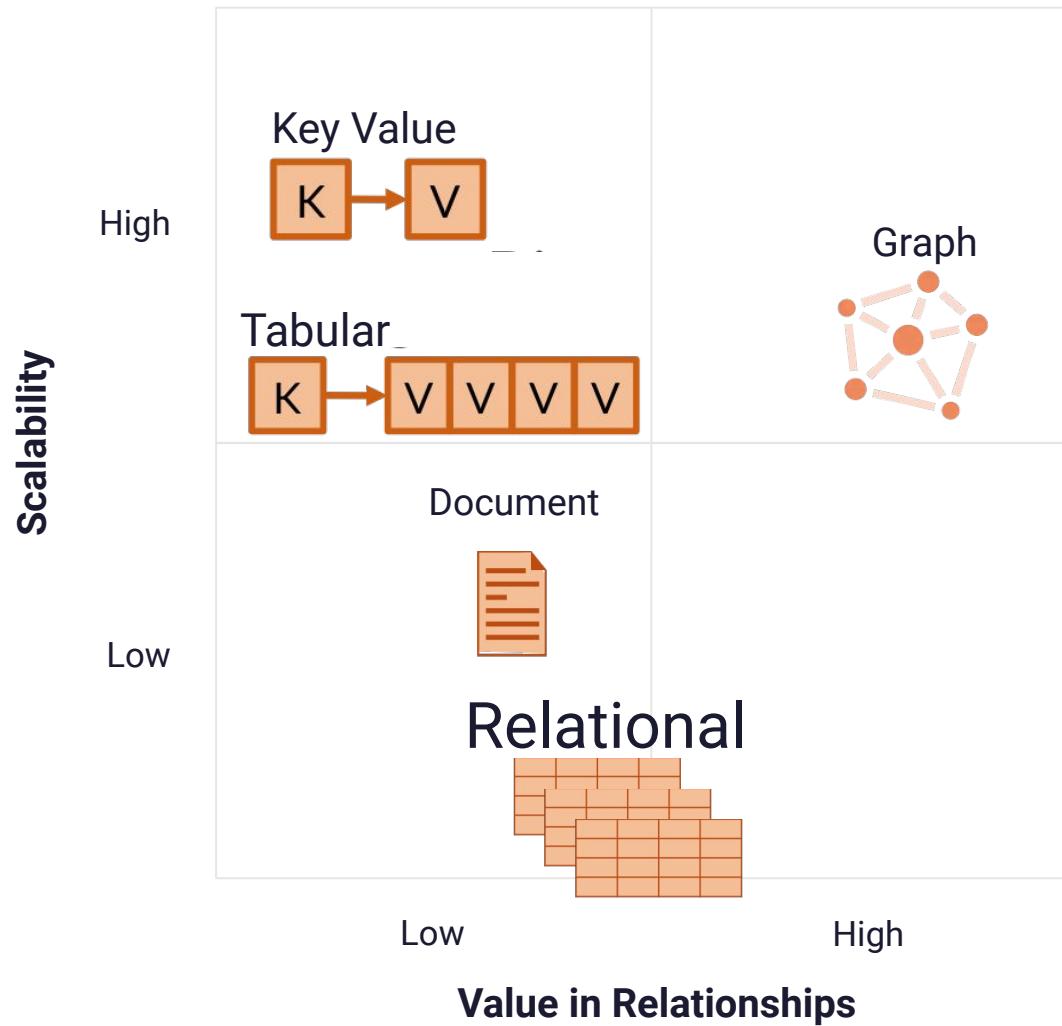
- Social Network, Customer 360
- Internet of Things
- Personalization and recommendation
- Health Care, Path finding, Security, Fraud Detection



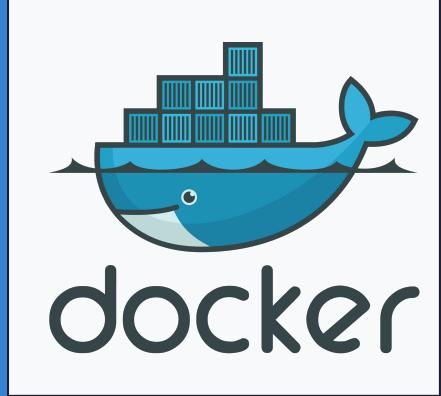
```
// What are the names of projects that were created by two friends?  
g.V().match(  
    as("a").out("knows").as("b"),  
    as("a").out("created").as("c"),  
    as("b").out("created").as("c"),  
    as("c").in("created").count().is(2)).  
    select("c").by("name")
```



# Positioning graphs Scalability and flexibility



# HandsOn #5 Graph Databases



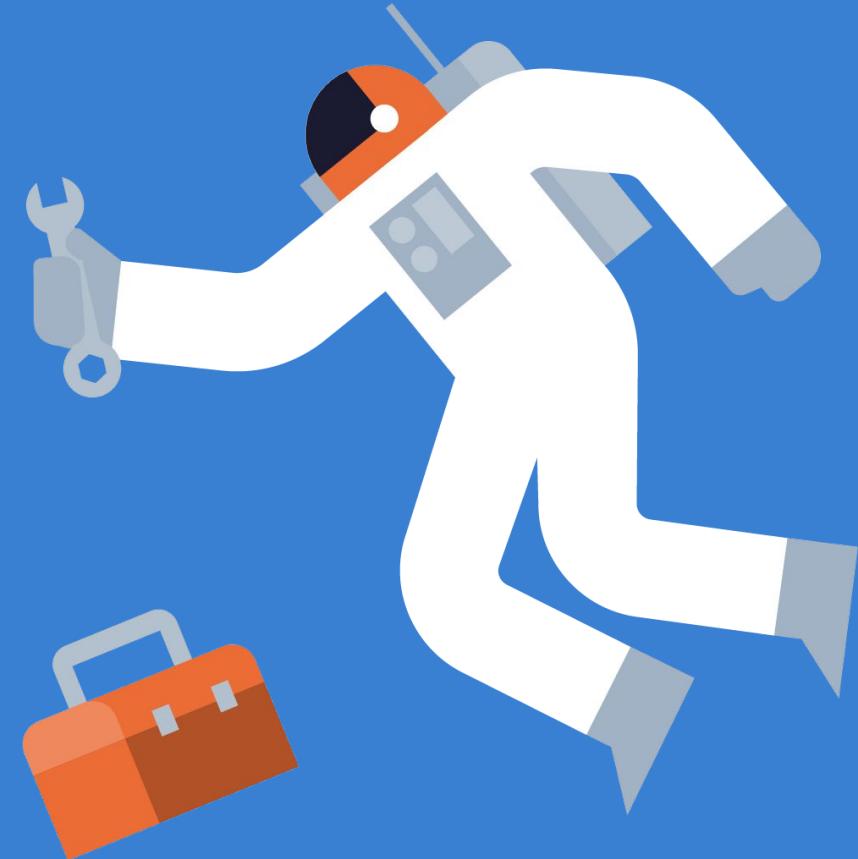
**Docker and Compose**



**GitHub**

**Repository:**

- <https://github.com/datastaxdevs/workshop-introduction-to-nosql>



# Agenda

**01**

Definitions and  
objectives of NoSQL

**02**

Tabular  
Databases



**03**

Document  
Databases



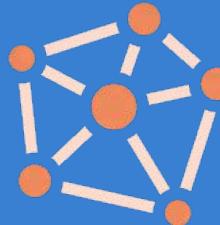
**04**



Key/values  
Databases

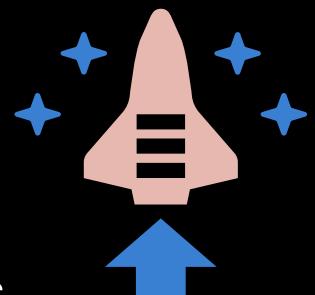
**05**

Graph  
Databases



**06**

Games  
TakeAways



# menti.com



Go to [www.menti.com](http://www.menti.com) and use the code 3491 9972

## Inequality predicates are allowed on ...

A bar chart titled "Inequality predicates are allowed on ...". The x-axis categories are "All table columns", "Partition key columns", "clustering key columns", and "No Inequality predicates are allowed". The y-axis ranges from 1 to 15. A green checkmark is above the "clustering key columns" bar at value 15. Red X's are above the other three bars at values 4, 3, and 1 respectively. The bars are colored light blue, yellow, green, and pink.

4 All table columns  
3 Partition key columns  
15 clustering key columns  
1 No Inequality predicates are allowed

2:10:19 / 2:26:05

Go to [www.menti.com](http://www.menti.com) and use the code 3491 9972

## Leaderboard

4821 p	spanda
4820 p	Agent X9
4775 p	fastest
4711 p	Sam
4468 p	CCedrickThePresenter
4371 p	shubham
3895 p	aaa
3877 p	vignesh
3861 p	adry
3812 p	Millie
	Puggie

2:11:07 / 2:26:05

DataStax

# Developer Resources



Subscribe

## LEARN

New hands-on learning at [www.datastax.com/dev](http://www.datastax.com/dev)  
Classic courses available at DataStax Academy

## ASK/SHARE

Join [community.datastax.com](http://community.datastax.com)  
Ask/answer community user questions – share your expertise

## CONNECT

Follow us [@DataStaxDevs](#)  
We are on Youtube – Twitter – Twitch!

## MATERIALS

Slides and exercises for this workshop are available at  
<https://github.com/DataStax-Academy/workshop-crud-with-python-and-node>

# Homework ([datastax.com/dev](https://datastax.com/dev))



## HOMEWORK

### [Complete hands-on #5]

- Install docker, play with the notebooks and show us some screenshots.

[Try another content] – <https://www.datastax.com/try-it-out>

- Go to `datastax/dev` and use the try-it-out



The screenshot shows the DataStax Developers homepage. At the top, there's a navigation bar with links for Products, Success, Learn, and Try For Free, along with a search icon. Below the navigation is a banner with the text "DATASTAX FOR DEVELOPERS" and "Learn How to Succeed with Apache Cassandra™". To the right of the banner is a large graphic of a person climbing a ladder with the words "LEVEL UP". Below the banner, there are three main sections: "Try It Out", "What is Cloud Native?", and "What is Cassandra?". The "Try It Out" section is highlighted with a red box. It features an icon of a hand writing on a screen, the text "Try It Out", and the subtext "Write your first Cassandra query and get your hands on the technology." It includes a "Try Out Cassandra" button. The other two sections also have their own icons and descriptive text.

# Certifications

<https://www.datastax.com/dev/certifications>



A promotional card with a dark blue header and footer and a white central section. The header features the DataStax eye logo and the word 'cassandra'. The footer features the Kubernetes logo and the word 'kubernetes'. In the center, the text 'COMING SOON! Apache Cassandra Operations in Kubernetes Certification' is displayed in yellow and white. Below this, a paragraph explains the program's purpose: 'As teams work to containerize and deploy applications using Kubernetes, there's increasing interest in running Cassandra in Kubernetes alongside applications as well. We're developing a new certification program to help teams level up their skills to run Cassandra successfully in cloud-native deployments.' It also mentions that the certification will cover topics like running Cassandra in Docker containers, understanding how Cassandra maps to Kubernetes, and deploying Cassandra on Kubernetes using Kubernetes operators and monitoring tools. At the bottom, a button says 'SIGN UP NOW'.

Vouchers (145\$ each), valid 3 months, with 2 attempts will be given to people who apply and register to the 3 episodes.

# Weekly Workshops

<https://www.datastax.com/workshops>

The image shows a composite view of the DataStax Developers YouTube channel and its website.

**YouTube Channel:** The left side displays the official DataStax Developers YouTube channel page. It features a large banner with the text "LEVEL UP with the DataStax Developers". Below the banner, there's a "DataStax Developers" profile picture, the channel name, and 8,1 k abonnés. The main video grid shows various workshop recordings, each with a thumbnail, title, and duration. For example, one video titled "Building Microservices with Cassandra + Spring" has a duration of 2:23:56 and 1,1 k vues. Another video titled "Advanced Data Modeling in Apache Cassandra™" has a duration of 2:41:51 and 1,3 k vues.

**Website:** The right side shows the DataStax Developers website. At the top, there's a "SUBSCRIBE" button and a call-to-action "for weekly content on building". Below this, a section titled "Upcoming Live Events" lists four workshops:

- Apache Cassandra™ Certification Preparation**: Multiple Dates | NoSQL | Beginner. Includes a "Register Now" button.
- Build Microservices with Apache Cassandra™!**: Feb 17 or Feb 18 | NoSQL | Beginner. Includes a "Register Now" button.
- Certification Exam Preparation Workshop**: MULTIPLE DATES. Includes a "Register Now" button.
- Cloud-Native Workshop: Build Spring Microservices with Apache Cassandra™**: MULTIPLE DATES. Includes a "Register Now" button.
- Learn how to build a Serverless Game!**: Feb 24 or Feb 25 | Game Development | Beginner. Includes a "START" button and a "Register Now" button.
- Build Microservices with Cassandra & Quarkus**: March 11 | Microservices | Beginner. Includes a "Register Now" button.

# Next Week:

#Frontend2021 – Build a Netflix clone  
with GraphQL, React and a NoSQL DB!

LIVE hands-on workshop </>

Building a  
**NETFLIX Clone**

June 9 or June 10 | Application Development



Register here: [dtsx.io/workshop-netflix-clone](https://dtsx.io/workshop-netflix-clone)

# Join our 10k Discord Community

## The Fellowship of the RINGS

<https://dtsx.io/discord>

A screenshot of the DataStax Developers Discord server's main chat room. The interface shows a sidebar with various channels like 'server-conduct', 'upcoming-events', 'useful-resources', 'moderator-only', 'WORKSHOPS' (which contains 'main-chat-room' and several breakout rooms), and 'TOPICS' (including '# grafana-cassandra-data...' and '# docker-training'). The main channel 'main-chat-room' is active, displaying messages from users like Jack Fryer and Cedrick Lunven. A large event announcement for 'Apache Cassandra™ meets Kubernetes!' is pinned at the top of the channel. On the right side, there's a sidebar for 'PRESENTER—3' featuring three individuals: Aleks Volochnev, David Jones-Gilardi, and jscarp. Below that is a section for 'HELPER—1' with John Sanda. The bottom right shows a list of 227 active users, including Abhiprada, Absurdism, hiya, Adalberto, aditya\_dhunna, adnaneCord, Adrigunz, Aemilius Gaurav, Aemilius gaurav, Aguvas, ajscilingo, and akashTaxvisor. At the bottom of the screen, there's a message input field and some reaction icons.

# DataStax Developers

## Thank you!



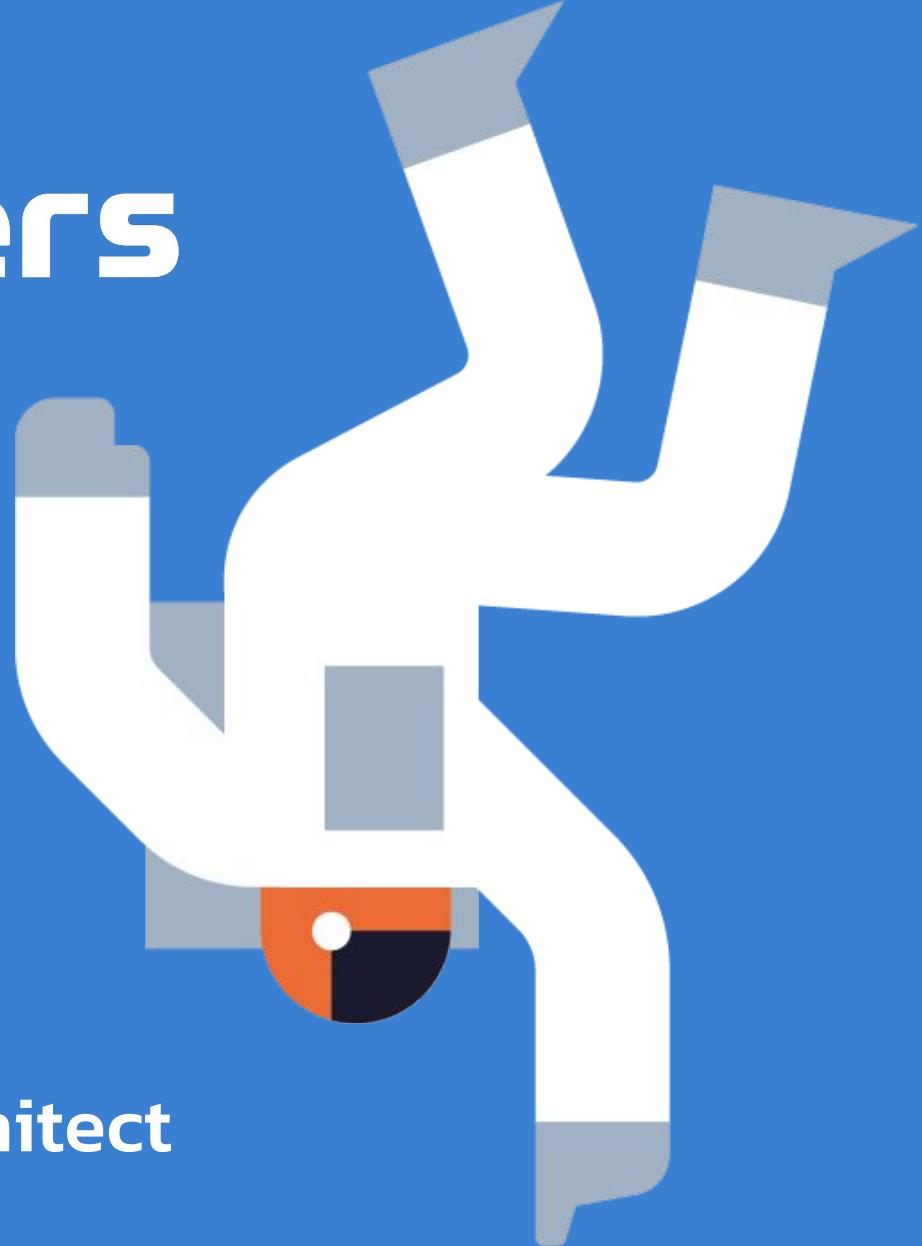
@hadesarchitect  
@clun  
@SonicDMG



@hadesarchitect  
@clunven  
@SonicDMG



@hadesarchitect  
@clunven  
@david-gilardi



# DataStax Developers

Thank you!



Subscribe

