

DataStax

Developers

Introduction to NoSQL Databases

Building efficient applications with Apache Cassandra





Cedrick
Lunven

Aleksandr
Volochnev

Jack
Fryer

Kirsten
Hunter

Stefano
Lottini

David
Gilardi

Ryan
Welford

Rags
Srinivas

Sonia
Siganporia

R

S



DataStax Developers Crew

Senior Developer Advocate



STARGATE



david-gilardi



SonicDMG



SonicDMG

- Training
- Teaching
- Support
- Public Speaking
- Reference Applications
- Cassandra Certified
- *Cough (Oracle 8 certified)
- Happy developer for 25 years
- Coded in a ton, but Java/js mainstays
- Advanced User



David Jones-Gilardi



Developer Advocate



STARGATE



{ REST }



ryanwelford



RyanWelford

- Front-End dev
- Training
- Teaching
- Support

- Happy developer for 5 years
- Love developer awesome user experiences



Ryan Welford

01



Intro to Bootcamp 2022
HouseKeeping

02

Intro to NoSQL Databases
Power of Purpose-built DB

03

Intro to Cassandra
Peer-to-Peer Database

04

CAP Theorem
Law of Distributed Systems

05

Tables and Partitions
How do you structure data

06

What's next?
Quiz, Homework, Next week



Agenda



WEEK 1

January 5th - January 11th



WEEK 2

January 12th - January 18th



WEEK 3

January 19th-January 25th



WEEK 4

January 26th-February 1st



Introducing 2022 Bootcamp



#1 Introduction to NoSQL Databases and Apache Cassandra™

- ~2H - Live lessons (same but different timezones)
 - January 5th 6 PM CET (= 9am PST = 11 am CST = 12 pm EST)
 - January 6th 9 AM CET (= 12am PST = 1am CST = 3am EST)
- ~3H - Homeworks
 - Hands-on with DBaaS Cassandra service AstraDB
 - Learning materials: Cassandra Fundamentals exercises 1 to 8
 - Fill the week 1 test



WEEK 2
January 12h - January 18th

WEEK 3
January 19th-January 25th



WEEK 4
January 26th-February 1st



Introducing 2022 Bootcamp

#1 Introduction to NoSQL Databases and Apache Cassandra™

- ~2H - Live lessons (same but different timezones)
 - January 5th 6 PM CET (= 9am PST = 11 am CST = 12 pm EST)
 - January 6th 9 AM CET (= 12am PST = 1am CST = 3am EST)
- ~3H - Homeworks
 - Hands-on with DBaaS Cassandra service AstraDB
 - Learning materials: Cassandra Fundamentals exercises 1 to 8
 - Fill the week 1 test



WEEK 3

January 19th-January 25th

#2 Design Cassandra Data Models for Full-Stack Applications

- ~2H - Live lessons (twice the same for different timezones)
 - January 12th 6 PM CET (= 9am PST = 11 am CST = 12 pm EST)
 - January 13th 9 AM CET (= 12am PST = 1am CST = 3am EST)
- ~3H - Homeworks
 - Hands-on with DBaaS Cassandra service AstraDB
 - Learning materials: Cassandra Fundamentals exercises 9 to 14
 - Learning materials: DS220 Modules 1 to 20 (Table Features)
 - Fill the week 2 test



WEEK 4

January 26th-February 1st



Introducing 2022 Bootcamp

#1 Introduction to NoSQL Databases and Apache Cassandra™

- ~2H - Live lessons (same but different timezones)
 - January 5th 6 PM CET (= 9am PST = 11 am CST = 12 pm EST)
 - January 6th 9 AM CET (= 12am PST = 1am CST = 3am EST)
- ~3H - Homeworks
 - Hands-on with DBaaS Cassandra service AstraDB
 - Learning materials: Cassandra Fundamentals exercises 1 to 8
 - Fill the week 1 test

#2 Design Cassandra Data Models for Full-Stack Applications

- ~2H - Live lessons (twice the same for different timezones)
 - January 12th 6 PM CET (= 9am PST = 11 am CST = 12 pm EST)
 - January 13th 9 AM CET (= 12am PST = 1am CST = 3am EST)
- ~3H - Homeworks
 - Hands-on with DBaaS Cassandra service AstraDB
 - Learning materials: Cassandra Fundamentals exercises 9 to 14
 - Learning materials: DS220 Modules 1 to 20 (Table Features)
 - Fill the week 2 test

#3 Build a Full-Stack Backend with a NoSQL database

- ~2H - Live lessons (same but different timezones)
 - January 19th 6 PM CET (= 9am PST = 11 am CST = 12 pm EST)
 - January 20th 9 AM CET (= 12am PST = 1am CST = 3am EST)
- ~3H - Homeworks
 - Coding with DBaaS Cassandra service AstraDB
 - DS220 Modules 21 to 40 (Collections-> Physical Data Model)
 - Fill the week 3 test

WEEK 4
January 26th-February 1st



#1 Introduction to NoSQL Databases and Apache Cassandra™

- ~2H - Live lessons (same but different timezones)
 - January 5th 6 PM CET (= 9am PST = 11 am CST = 12 pm EST)
 - January 6th 9 AM CET (= 12am PST = 1am CST = 3am EST)
- ~3H - Homeworks
 - Hands-on with DBaaS Cassandra service AstraDB
 - Learning materials: Cassandra Fundamentals exercises 1 to 8
 - Fill the week 1 test

#2 Design Cassandra Data Models for Full-Stack Applications

- ~2H - Live lessons (twice the same for different timezones)
 - January 12th 6 PM CET (= 9am PST = 11 am CST = 12 pm EST)
 - January 13th 9 AM CET (= 12am PST = 1am CST = 3am EST)
- ~3H - Homeworks
 - Hands-on with DBaaS Cassandra service AstraDB
 - Learning materials: Cassandra Fundamentals exercises 9 to 14
 - Learning materials: DS220 Modules 1 to 20 (Table Features)
 - Fill the week 2 test

#3 Build a Full-Stack Backend with a NoSQL database

- ~2H - Live lessons (same but different timezones)
 - January 19th 6 PM CET (= 9am PST = 11 am CST = 12 pm EST)
 - January 20th 9 AM CET (= 12am PST = 1am CST = 3am EST)
- ~3H - Homeworks
 - Coding with DBaaS Cassandra service AstraDB
 - DS220 Modules 21 to 40 (Collections-> Physical Data Model)
 - Fill the week 3 test

#4 Build Api and Microservices for Apache Cassandra™

- ~2H - Live lessons (same but different timezones)
 - January 26th 6 PM CET (= 9am PST = 11 am CST = 12 pm EST)
 - January 27th 9 AM CET (= 12am PST = 1am CST = 3am EST)
- ~3H - Homeworks
 - Coding with DBaaS Cassandra service AstraDB
 - Learning materials: DS220 Modules 42 to 56
 - Fill the week 4 test



Introducing 2022 Bootcamp



1

Attend one of the 2 LIVE STREAMED workshop
(Wednesday or Thursday)



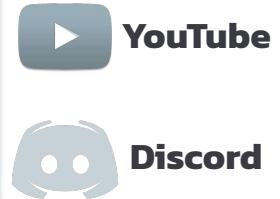
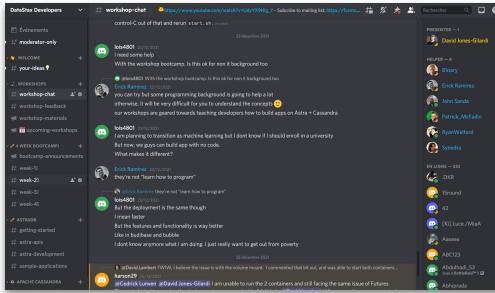
4-week bootcamp Housekeeping



Livestream: youtube.com/DataStaxDevs

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

Agenda



Games and quizzes: menti.com

How much experience do you have with the Spring Framework ?



- 1 Attend one of the 2 LIVE STREAMED workshop
(Wednesday or Thursday)
- 2 Complete the workshop labs



4-week bootcamp Housekeeping



Database + GraphQL + PlayGround



DataStax
Astra DB



The screenshot shows the Gitpod IDE interface. On the left is the Explorer view displaying project files like 'StargateDemoApplication.java', 'ManifestDemoApplication.java', and 'gitpod.Dockerfile'. The main area contains two code editors with Java code. Below the editors are tabs for 'PROBLEMS', 'OUTPUT', 'TERMINAL', and 'DEBUG CONSOLE'. The terminal shows the command 'gitpod /workspace/workspace-spring-stargate \$'. To the right of the code editors are logos for npm, node.js, Maven, and Docker. The bottom of the interface features a navigation bar with 'gitpod' and other options.

Gitpod

The screenshot shows a GitHub repository page for 'DataStax-Examples / todo-astra-jamstack-netlify'. The page displays a list of branches (master, 5 branches, 5 tags), pull requests, and code. It includes sections for 'Code', 'Actions', 'Projects', 'Wiki', 'Security', 'Insights', and 'Settings'. A 'GitHub Out' button is visible. The repository details show 177 commits ahead of 'tjke/master'. The 'Code' tab is selected, showing files like 'gh-pages', 'functions', 'public', 'src', 'env-template', 'gitignore', and 'gitpod.yml'. The 'Readme' section contains a link to 'astra.datastax.com/register'. The 'Releases' section indicates 'No releases published'.

GitHub

O'REILLY
Katacoda KATACODA OVERVIEW & SOLUTIONS

Connect to Astra (Cassandra as a Service) with CQL Shell

Step 1 of 5

Create your Astra DB Database

If you don't have an Astra account, set one up - it's easy!

Go to the Astra DB page in your browser astra.datastax.com.

Let's create the database. Follow the steps outlined here. To make it easy, we have recommended the values you should use for this scenario.

NOTE: If you already have an Astra DB database with values that differ from what we suggested, you may have to adapt some of the operations in this scenario accordingly, or create an additional response with the designated values.

If you don't already have an Astra DB database, when you log in the first time you'll see a screen that looks like the image below.

Choose Plan & Provider



Katacoda



2

Complete Workshops Labs

- 1** Attend one of the 2 LIVE STREAMED workshop
(Wednesday or Thursday)
- 2** Complete the workshop labs
- 3** Do your homeworks (learning materials)



4-week bootcamp Housekeeping

The screenshot shows the DataStax Academy course page for DS220. At the top, it says "Course Content" and "Introduction". Below that, there are three sections: "Data Modeling Overview", "Data Modeling Overview Quiz", and "Relational Vs. Apache Cassandra". Each section has a "Start" button. To the right, there's a progress bar showing "Not Started 0/56" and a circular progress indicator at 0%. Below the progress bar are sections for "Badges" and "Competencies".

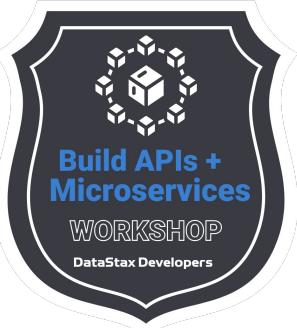
The screenshot shows the "Cassandra Fundamentals" learning series page. It features a "GET STARTED" button and a list of topics: 01 Introduction to Apache Cassandra™, 02 Cassandra Query Language, 03 Keyspaces and Data Replication Strategies, 04 Tables with Single-Row Partitions, and 05 Tables with Multi-Row Partitions. The background has a blue hexagonal pattern.

The screenshot shows a GitHub repository named "DataStax-Examples / todo-astra-jamstack-netlify". It displays a list of branches (master, 1.2, 1.3) and pull requests. A green "Group by Pull Request" button is highlighted. On the right, there's a "astra datastax.com/register" section with links to "net", "cassandra", "rest api", "dev", "database", "astra", and "building-sample-apps". Below that is a "Readme" section and a "Releases" section with a note about no releases published.

3

Do your homeworks

- 1** Attend one of the 2 LIVE STREAMED workshop (Wednesday or Thursday)
- 2** Complete the workshop labs
- 3** Do your homeworks (learning materials)
- 4** Pass the weekly test (google form)



Intro to NoSQL Homework

Welcome and thank you! Here you can submit your homework for the datastax developers "Intro to NoSQL" workshop. In case of any questions please contact organisers at <https://dtsx.io/aleks> or just send an email to aleksandr.volochnev@datastax.com

- Workshop materials: <https://github.com/datastaxdevs/workshop-introduction-to-nosql>
- Discord chat: <https://dtsx.io/discord>

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

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 *

Your email

4

Pass the Weekly Test

menti.com



Go to www.menti.com and use the code 3491 9972

Inequality predicates are allowed on ...

A bar chart titled "Inequality predicates are allowed on ...". The y-axis represents the count of inequality predicates, ranging from 1 to 15. The x-axis categories are "All table columns", "Partition key columns", "clustering key columns", and "No inequality predicates are allowed".

Column Type	Count
All table columns	4
Partition key columns	3
clustering key columns	15
No inequality predicates are allowed	1

Below the chart, there is a video player interface showing a video of a person speaking. The video player includes controls like play/pause, volume, and a progress bar indicating 2:10:19 / 2:26:05. The title of the video is "Big paycheck".

Go to www.menti.com and use the code 3491 9972

Leaderboard

User ID	User Name	Profile Icon
4821 p	spanda	Avatar icon
4820 p	Agent X9	Avatar icon
4775 p	Sam	Avatar icon
4711 p	CCedrickThePresenter	Avatar icon
4468 p	shubham	Avatar icon
4371 p	aaa	Avatar icon
3895 p	vignesh	Avatar icon
3877 p	adry	Avatar icon
3861 p	Millie	Avatar icon
3812 p	Puggie	Avatar icon

Below the leaderboard, there is a video player interface showing a video of a person speaking. The video player includes controls like play/pause, volume, and a progress bar indicating 2:11:07 / 2:26:05. The title of the video is "Big paycheck".

01



Intro to Bootcamp 2022
HouseKeeping

02

Intro to NoSQL Databases
Power of Purpose-built DB

03

Intro to Cassandra
Peer-to-Peer Database

04

CAP Theorem
Law of Distributed Systems

05

Tables and Partitions
How do you structure data

06

What's next?
Quiz, Homework, Next week



Agenda

Why NoSQL was invented?

Relational Databases dominated the database market for decades. Why change?



**Relational Database Management Systems weren't
ready to cope with the new data volume.**



**Relational Database Management
Systems weren't ready to cope with the
new performance requirements.**



Velocity



Relational Database Management
Systems weren't ready to cope with the requirements of data types and data relations.

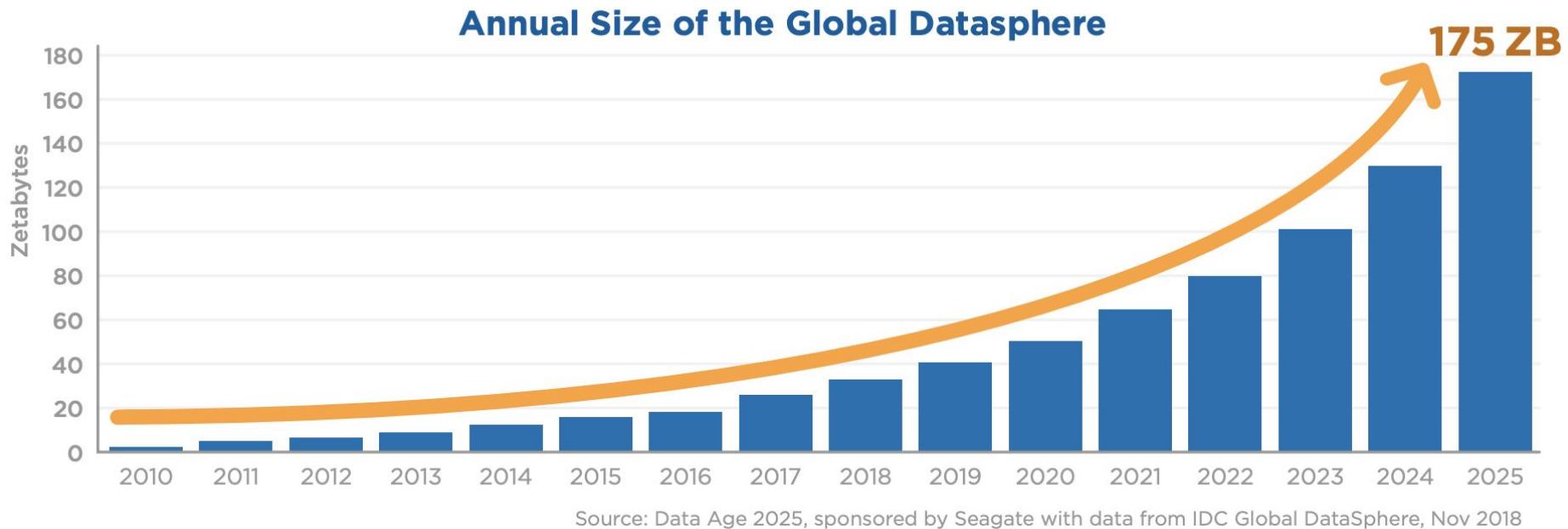
- Schemaless
- Relation as an entity
- Etc.



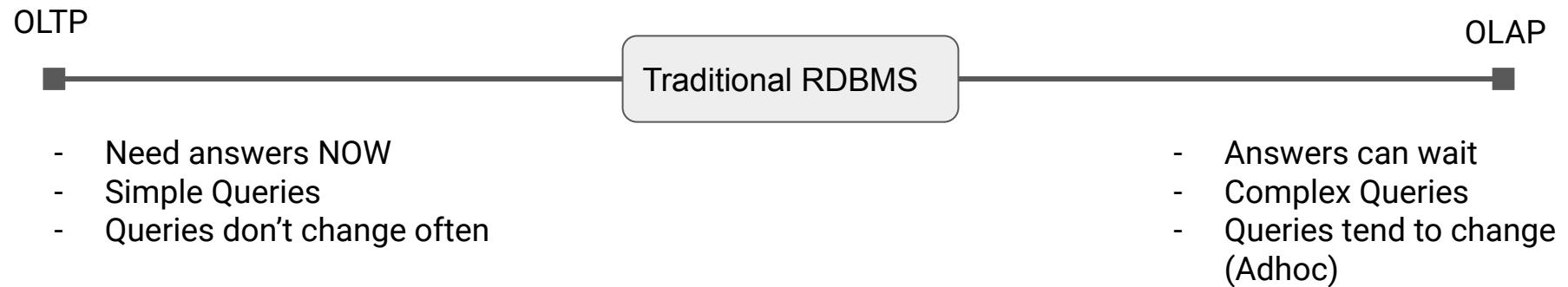
Variety



Requirements changed



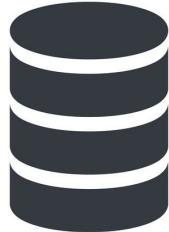
- OnLine Transaction Processing
- OnLine Analytical Processing



OLTP / OLAP



Relational Databases



Any database that supports SQL and based on the relational model of data

Not Only SQL

Time-Series Databases

Influx, OpenTSDB, Prometheus

Document Databases

MongoDB, Elastic, DocDb

Tabular Databases

Cassandra, Hbase, Bigtable

Ledger Databases

Amazon QLDB, ksql_db

Graph Databases

Neo4j, Datastax Graph, Titan

Key/Value Databases

Redis, Dynamo, Memcache



SQL to NoSQL

01



**Intro to Bootcamp 2022
HouseKeeping**

02

**Intro to NoSQL Databases
Power of Purpose-built DB**

03

**Intro to Cassandra
Peer-to-Peer Database**

04

**CAP Theorem
Law of Distributed Systems**

05

**Tables and Partitions
How do you structure data**

06

**What's next?
Quiz, Homework, Next week**



Agenda

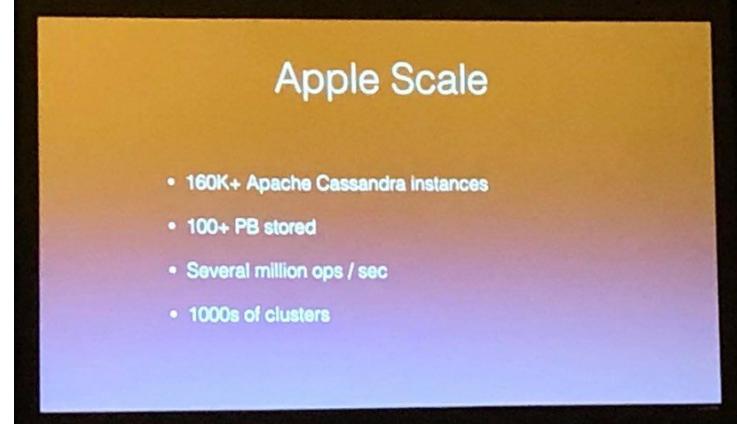
Apache Cassandra @ Netflix

- . 98% of streaming data is stored in Apache Cassandra
- . Data ranges from customer details to viewing history to billing and payments
- . Foundational datastore for serving millions of operations per second

- 30 million ops/sec on most active single cluster
- 500 TB most dense single cluster
- 9216 CPUs in biggest cluster

O(100) Clusters
O(10000) Instances
O(10,000,000) Replications per second
O(100,000,000) Operations per second
O(1,000,000,000,000,000) Petabytes of data

dtsx.io/cassandra-at-netflix



And many others...



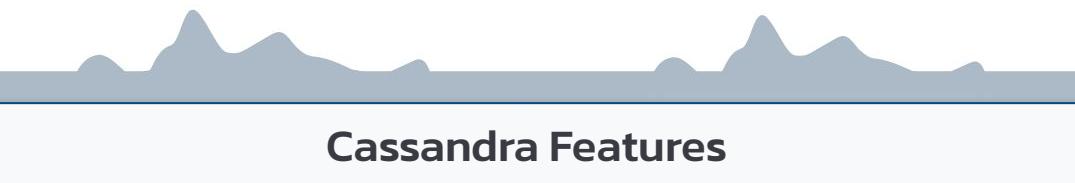
Cassandra Biggest Users (and Developers)

Apache Cassandra

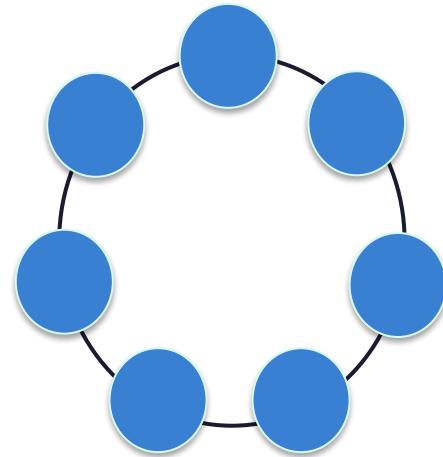
NoSQL Distributed Decentralised Database Management System



- Big Data Ready
- Read / Write Performance
- Linear Scalability
- Highest Availability
- Self-Healing and Automation
- Geographical Distribution
- Platform Agnostic
- Vendor Independent



Partitioning over distributed architecture makes the database capable to handle data of any size: we mean petabytes scale. Need more volume? Add more nodes.

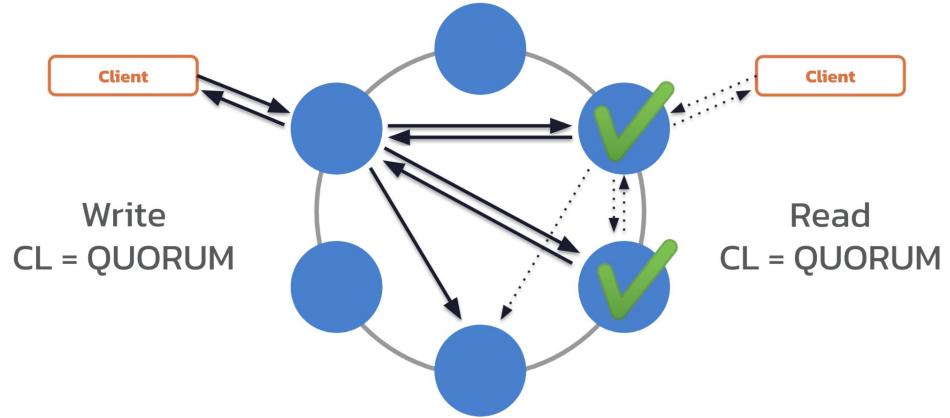


Big Data Ready



Even a single Cassandra node is very performant but a cluster consisting of multiple nodes and data centers brings throughput to the next level.

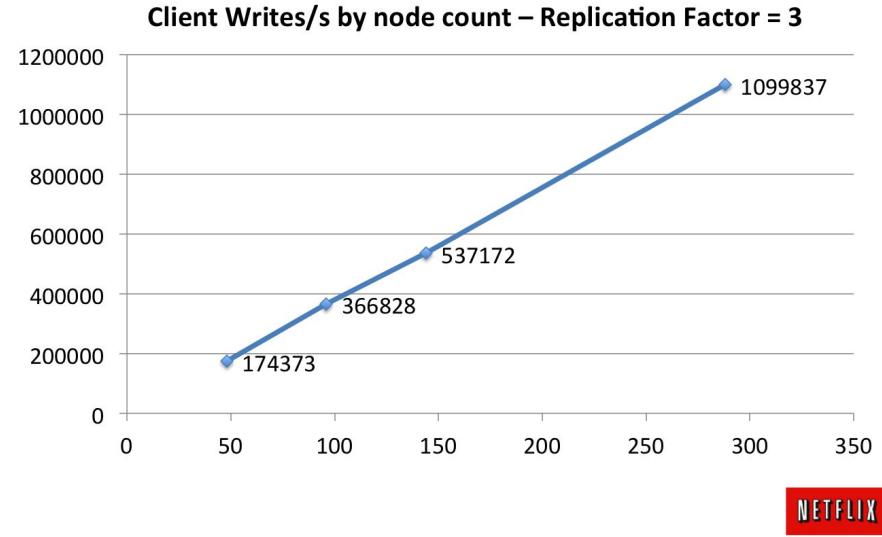
Decentralisation (masterless architecture) means that every node is able to deal with any request, read or write.



Read / Write Performance



- For volume or velocity, there are no limitations
- Linear - No overhead on new nodes, scales with your needs*

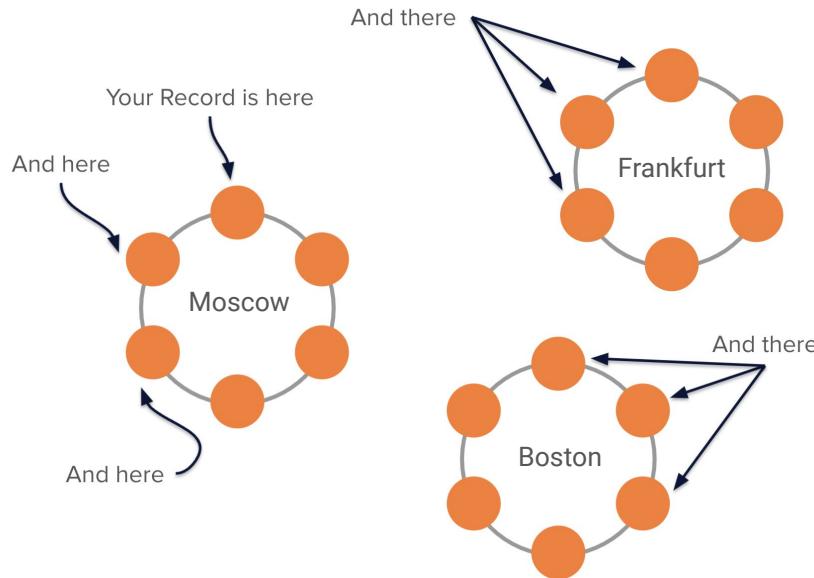


Linear Scalability



Replication, Decentralisation, and Topology-Aware Placement Strategy take care of possible downtimes:

- Multiple Live Replicas
- No Single Point of Failure
- Network topology-aware data placement
- Client-side Smart Reconnection and Strong Retry Mechanism



Highest Availability



Operations for a huge cluster can be exhausting so Apache Cassandra clusters are smart and able to scale, change data placement and recover automatically.

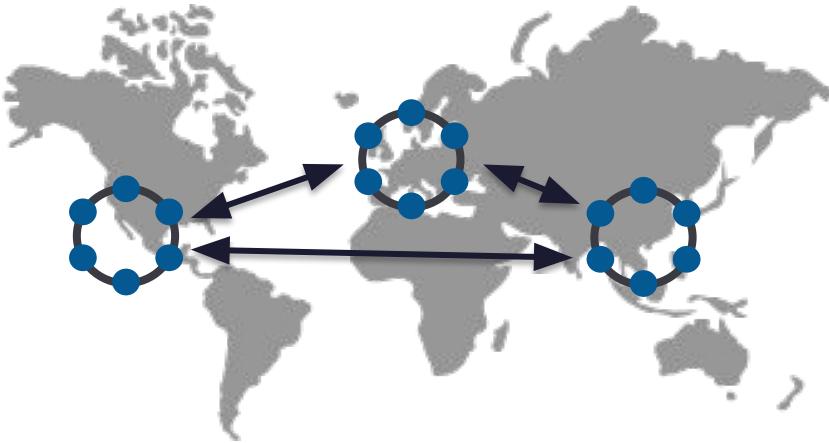


Self-Healing and Automation



Cassandra's trademark is multi-datacenter deployments, granting you an exceptional capability for disaster tolerance while keeping your data close to your clients - worldwide.

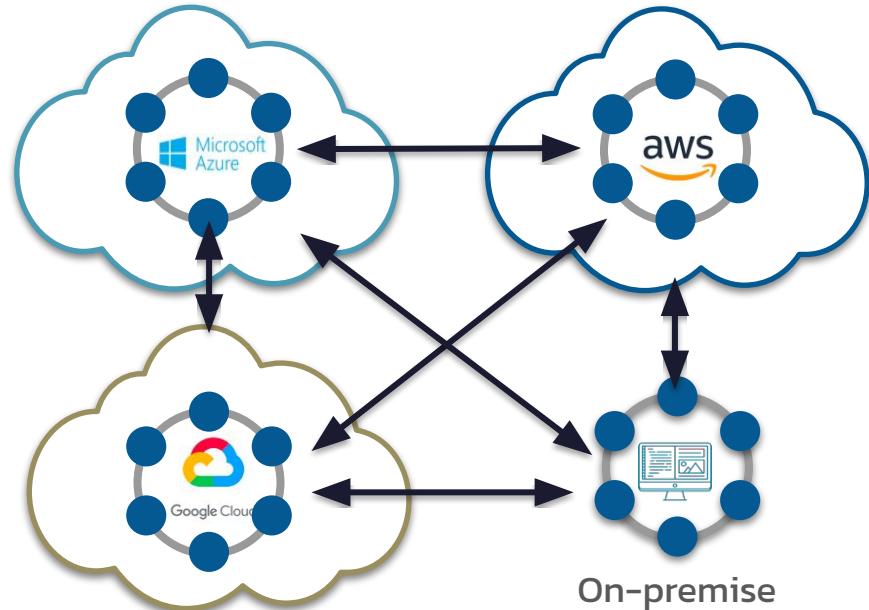
All DCs are active (available for both writes and reads)!



Geographical Distribution



Apache Cassandra is not bound to any platform or service provider, helping you build hybrid-cloud and multi-cloud solutions with ease.



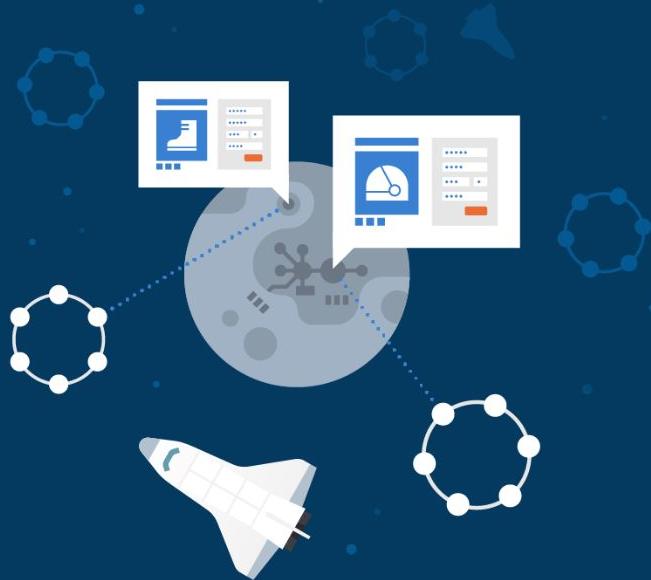
Platform Agnostic

Cassandra doesn't belong to any of commercial vendors but controlled by a non-profit Open Source **Apache Software Foundation**, already familiar to you by *Hadoop*, *Spark*, *Kafka*, *Zookeeper*, *Maven* and many other projects.



Vendor Independent





Looks cool, right?
How does it all work?

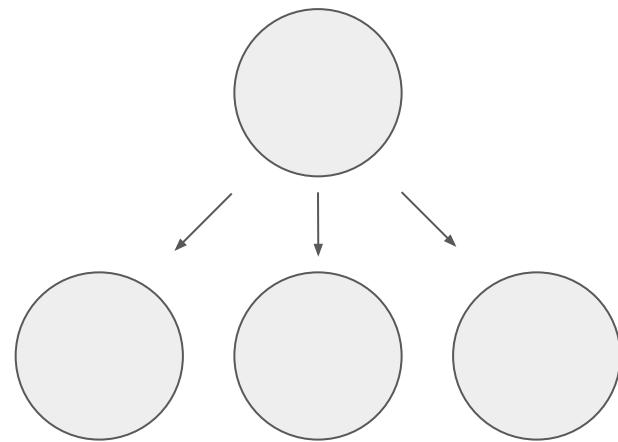


All servers are created equal

Cassandra Holy Book, P.I Paragraph I



Leader Server (Write/Read)

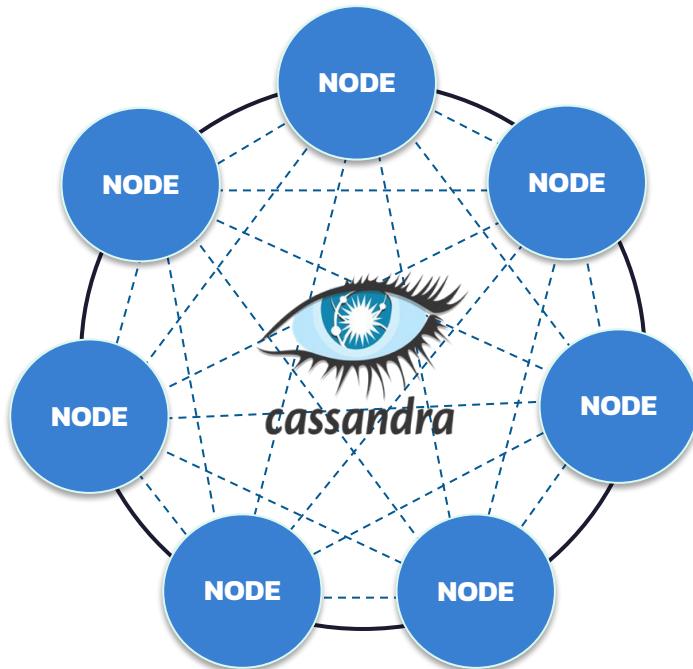


1. Single Point of Failure
2. Hard to scale for writes
3. Application needs to know where to write



Traditional Architecture





1. NO Single Point of Failure
2. Scales for writes and reads
3. Application can contact any node
(in case of failure - just contact next one)



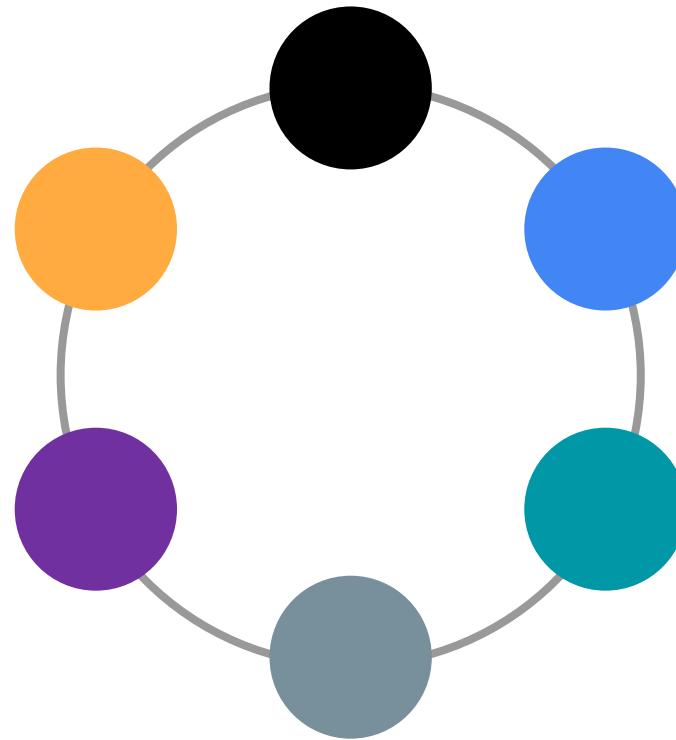
Master-less (Peer-to-Peer) Architecture

Data is Replicated



$RF = ?$

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



Replication Factor

```
CREATE KEYSPACE population  
    WITH REPLICATION = {  
        'class' : 'NetworkTopologyStrategy',  
        'us-west-1' : 3,  
        'eu-east-2' : 5  
    };
```

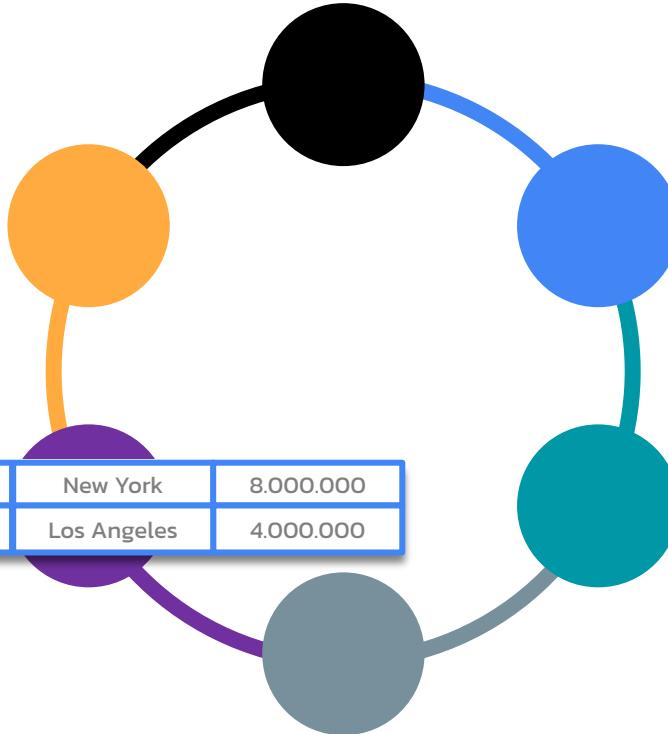
keyspace replication strategy



RF = 1

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

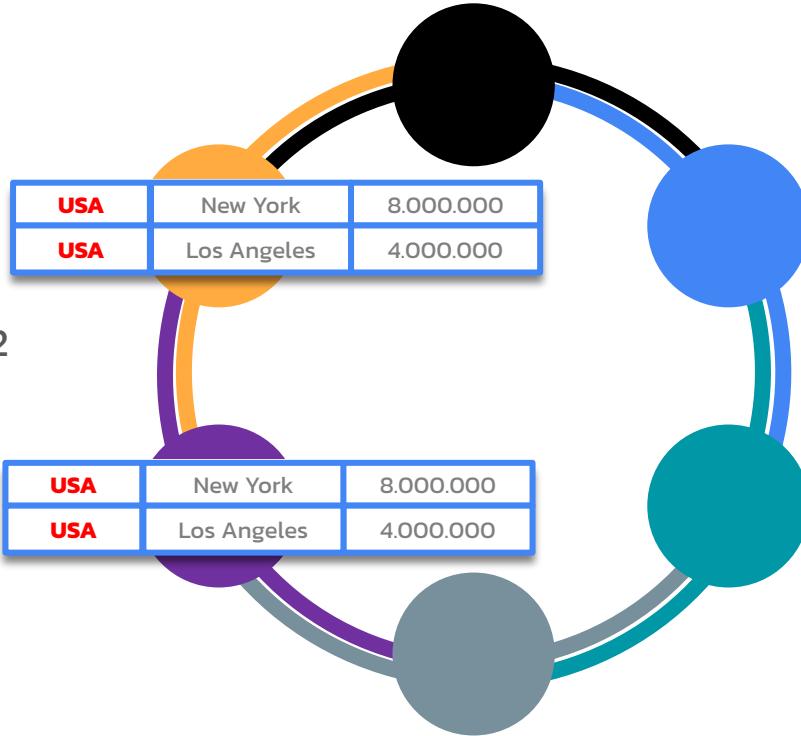
USA	New York	8.000.000
USA	Los Angeles	4.000.000



Replication Factor

RF = 2

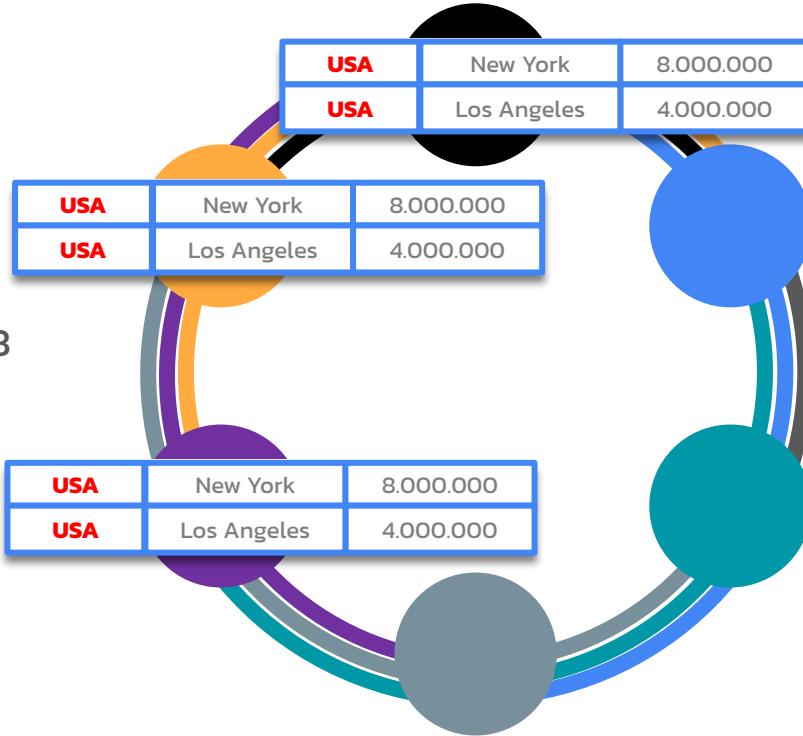
Replication Factor 2
means that every
partition is stored
on 2 nodes



Replication Factor

RF = 3

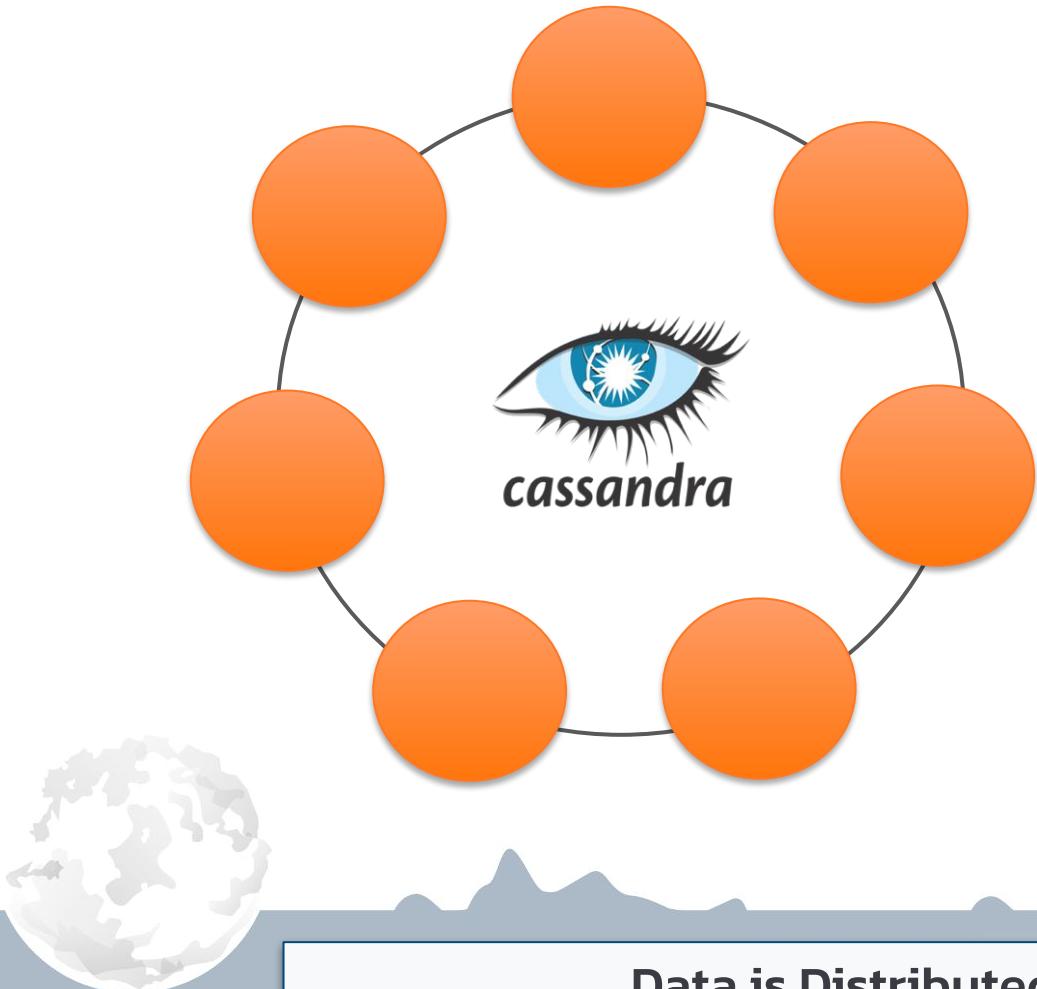
Replication Factor 3
means that every
partition is stored
on 3 nodes



Replication Factor

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

USA	New York	8.000.000
USA	Los Angeles	4.000.000

Country City Population

DE	Berlin	3.350.000
DE	Nuremberg	500.000

FR	Paris	2.230.000
FR	Toulouse	1.100.000

UK	London	9.200.000
-----------	--------	-----------

JP	Tokyo	37.430.000
-----------	-------	------------

AU	Sydney	4.900.000
IN	Mumbai	20.200.000

CA	Toronto	6.200.000
CA	Montreal	4.200.000



cassandra



Data is Distributed

```
Keyspace  
↓  
CREATE TABLE population.population_by_country_city (  
    country text,  
    city text,  
    population integer,  
    PRIMARY KEY ((country), city)  
);  
Table  
↓  
↑  
Partition key
```

Key-based Partitioning

Country	City	Population
USA	New York	8.000.000
USA	Los Angeles	4.000.000
CA	Toronto	6.200.000
CA	Montreal	4.200.000
DE	Berlin	3.350.000
DE	Nuremberg	500.000

Partition Keys

Partitioner
Murmur3 Hashing

Country	City	Population
59	New York	8.000.000
59	Los Angeles	4.000.000
12	Toronto	6.200.000
12	Montreal	4.200.000
45	Berlin	3.350.000
45	Nuremberg	500.000

Tokens

Cassandra Nodes

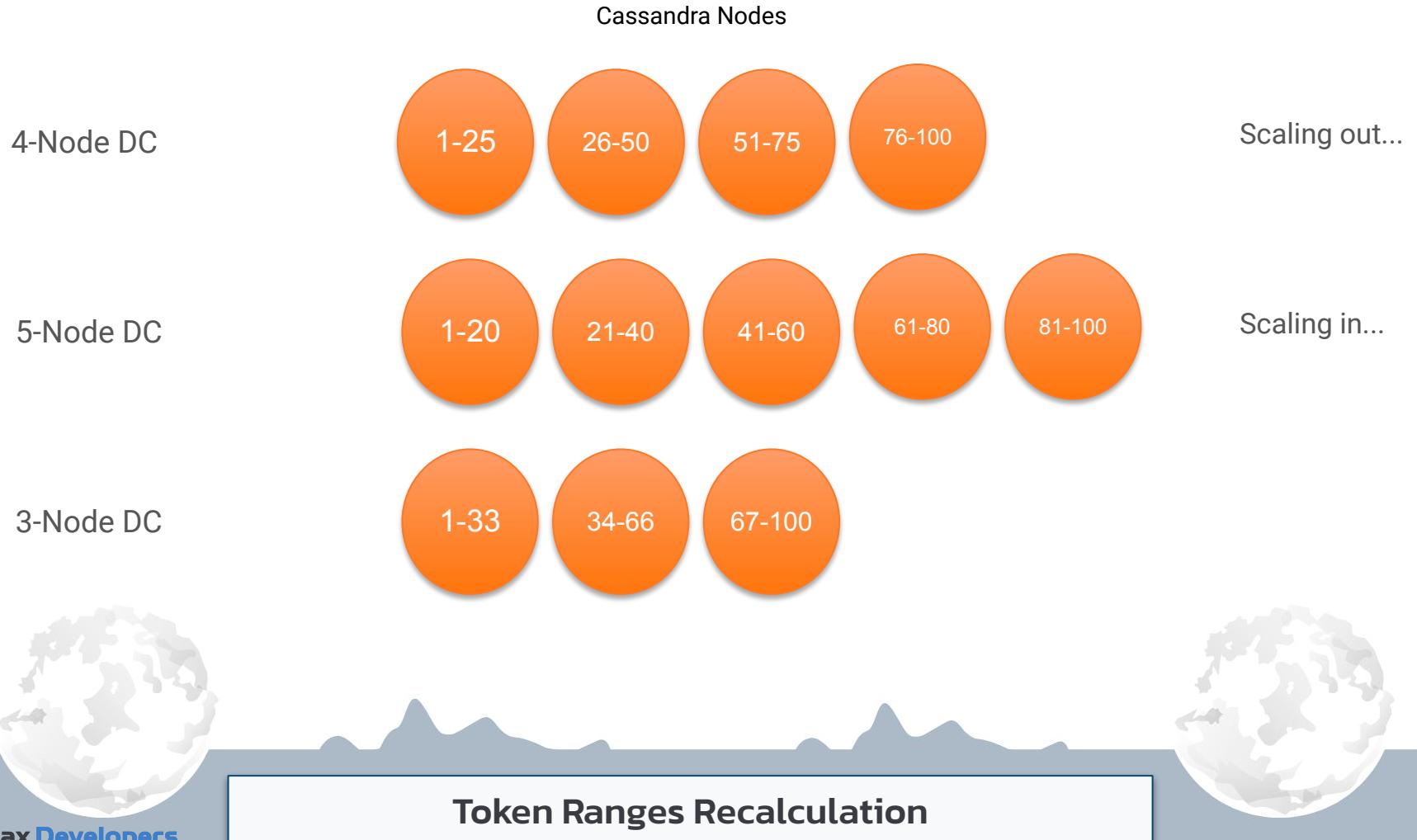


Partitioning and Token Ranges

Why partitioning?
Because scaling doesn't have to be [s]hard!

Big Data doesn't fit to a single server, splitting it into chunks we can easily spread them over dozens, hundreds or even thousands of servers, adding more if needed.





Replication Factor = 1

1-25

26-50

51-75

76-100

Replication Factor = 2

1-25,
26-50

26-50,
51-75

51-75,
76-100

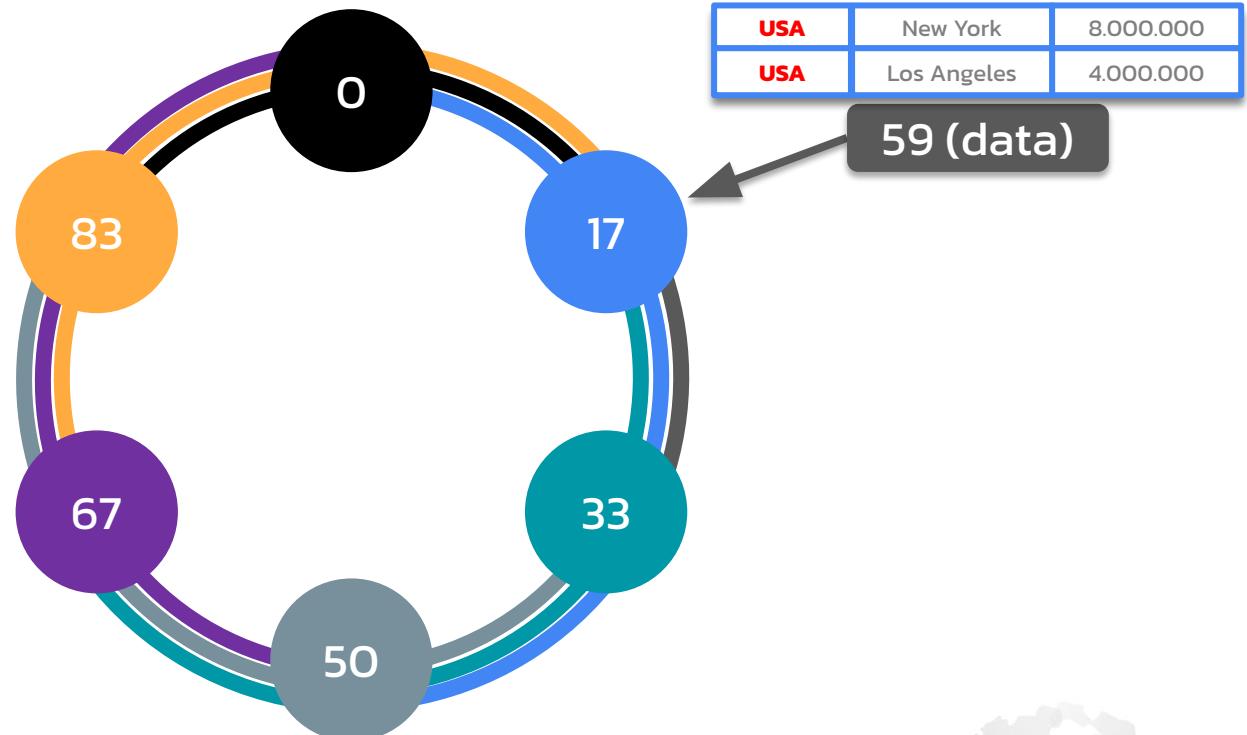
76-100,
1-25

And so on...



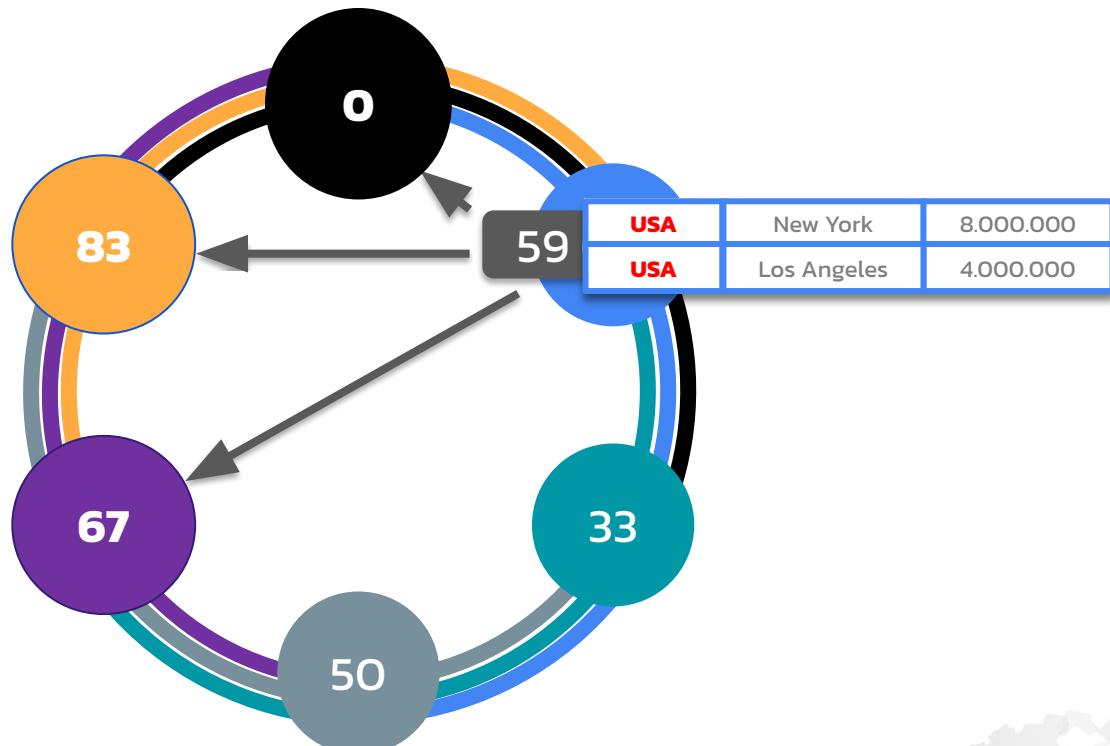
Partitioning + Replication

RF = 3



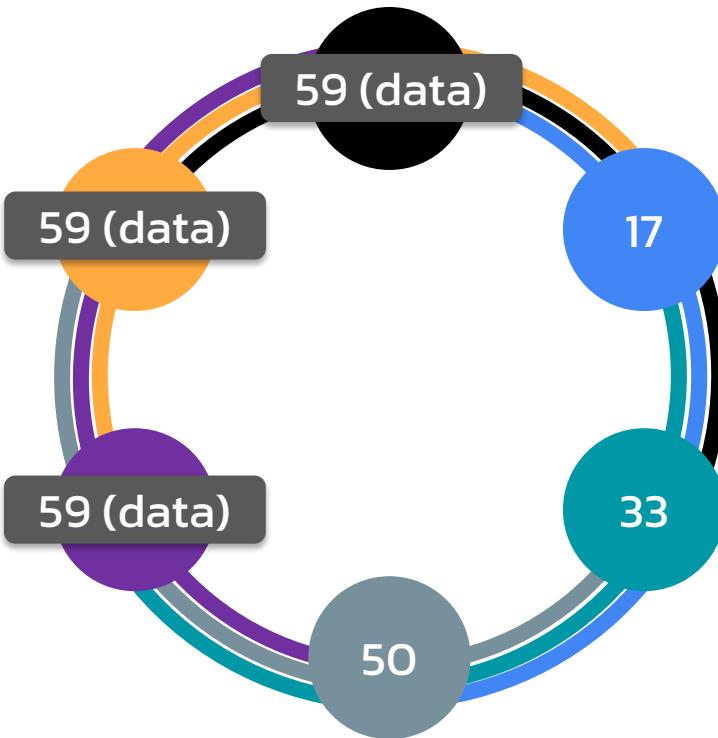
Partitioning + Replication

RF = 3



Partitioning + Replication

RF = 3



Partitioning + Replication

IMPORTANT

Each Cassandra node and even each Cassandra driver knows Data Allocation in a cluster (it's called Token-Aware), so your application can contact literally ANY server and still get the answer.

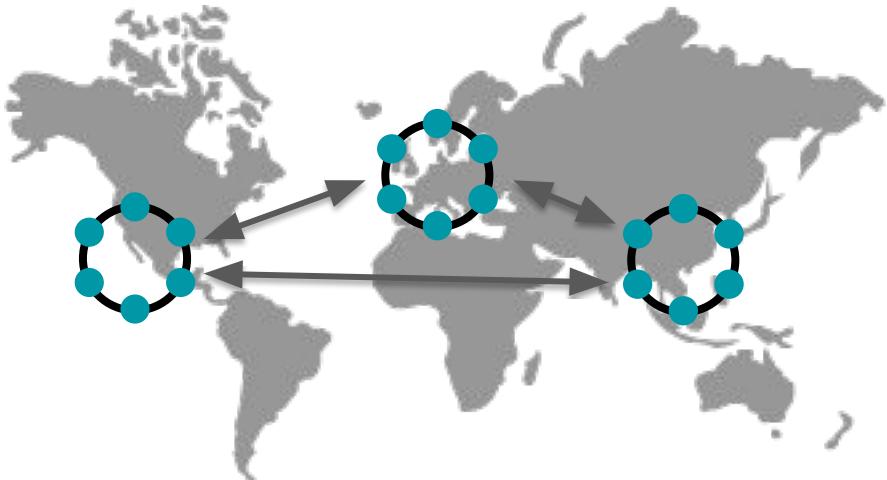




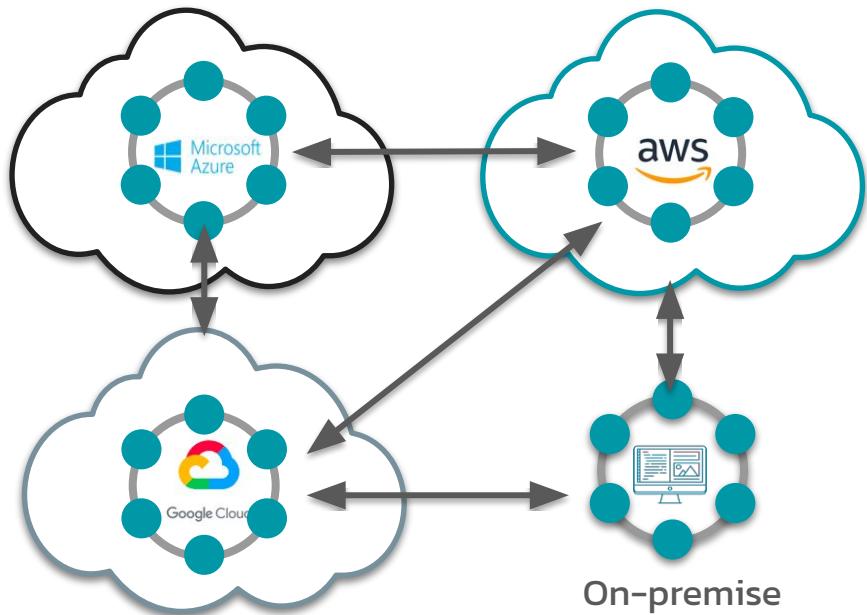
Data is GLOBALLY distributed



Geographical Distribution



Hybrid-Cloud and Multi-Cloud



On-premise

Data is globally distributed

01



**Intro to Bootcamp 2022
HouseKeeping**

02

**Intro to NoSQL Databases
Power of Purpose-built DB**

03

**Intro to Cassandra
Peer-to-Peer Database**

04

**CAP Theorem
Law of Distributed Systems**

05

**Tables and Partitions
How do you structure data**

06

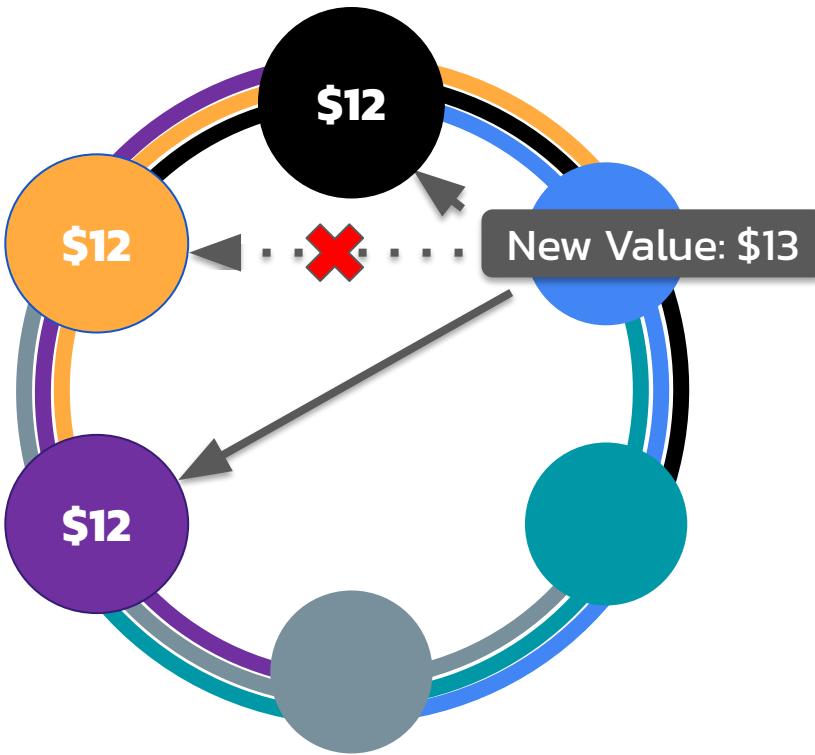
**What's next?
Quiz, Homework, Next week**



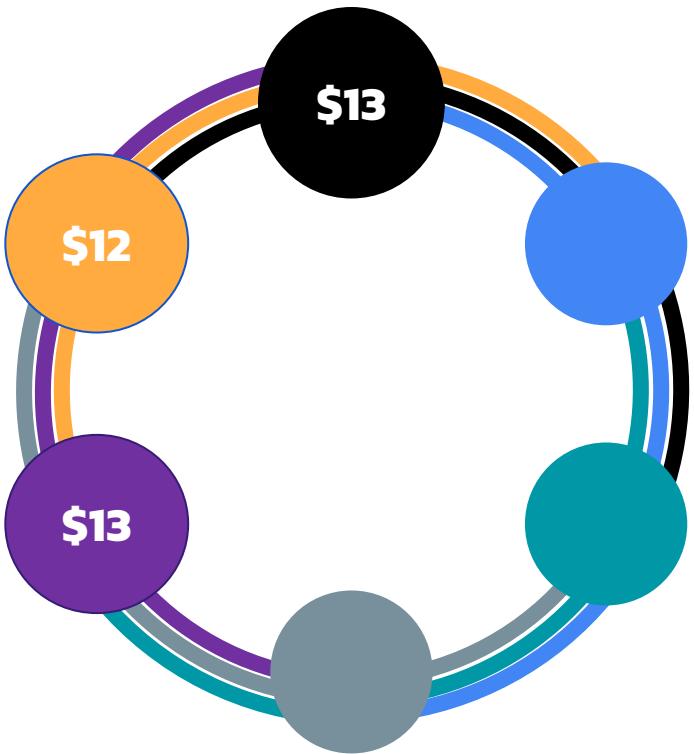
Agenda



What is the biggest problem of replication?

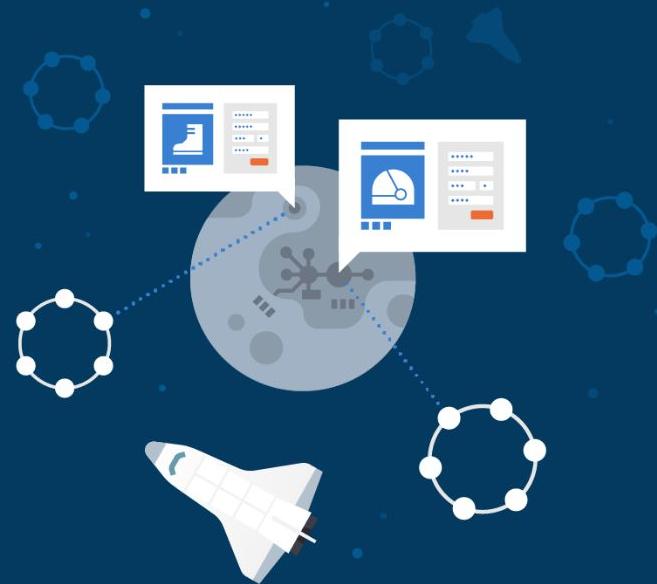


Potential Inconsistency

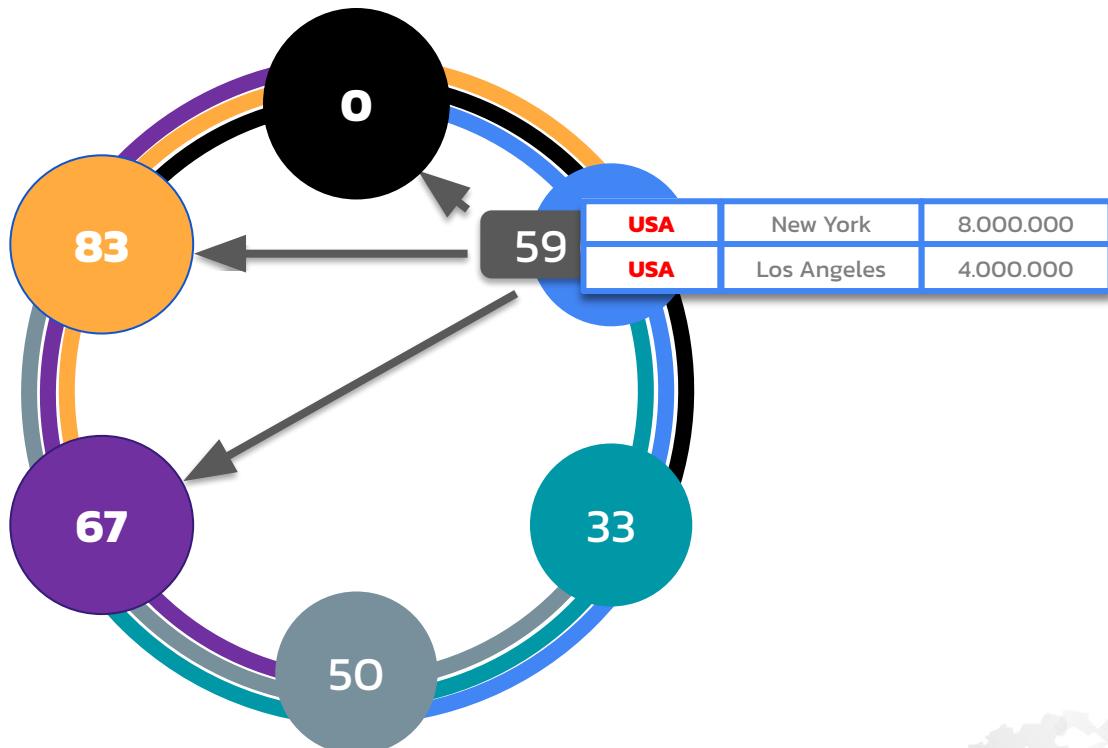


Potential Inconsistency

Cassandra Layered Self-Defence

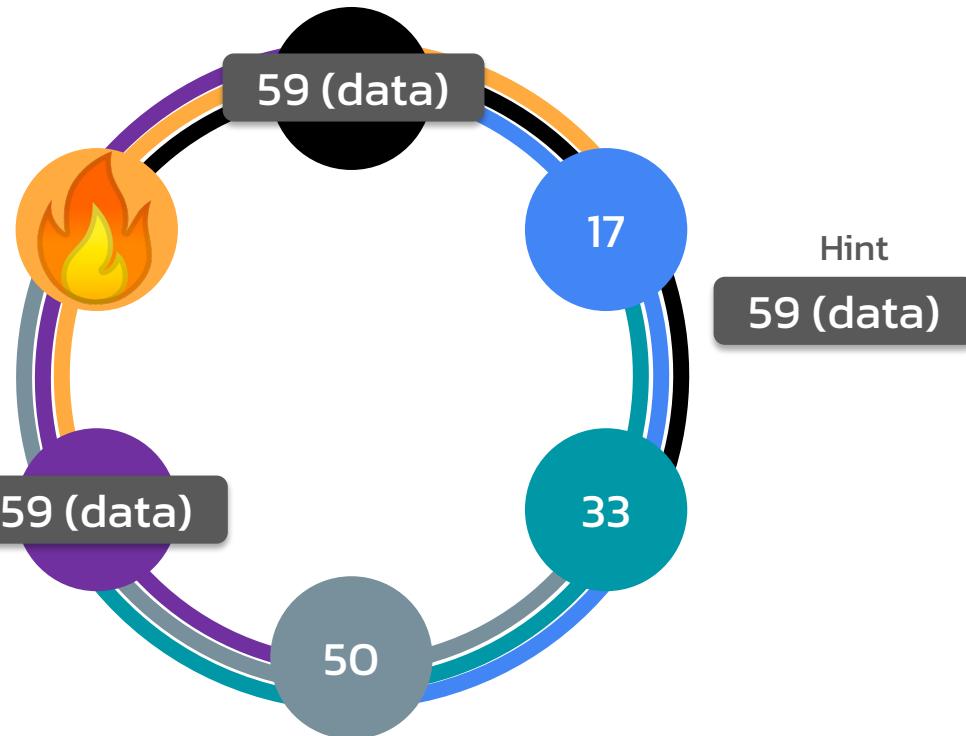


RF = 3



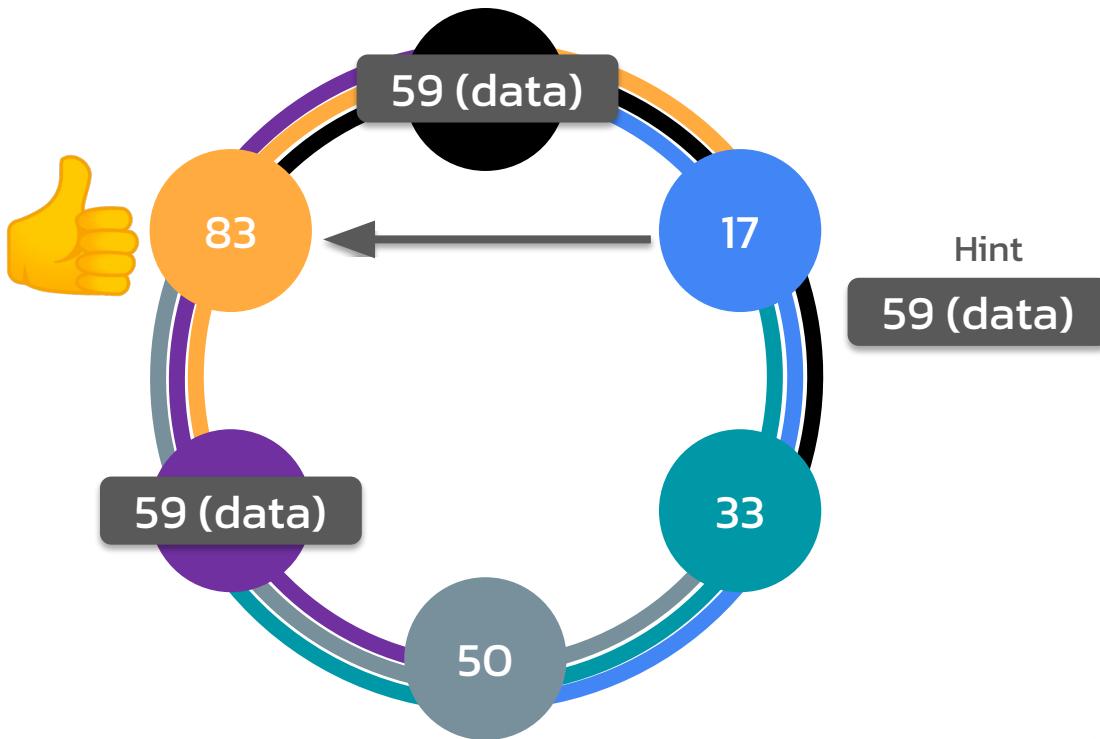
Hinted Hand-offs

RF = 3



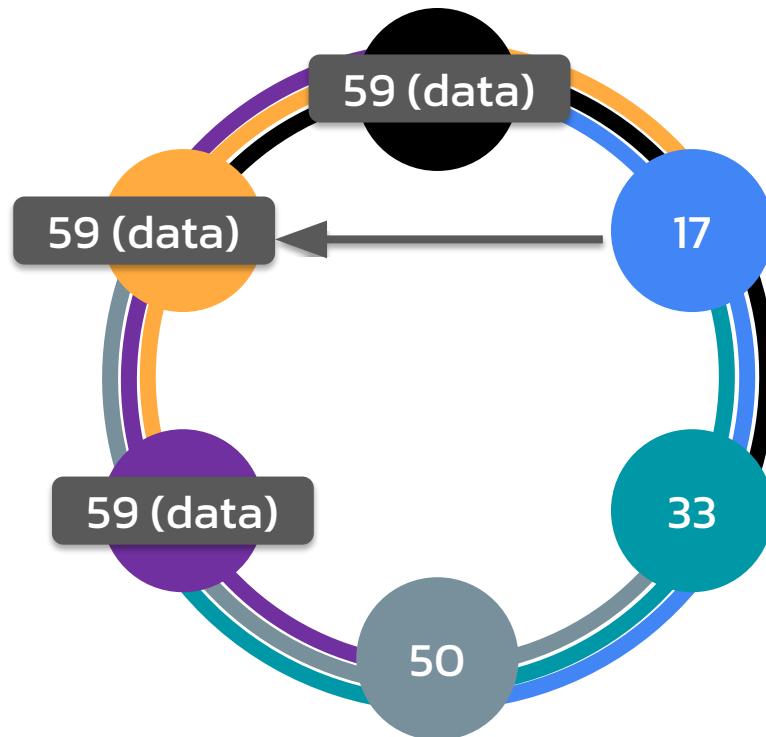
Hinted Hand-offs

RF = 3



Hinted Hand-offs

RF = 3



Hinted Hand-offs

CAP Theorem



CAP Theorem operates three features:

1. Availability
2. Consistency
3. Partition Tolerance

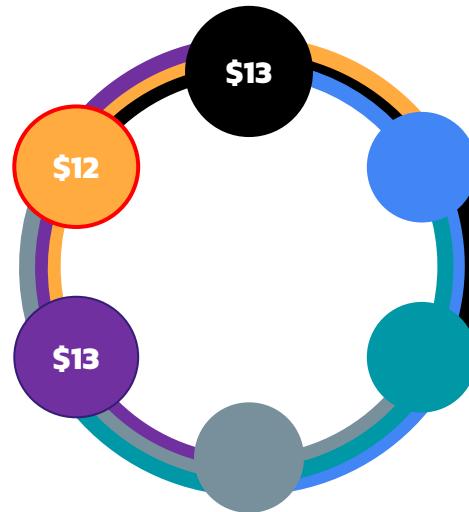


Availability basically means “Uptime”. You ask the question, you get the answer. If failure of a single or even multiple servers doesn’t lead to no response (no downtime), your system is **available**.



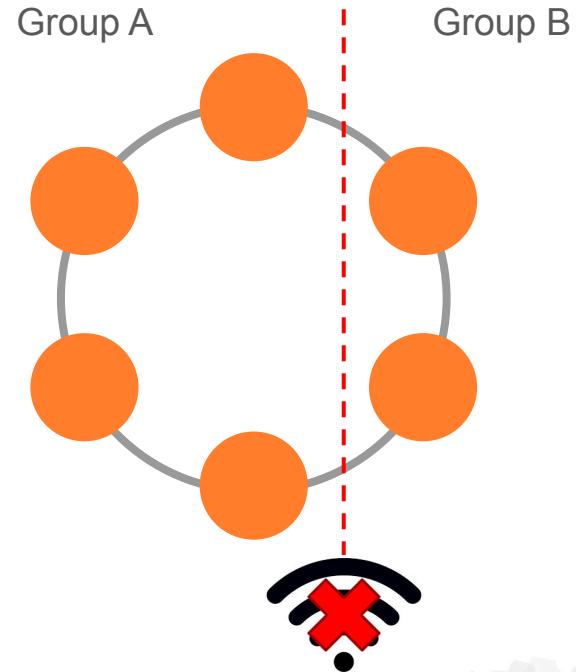
CAP Theorem: Availability

Consistency means “no stale data”. You ask for something, you get the most recent value. If one of your servers return outdated information, your system is **inconsistent**.

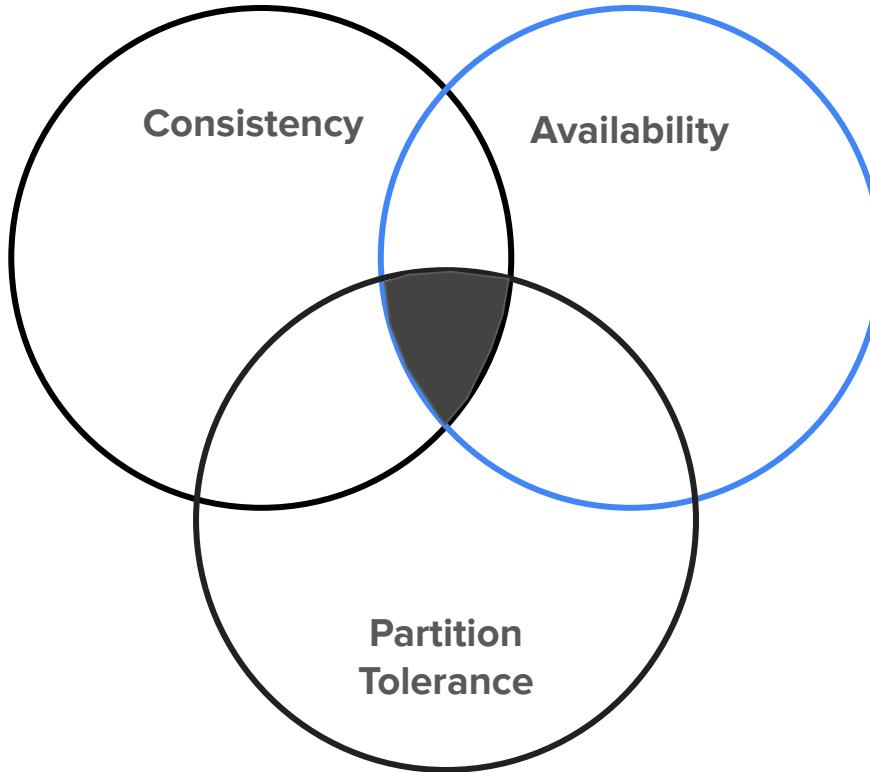


CAP Theorem: Consistency

Partition Tolerance is the ability of a distributed system to survive “network partitioning”. Network partitioning means that the part of the servers can not reach the second part.



In the distributed environment **in case of emergency** you can have only two guaranteed qualities out of three :(



CAP Theorem

Cassandra is configurably consistent. In any moment of the time, for any particular query you can set the Consistency Level you require to have. It defines how many **CONFIRMATIONS** you'll wait before the response is dispatched;

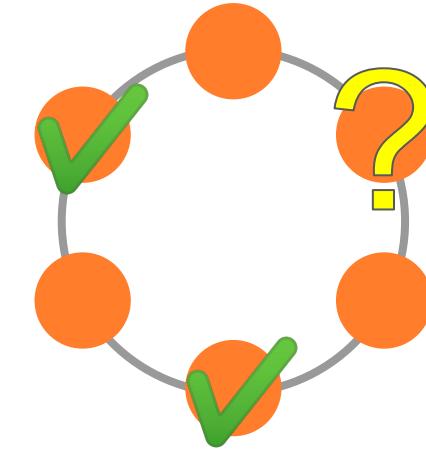
```
PreparedStatement pstmt = session.prepare(  
    "INSERT INTO product (sku, description) VALUES (?, ?)"  
);  
pstmt.setConsistencyLevel(ConsistencyLevel.ONE);
```

```
cqlsh> CONSISTENCY
```

```
Current consistency level is QUORUM.
```

```
cqlsh> CONSISTENCY ALL
```

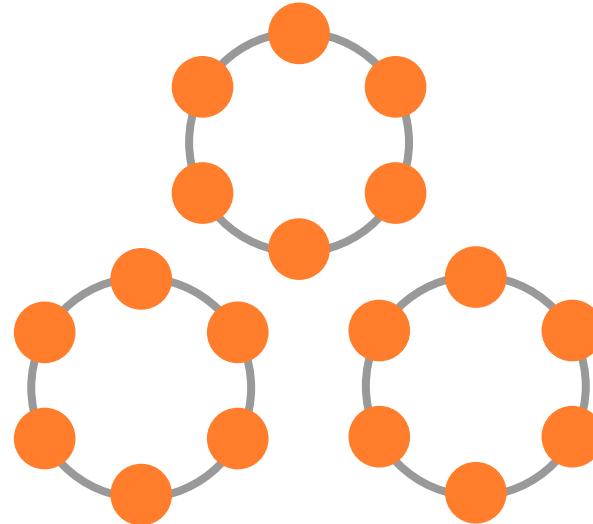
```
Consistency level set to ALL.
```



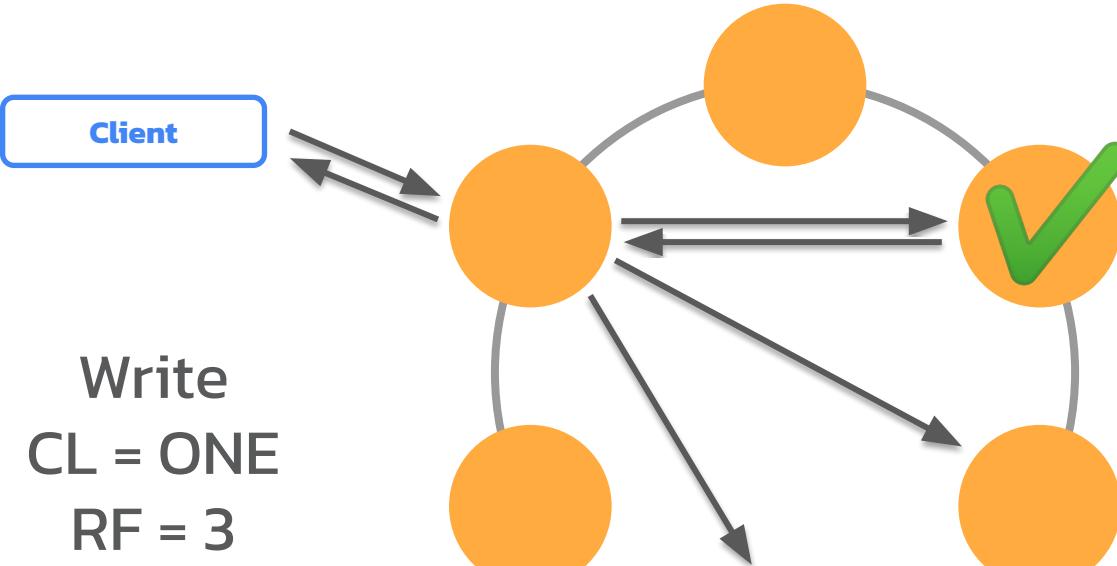
Is Cassandra AP or CP?



- ANY
- ONE
- LOCAL_ONE
- TWO, THREE
- QUORUM
- LOCAL_QUORUM
- EACH_QUORUM
- ALL

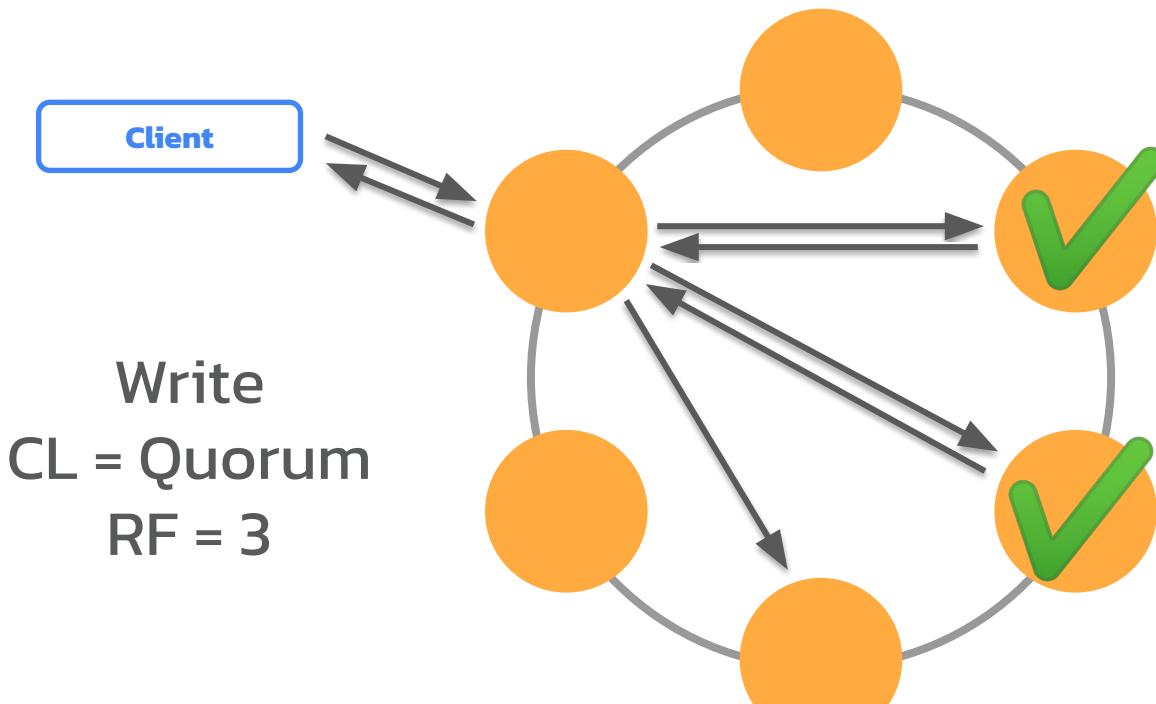


Query Consistency Levels



Consistency Level One

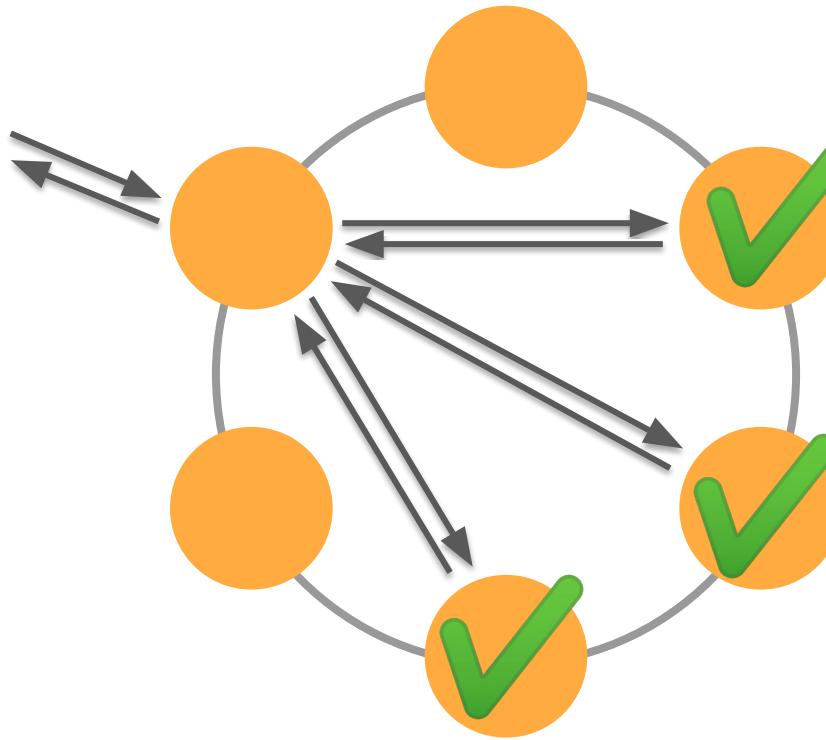




Consistency Level Quorum

Client

Write
CL = ALL
RF = 3



Consistency Level ALL



Client

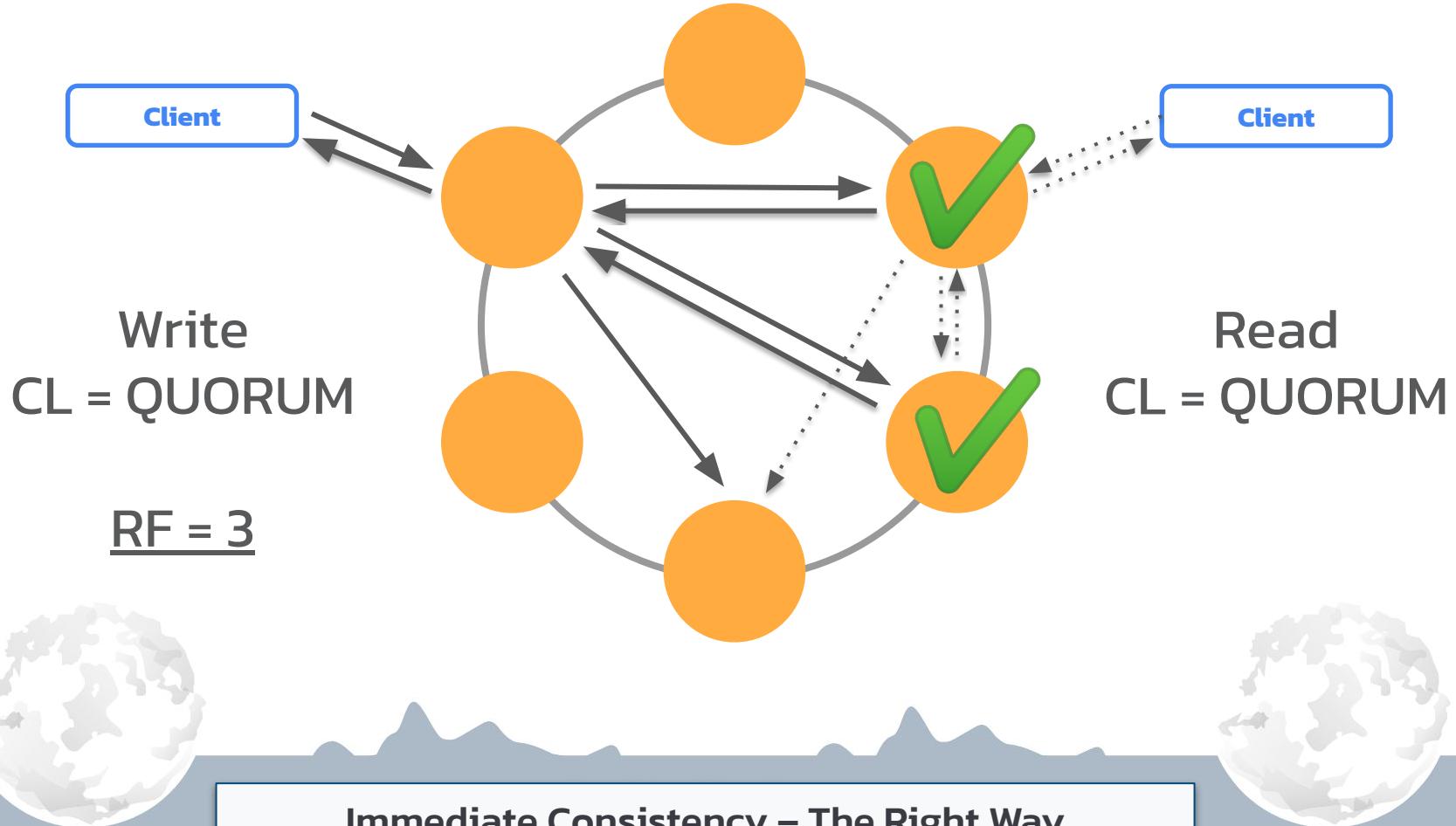
Writes
CL = ALL
RF = 3

IT'S A TRAP!

CAP THEOREM IS STILL HERE!

Consistency Level ALL

CL Write + CL Read > RF → Immediate Consistency



01



Intro to Bootcamp 2022
HouseKeeping

02

Intro to NoSQL Databases
Power of Purpose-built DB

03

Intro to Cassandra
Peer-to-Peer Database

04

CAP Theorem
Law of Distributed Systems

05

Tables and Partitions
How do you structure data

06

What's next?
Quiz, Homework, Next week



Agenda

Data Structure: a Cell



An intersection of a row
and a column, stores data.



DataStax Developers



Data Structure: a Row



A single, structured
data item in a table.

1	John	Doe	Wizardry
---	------	-----	----------



DataStax Developers



Data Structure: a Partition



A group of rows having the same partition token, a base unit of access in Cassandra.

IMPORTANT: stored together, all the rows are guaranteed to be neighbors.

ID	First Name	Last Name	Department
1	John	Doe	Wizardry
399	Marisha	Chavez	Wizardry
415	Maximus	Flavius	Wizardry



Data Structure: a Table



A group of columns and rows storing partitions.

ID	First Name	Last Name	Department
1	John	Doe	Wizardry
2	Mary	Smith	Dark Magic
3	Patrick	McFadin	DevRel



DataStax Developers



01



Intro to Bootcamp 2022
HouseKeeping

02

Intro to NoSQL Databases
Power of Purpose-built DB

03

Intro to Cassandra
Peer-to-Peer Database

04

CAP Theorem
Law of Distributed Systems

05

Tables and Partitions
How do you structure data

06

What's next?
Quiz, Homework, Next week



Agenda

menti.com



Go to www.menti.com and use the code 3491 9972

Inequality predicates are allowed on ...

A bar chart titled "Inequality predicates are allowed on ...". The y-axis represents the count of inequality predicates, ranging from 1 to 15. The x-axis categories are "All table columns", "Partition key columns", "clustering key columns", and "No inequality predicates are allowed".

Column Type	Count
All table columns	4
Partition key columns	3
clustering key columns	15
No inequality predicates are allowed	1

Below the chart, there is a video player interface showing a video of a person speaking. The video player includes controls like play/pause, volume, and a progress bar indicating 2:10:19 / 2:26:05. The title of the video is "Big paycheck".

Go to www.menti.com and use the code 3491 9972

Leaderboard

User ID	User Name	Profile Icon
4821 p	spanda	Avatar of a green sphere
4820 p	Agent X9	Avatar of a black mask
4775 p	Sam	Avatar of a cat
4711 p	CCedrickThePresenter	Avatar of a green leaf
4468 p	shubham	Avatar of a yellow bird
4371 p	aaa	Avatar of a yellow cat
3895 p	vignesh	Avatar of a red crown
3877 p	adry	Avatar of a smiling face
3861 p	Millie	Avatar of a brown bear
3812 p	Puggie	Avatar of a brown dog

Below the leaderboard, there is a video player interface showing a video of a person speaking. The video player includes controls like play/pause, volume, and a progress bar indicating 2:11:07 / 2:26:05. The title of the video is "Big paycheck".

SWAG WINNERS



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

To claim your prize, please send an email to:

jack.fryer@datastax.com

**** Include a screenshot of your menti screen**

Homework

!homework

🎓🔥 Intro to Cassandra for Developers
using DataStax Astra DB 🔥🎓



The course card has a dark blue header with a white eye icon and the text "DS201". The main body is white with dark blue text. It reads "DS201: Foundations of Apache Cassandra™ and DataStax" and "Online Course · 4h". At the bottom is a blue button with the word "Resume" and a small white icon.



DataStax Developers



Next Week

Design a Data Model for a Full-Stack Application

12 January, 2022

#NoSQL, #Apache Cassandra, #APIs, #DataModeling #Certification

Python Java JS Apache Cassandra APIs DataModeling Certification

WED JAN 12 2022

Bootcamp - Design a Data Model for a Full-Stack Application

[Register Now](#)

Join our 17k Discord Community

DataStax Developers



Discord interface showing the DataStax Developers community. The left sidebar lists channels: Événements, moderator-only, WELCOME, start-here, code-of-conduct, introductions, upcoming-events, useful-resources, memes, your-ideas, the-stage, WORKSHOPS (selected), workshop-chat (selected), workshop-feedback, workshop-materials, upcoming-workshops, ASTRADB, getting-started, astra-apis, astra-development, sample-applications, APACHE CASSANDRA, and Cedrick Lun... The main channel, #workshop-chat, contains a pinned message from RIGGITYREKT about mixed DSE version testing. Other messages from Erick Ramirez and Cedrick Lunven provide technical advice. On the right, a list of presenters and helpers is shown, along with a count of 560 users online.

workshop-chat <https://www.youtube.com/watch?v=MwU5xkFVVI> - Subscribe to mailing list: [http...](http://)

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

PRESENTER — 1
DAVID JONES-GILARDI

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

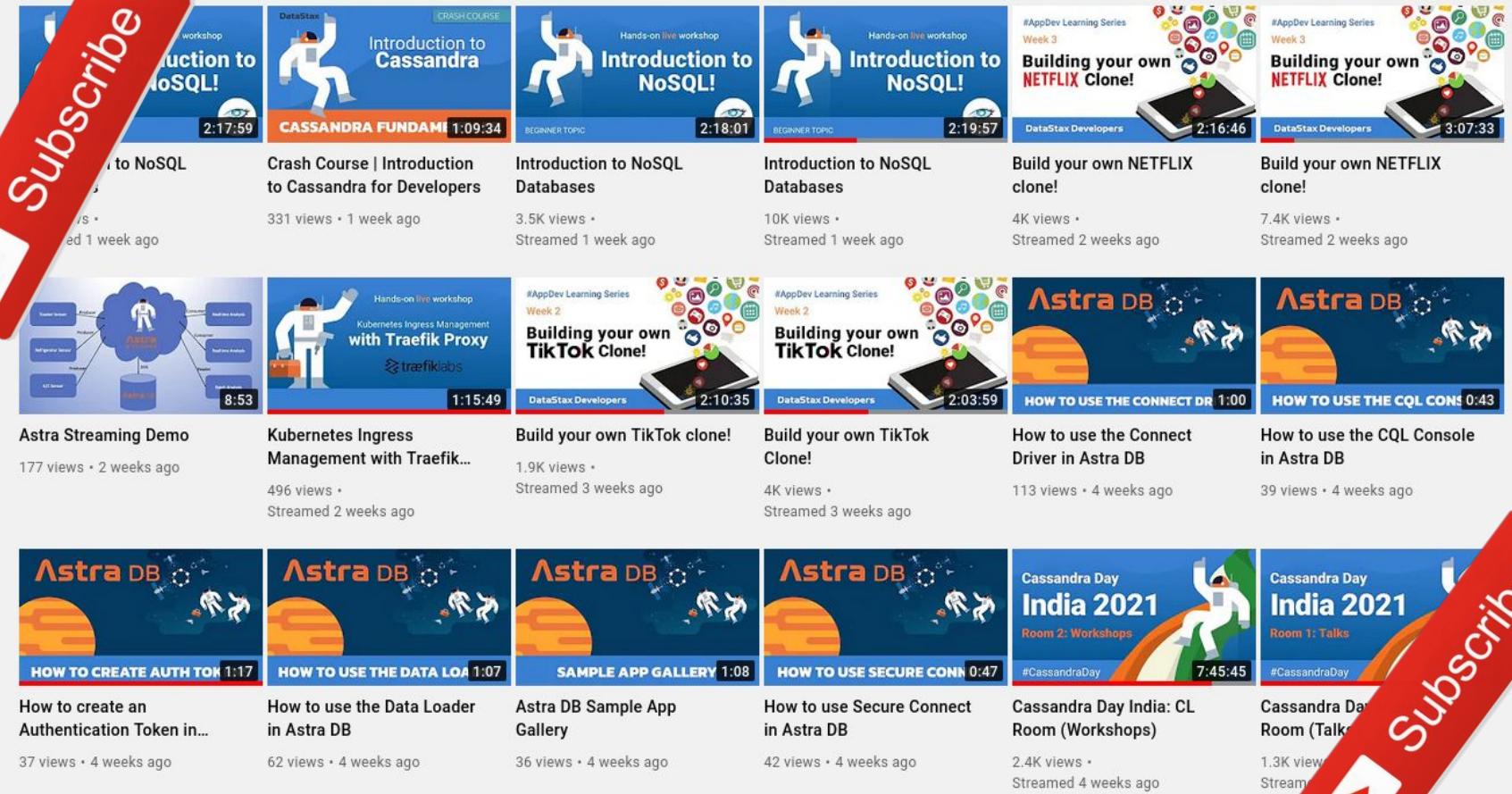
EN LIGNE — 560
-samu-
6304-42J8
Aahlya
Abdurahim
abhi3pathi
Abhiis.s
Abhineet
Abirsh

!discord

dtsx.io/discord

DataStax Developers

Subscribe



Subscribe

DataStax Developers

Thank You!

