



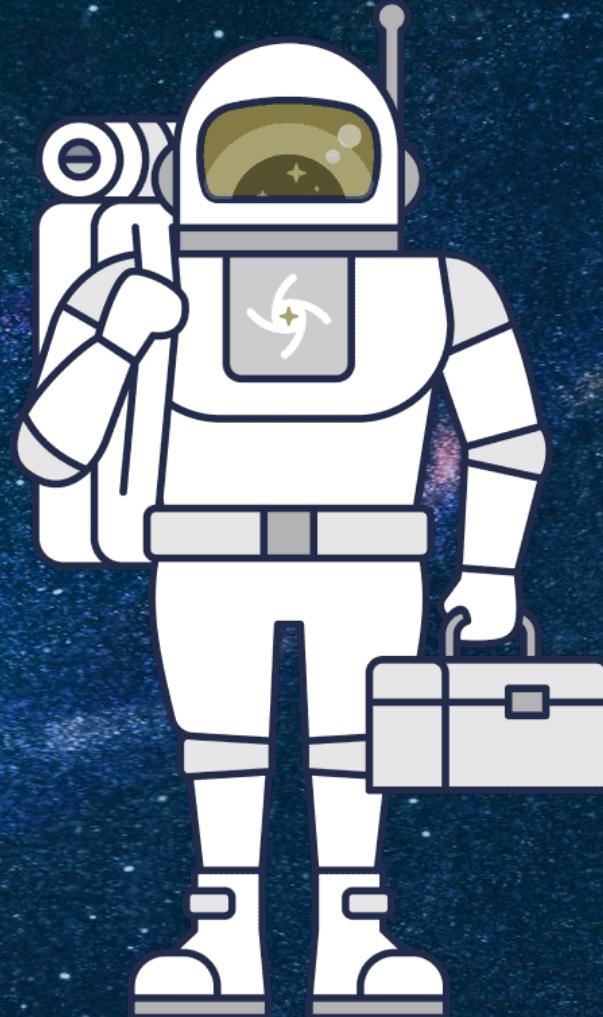
# STARGATE

A gateway for multi-models Data API

24th October – version 0.2

Cedrick Lunven| DataStax, Inc.

@clunven | cedrick.lunven@datastax.com



# About Me



**CÉDRICK LUNVEN**

Director Of Developer  
Advocacy At Datastax



- ✓ Training
- ✓ Teaching
- ✓ Reference applications
- ✓ Support
- ✓ Public Speaking



Creator and  
Maintainer of FF4j  
[ff4j.org](http://ff4j.org)

# Agenda



## Why: The rise of Data Gateways

- ✓ What: Architecture and Implementation
- ✓ How: Apache Cassandra as Multi-model DB
- ✓ Demo, talk is cheap
- ✓ Astra, NEXT Generation Dbaas..with Stargate

# “Start with Why” Simon Sinek

- Developers

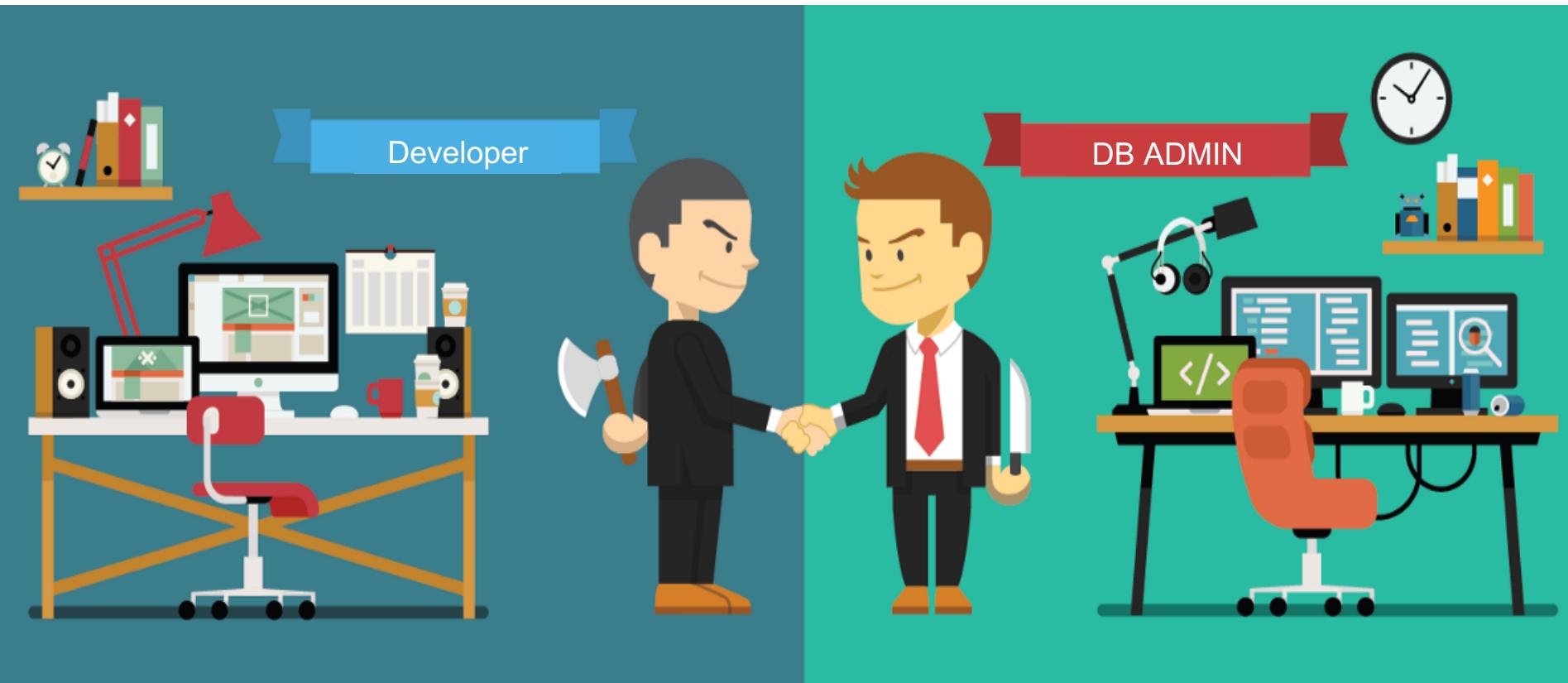
- Do you like learning query languages (CQL, N1QL, GQL, cypher, gremlin....)
  - No. Save my JSON, give it to me back when I need it
  - ORM / Spring Data are so popular nowadays
- Do you care about how your data is stored ?
  - Physical data model part of the interface...yike
  - Create structures based on queries
- Do you like installing and running Databases locally
  - especially distributed databases
  - especially with datasets and integration tests

# “Start with Why” Simon Sinek

- **Operators and Databases Administrators...**

- Do you allow developers to execute direct queries against your DB ?
- Do you like opening port ranges like 0-65536 to allow communications with applications, especially in the cloud.
- Do you like creating dedicated projects and hiring people just to create APIs to expose an existing treatment (digital transformation FTW)

# And when those 2 meet each other

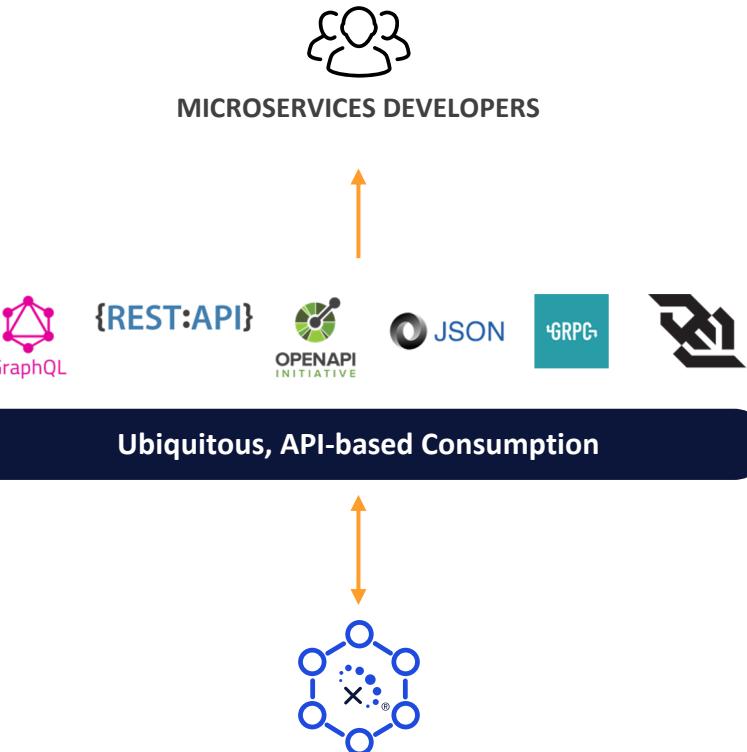


# API Data Gateways rational



Developers want the option to use modern APIs and development gateways.

Cassandra is a database designed for the new standard and the associated APIs that facilitate the first choice of developers.



# Agenda



Why: The rise of Data Gateways



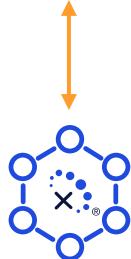
## What: Architecture and Implementation

- ✓ How: Apache Cassandra as Multi-model DB
- ✓ Demo, talk is cheap
- ✓ Astra, NEXT Generation Dbaas..with Stargate

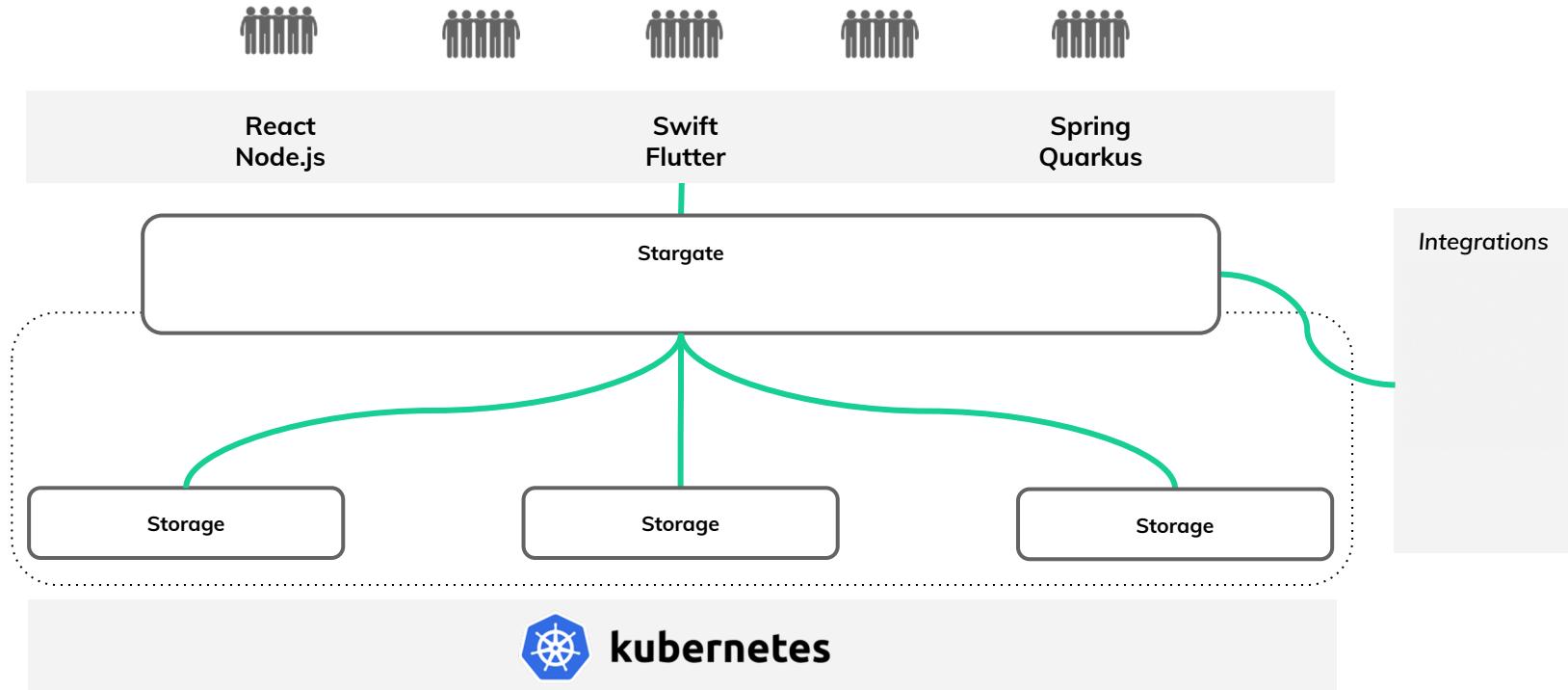
# Stargate Overview

An open source API framework for data

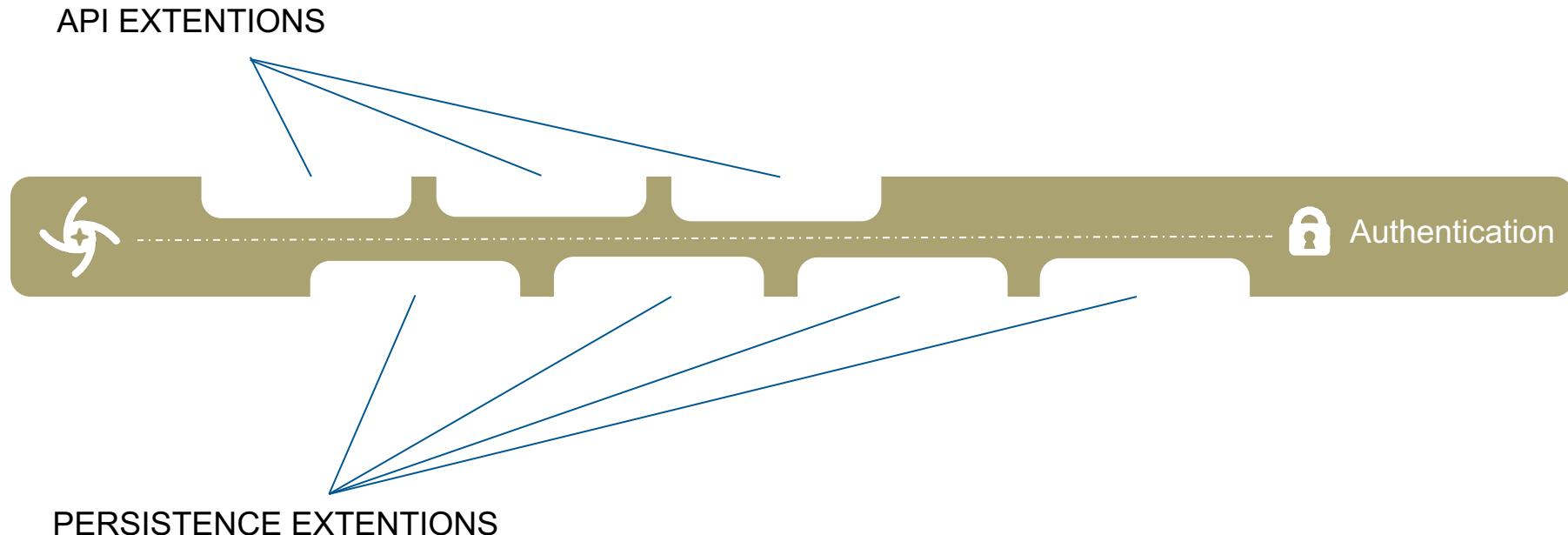
Stargate makes it easy to use a database for any application workload by adding plugin support for new APIs, data types, and access methods



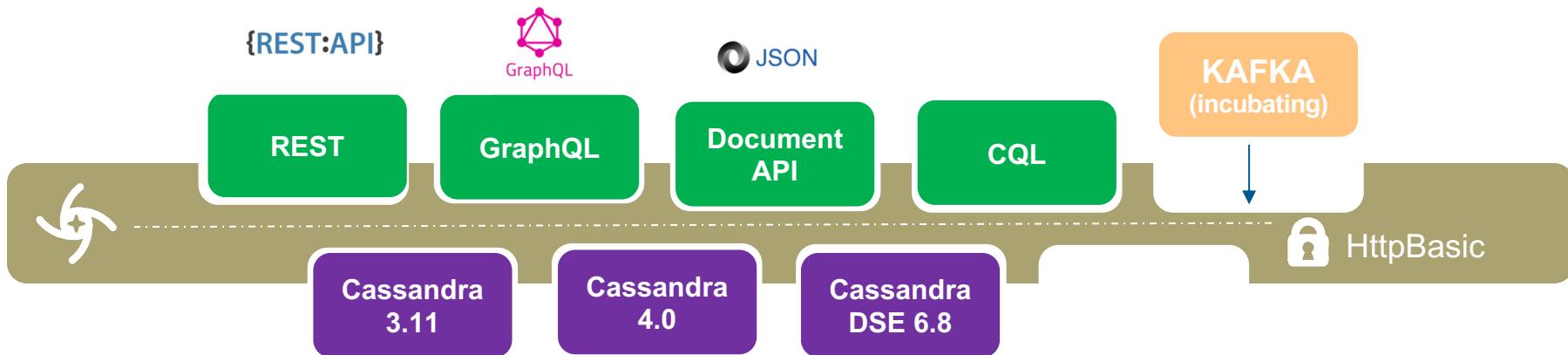
# Logical Architecture, Yet another Proxy



# API Extension and Persistence Extensions



# Modules



# Agenda



Why: The rise of Data Gateways

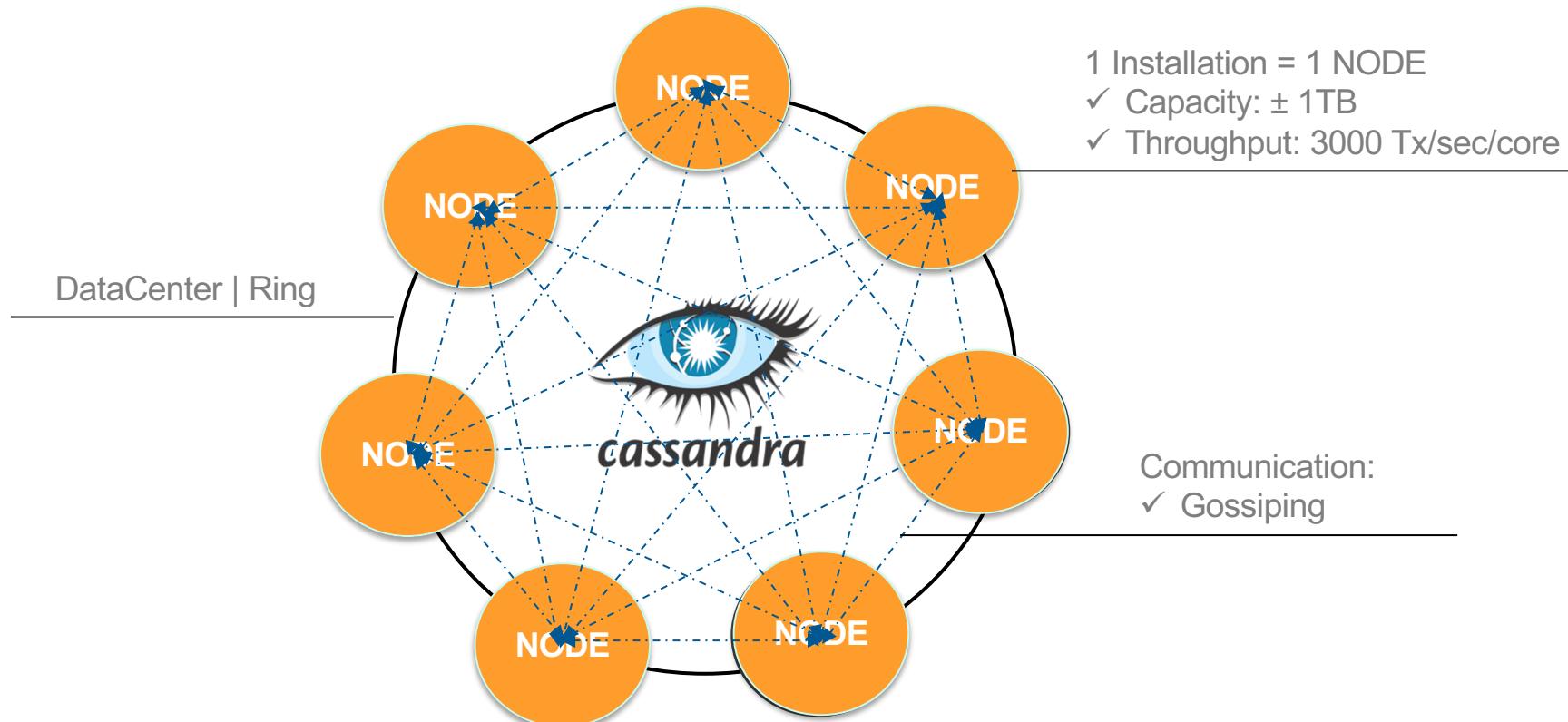
What: Architecture and Implementation



**How: Apache Cassandra as Multi-model DB**

- ✓ Demo, talk is cheap
- ✓ Astra, NEXT Generation Dbaas..with Stargate

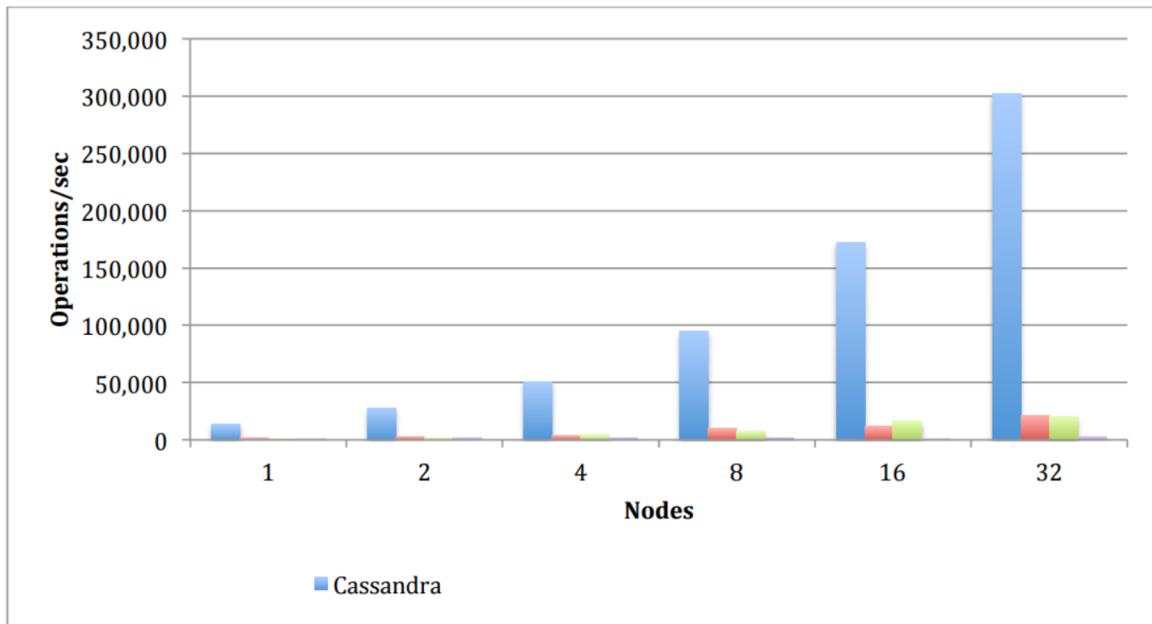
# Apache Cassandra™ = NoSQL Distributed Database



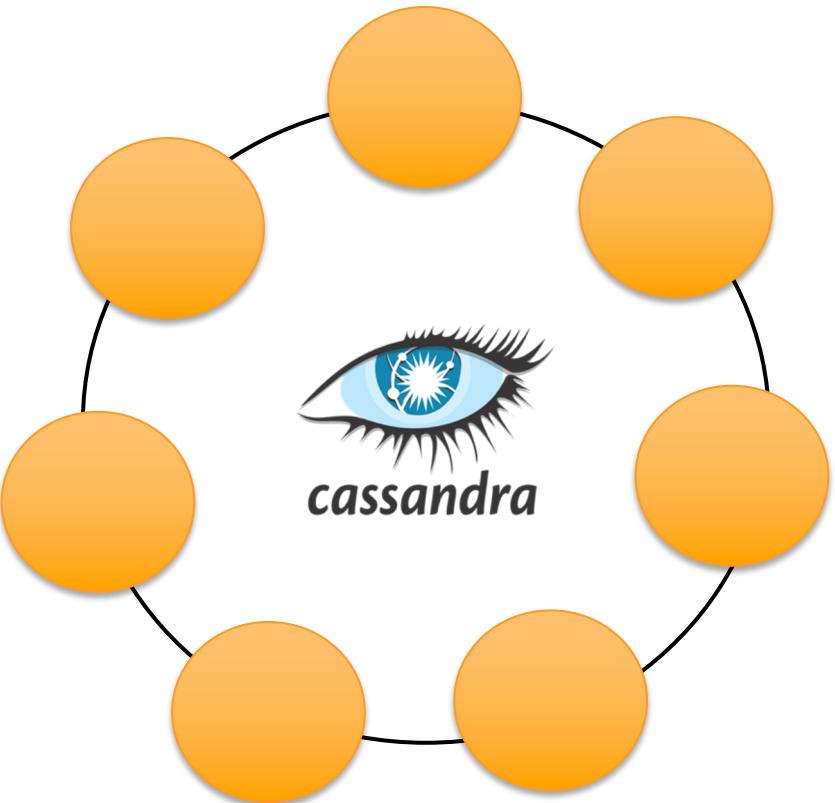
# Scales Linearly

- Need more capacity?
- Need more throughput?
- Add nodes!

Balanced Read/Write Mix



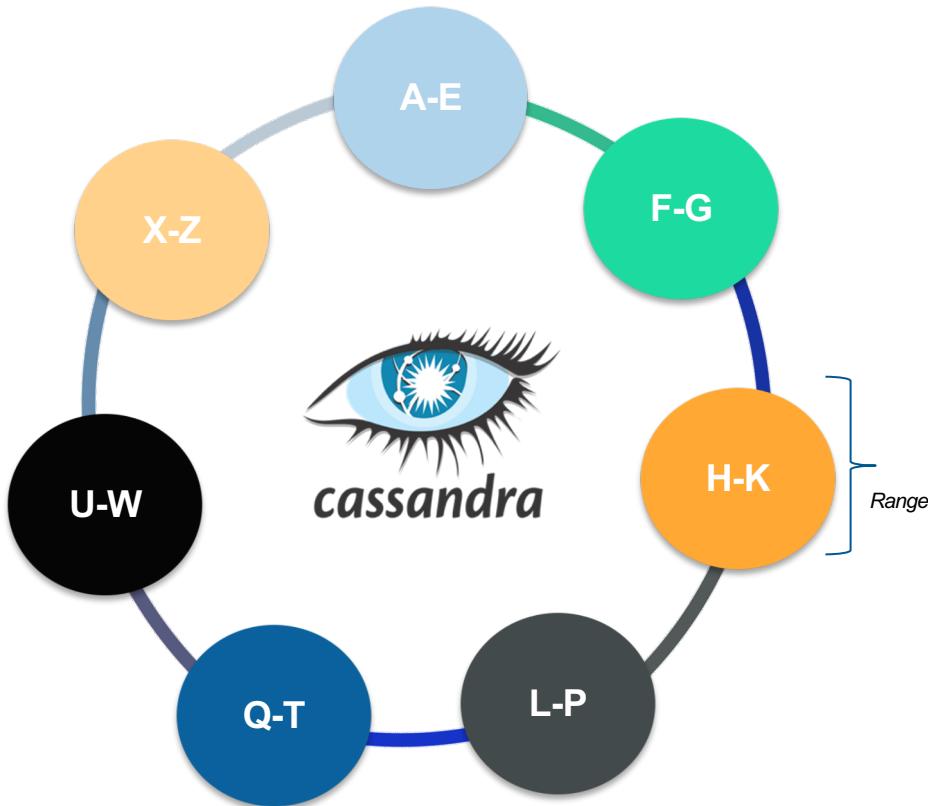
# Data is Distributed



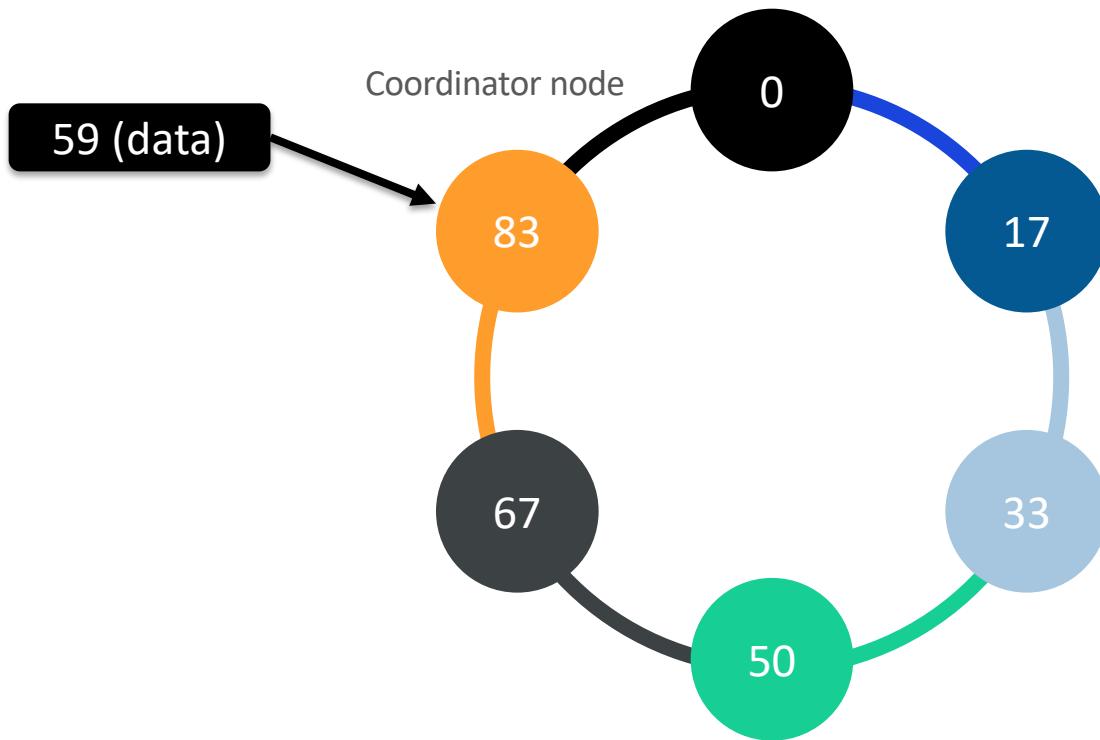
Country	City	Habitant
USA	New York	8.000.000
USA	Los Angeles	4.000.000
FR	Paris	2.230.000
DE	Berlin	3.350.000
UK	London	9.200.000
AU	Sydney	4.900.000
DE	Nuremberg	500.000
CA	Toronto	6.200.000
CA	Montreal	4.200.000
FR	Toulouse	1.100.000
JP	Tokyo	37.430.000
IN	Mumbai	20.200.000

Partition Key

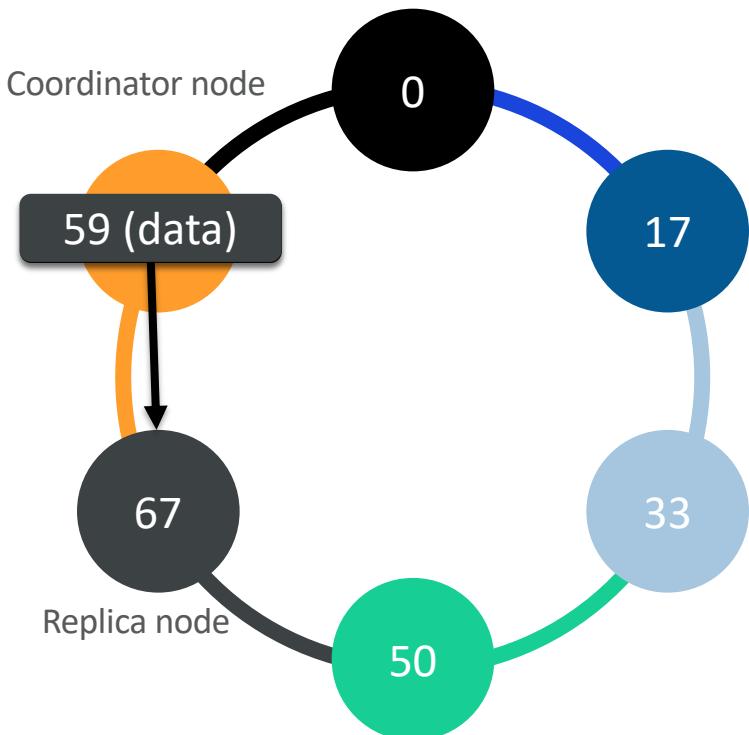
# Data is Distributed



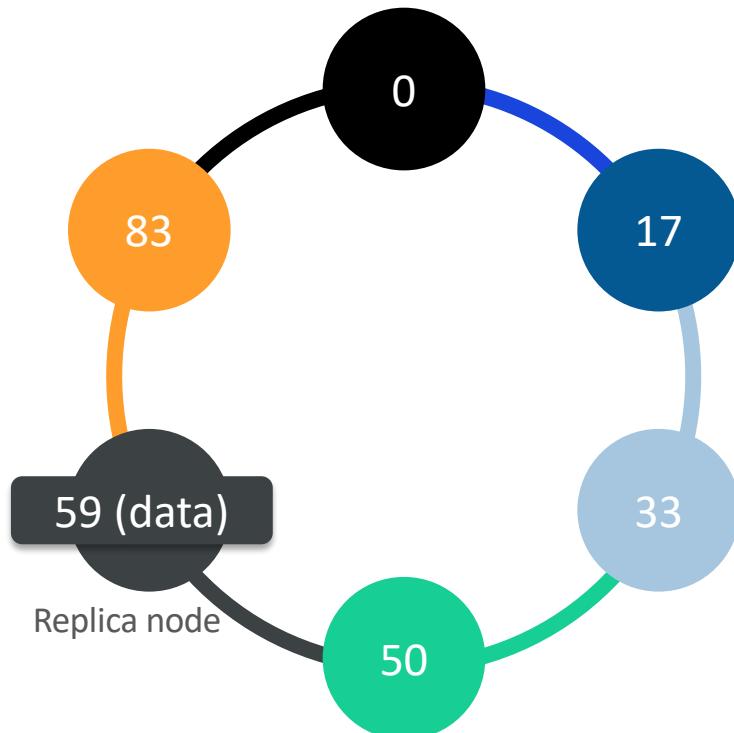
# How the Ring Works



# How the Ring Works

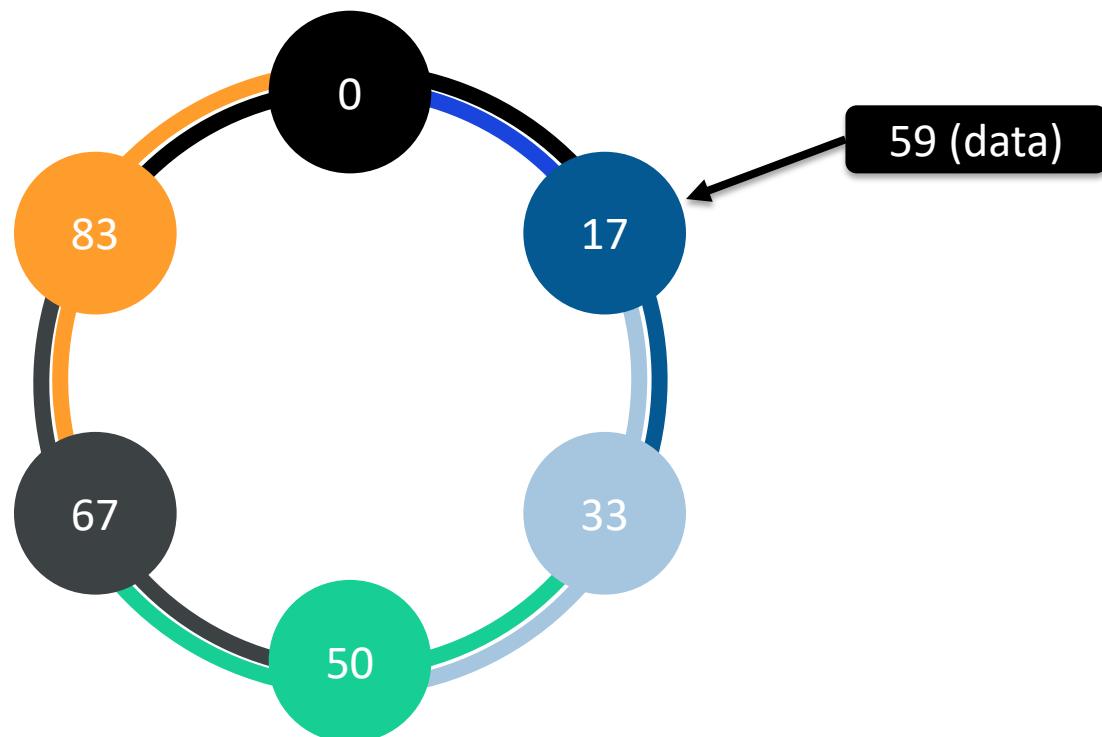


# How the Ring Works



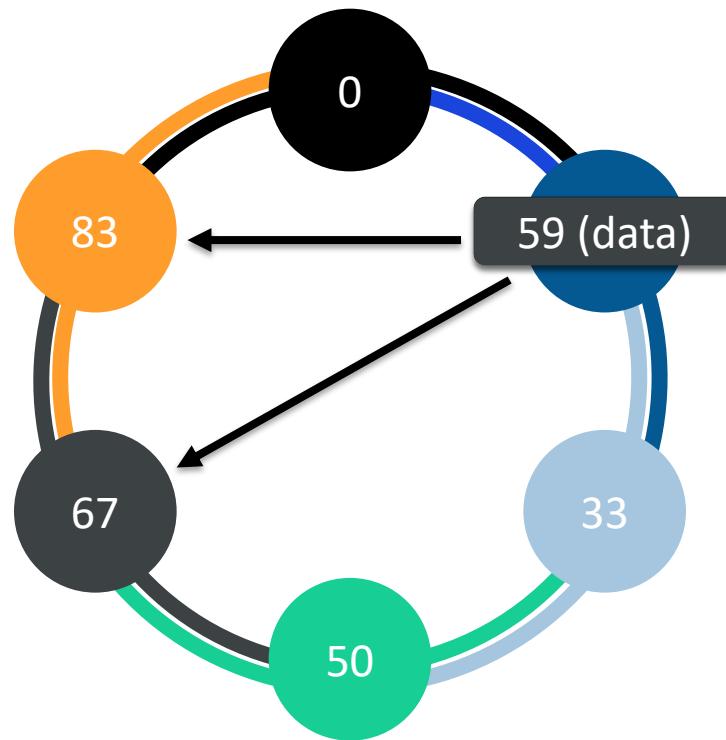
# Replication within the Ring

RF = 2

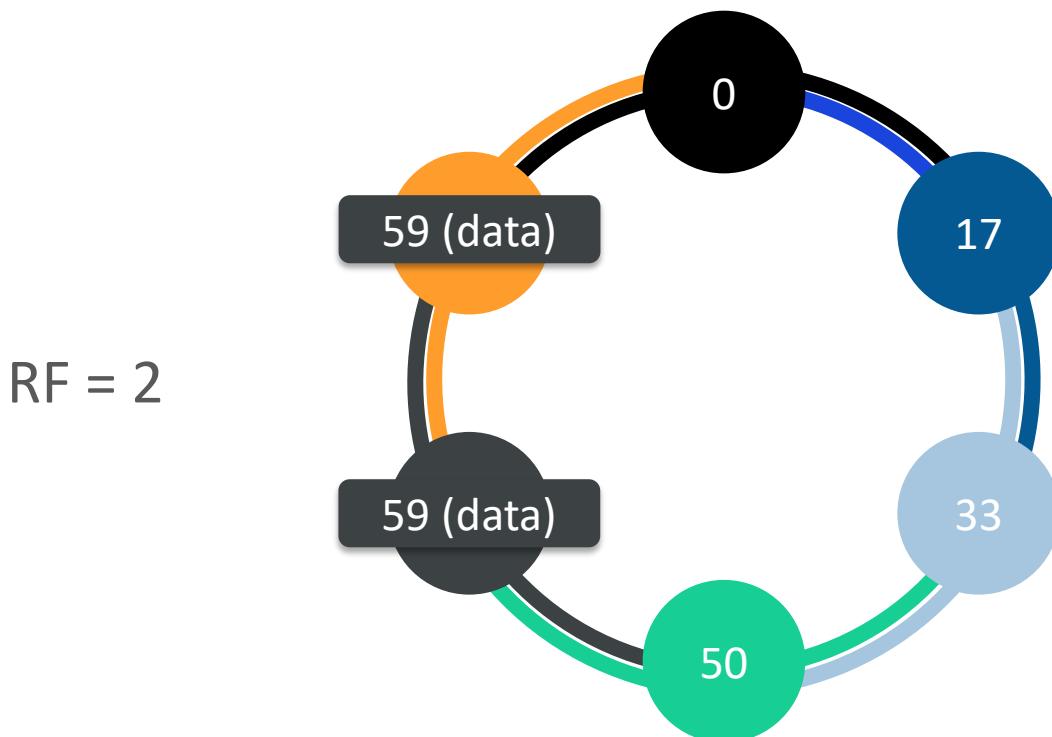


# Replication within the Ring

RF = 2

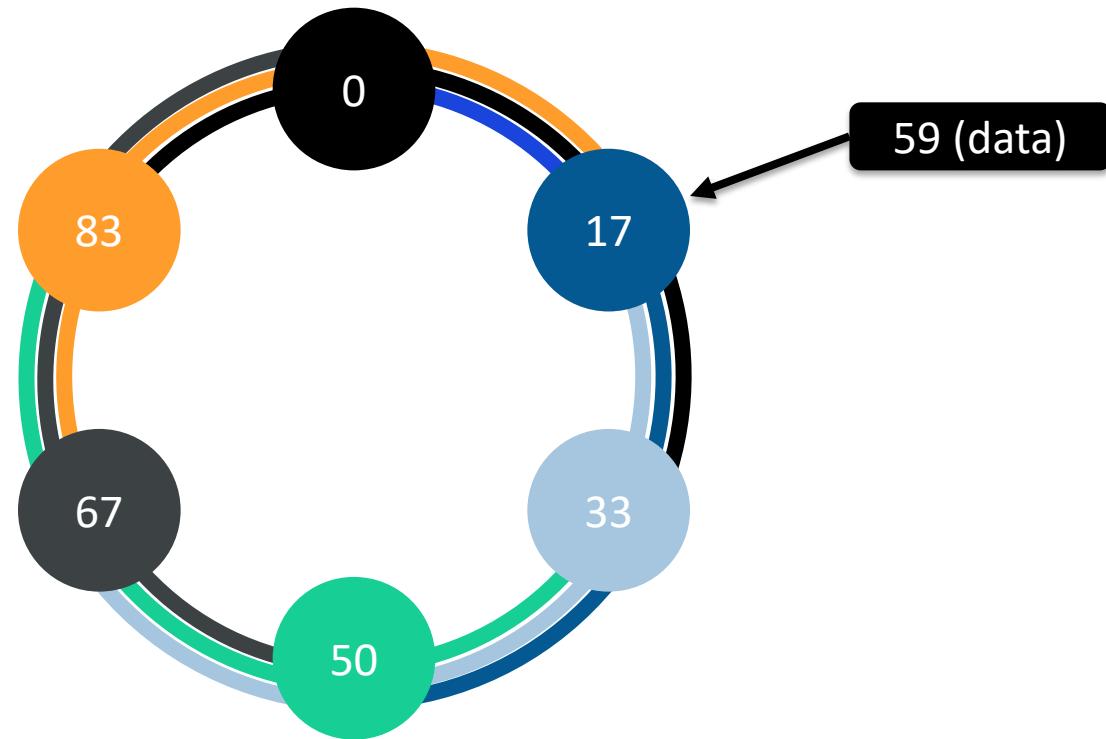


# Replication within the Ring



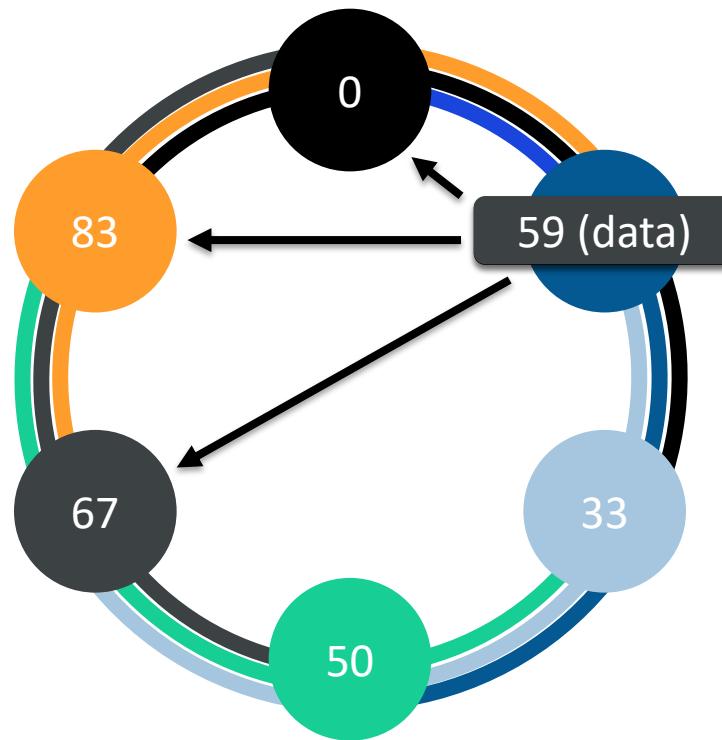
# Replication within the Ring

RF = 3



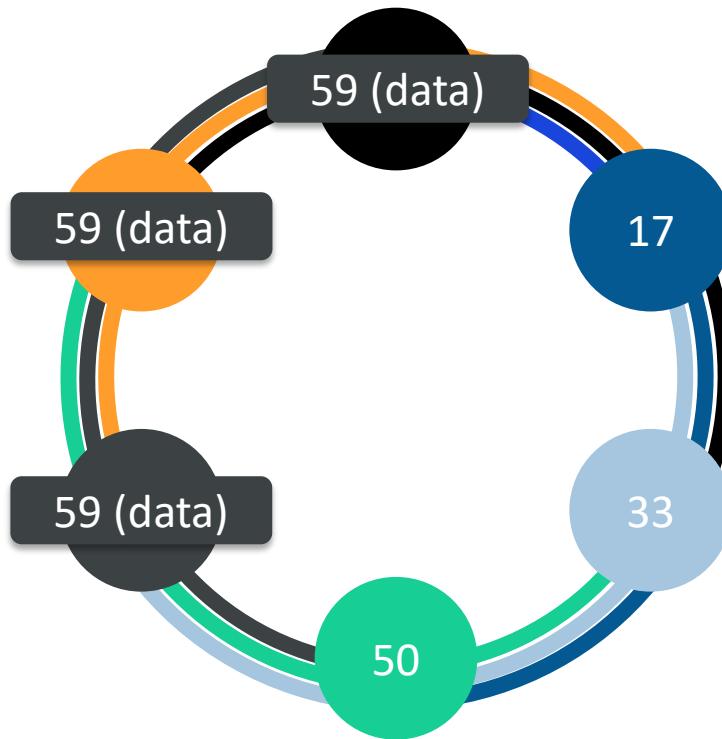
# Replication within the Ring

RF = 3



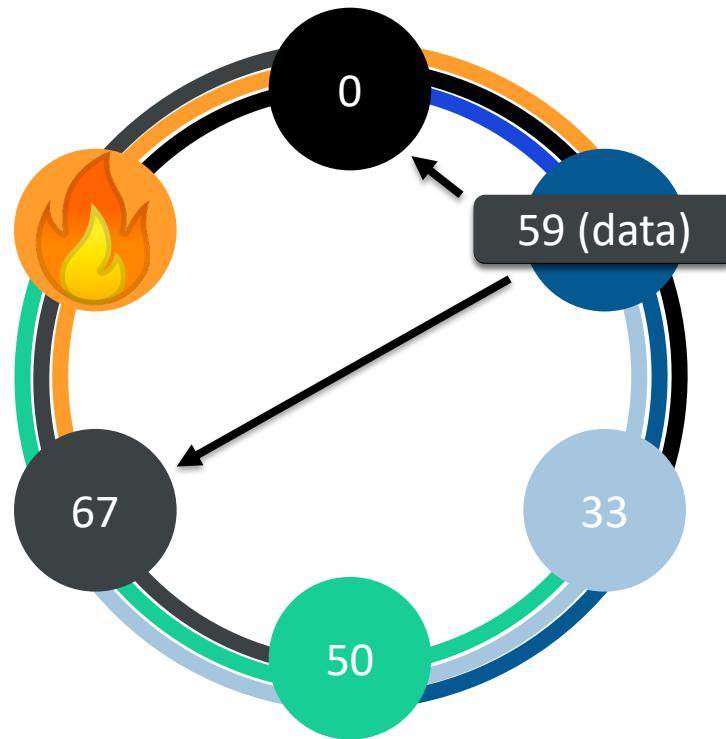
# Replication within the Ring

RF = 3



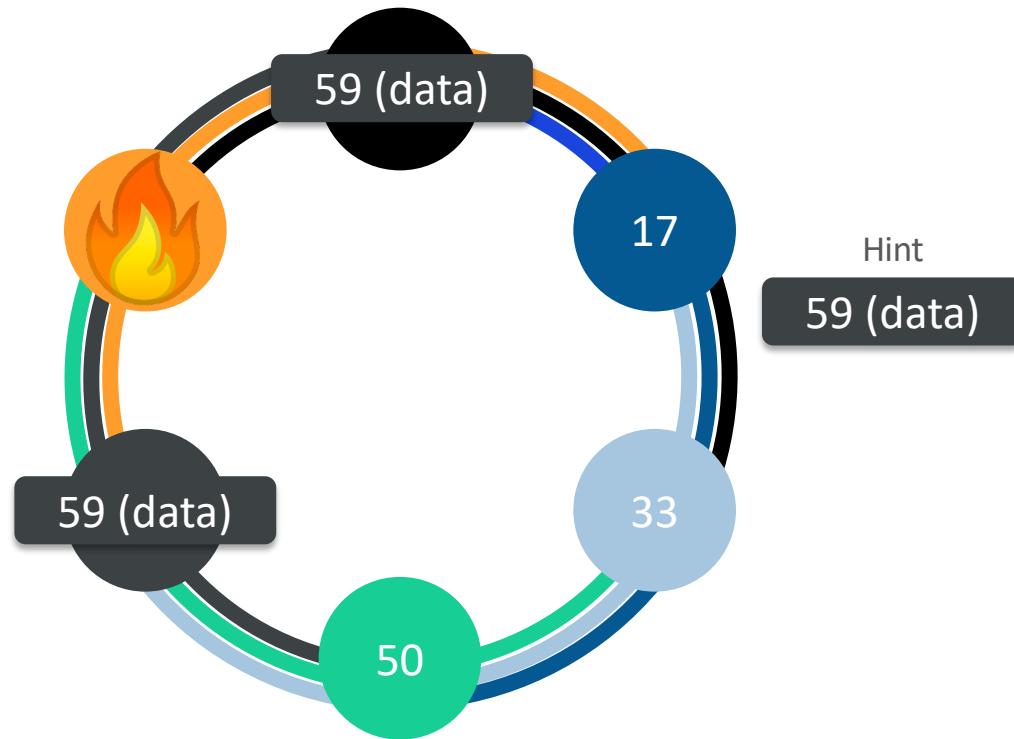
# Node Failure

RF = 3



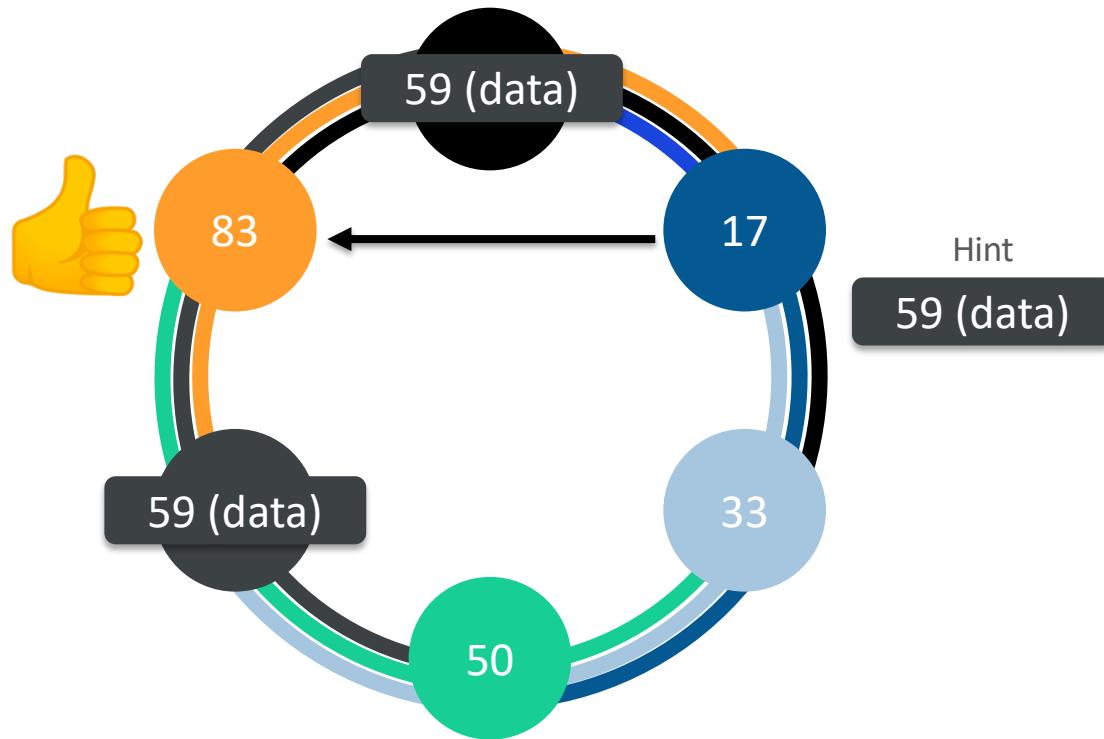
# Node Failure

RF = 3

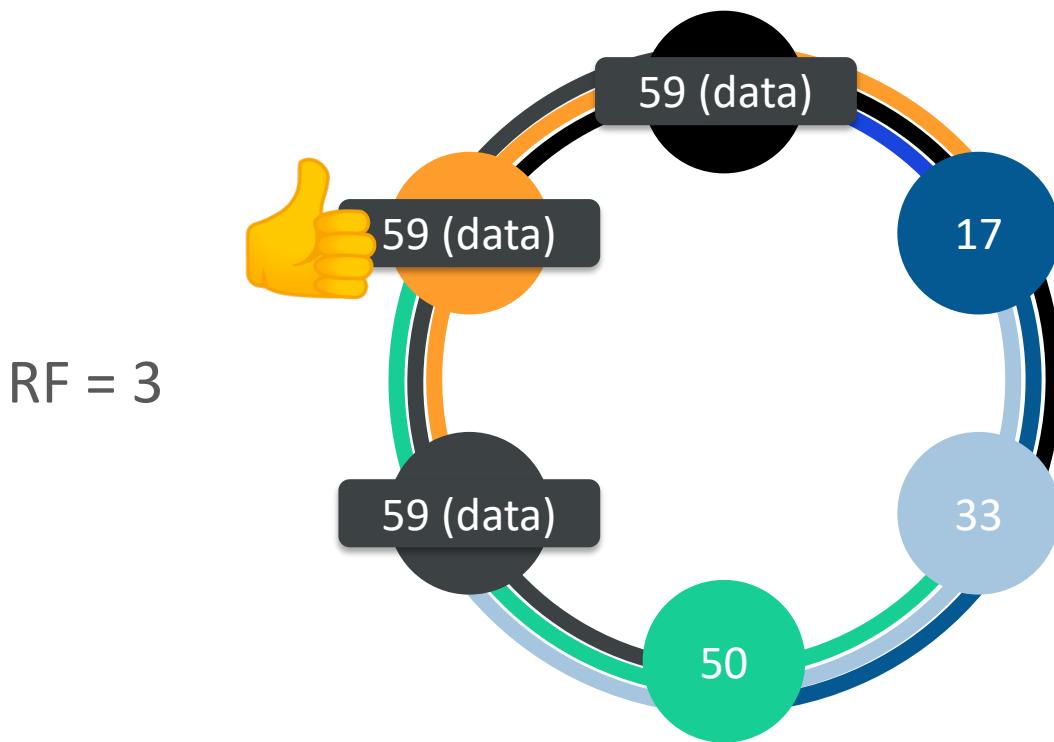


# Node Failure

RF = 3

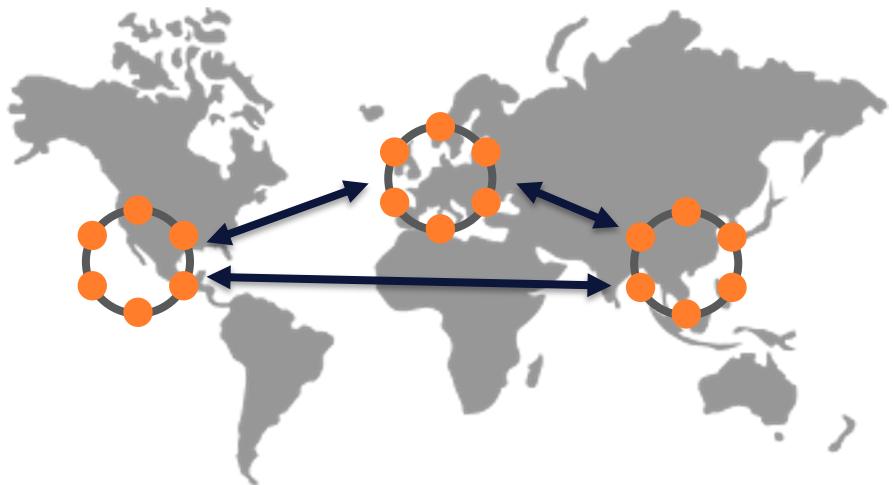


# Node Failure – Recovered!

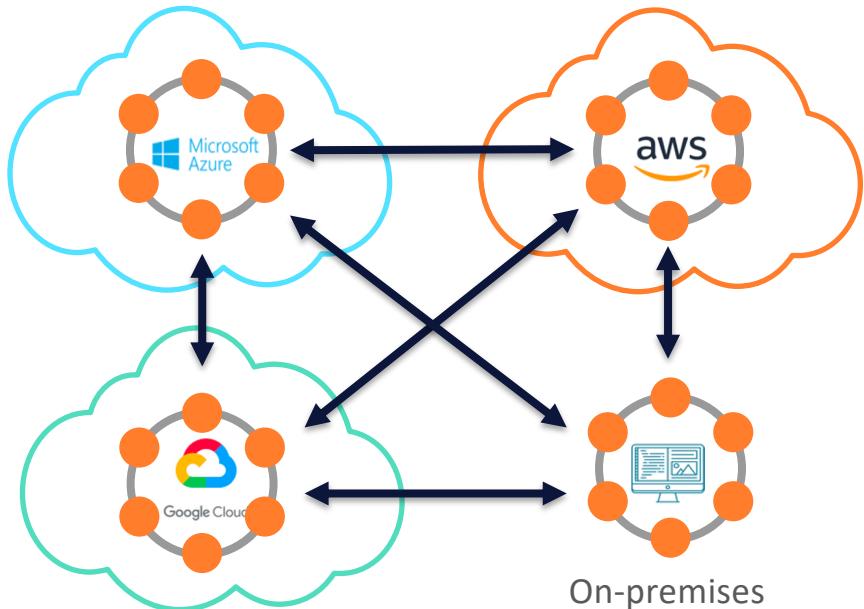


# Data Distributed Everywhere

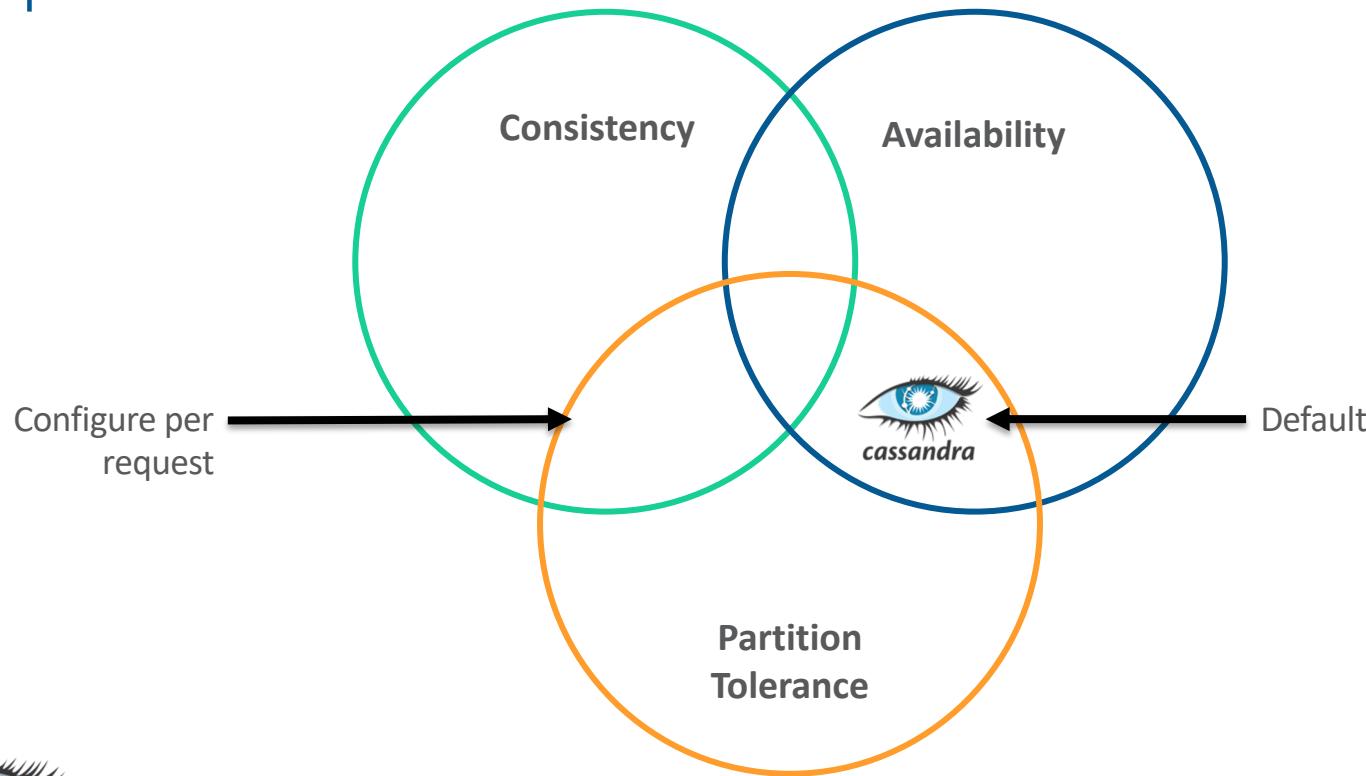
- Geographic Distribution



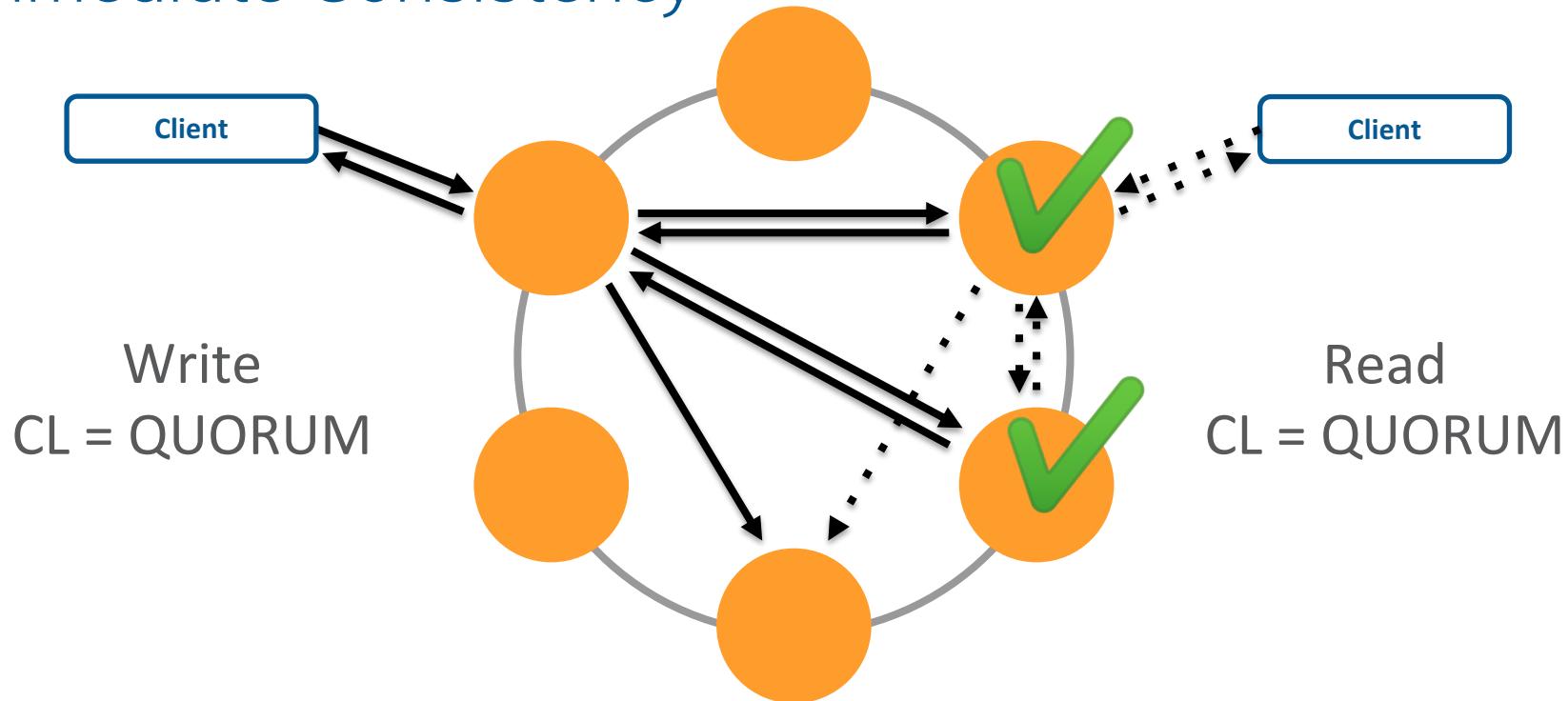
- Hybrid-Cloud and Multi-Cloud



# Cap Theorem

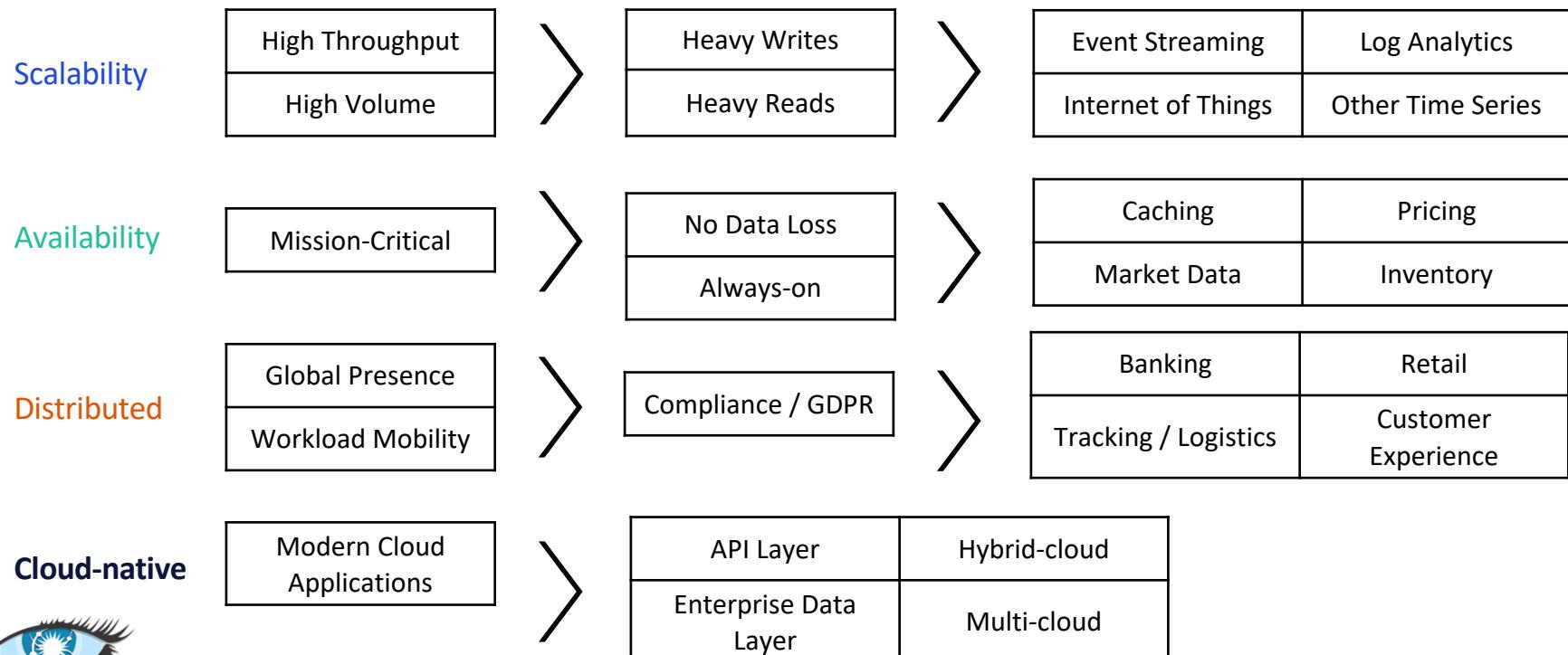


# Immediate Consistency



Immediate Consistent :  $CL_{read} + CL_{write} > RF$

# Understanding Use Cases



# Cassandra as a Multi-Models NoSsql Database



## Column Oriented

- Cassandra is contains **Tables** with a Partition key
- Defined as Column Oriented ( you can add columns)



## Key-Value

- This is like KEY and multiple values, fit the key-value pattern



## TimeSeries DataBase

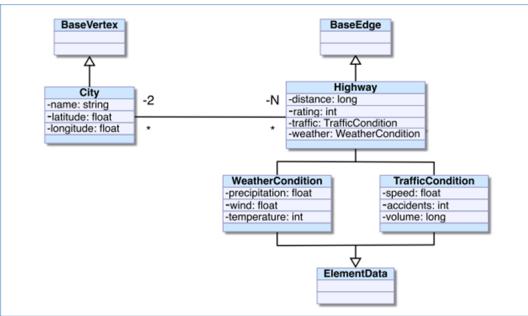
- Cassandra fits time series use cases (high throughput) and data models (CL)
- Provide: TTL, aggregation through Spark, split OLTP and OLAP



## Graph DataBase

- For 5 years Datastax has been implemented a graph Engine, fork of TitanDB on top of Cassandra storage, querin with **gremlin**.

# Cassandra as Document Store, the last frontier



- Cassandra already handles JSON and structures
  - INSERT JSON, SELECT JSON
  - Set<>, List<>, Map<>, and User Defined type (UDT) even nested
  - BUT everything is strongly coupled with a **SCHEMA**



# 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
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
x	c	[0]	d	2

# Agenda



Why: The rise of Data Gateways

What: Architecture and Implementation

How: Apache Cassandra as Multi-model DB



**Demo, talk is cheap**

✓ Astra, NEXT Generation Dbaas..with Stargate

<https://github.com/datastaxdevs/conference-2020-dataconla-stargate>

*How to actually learn any new programming concept*



*Essential*

Changing Stuff and  
Seeing What Happens

O RLY?

@ThePracticalDev



*Software can be chaotic, but we make it work*



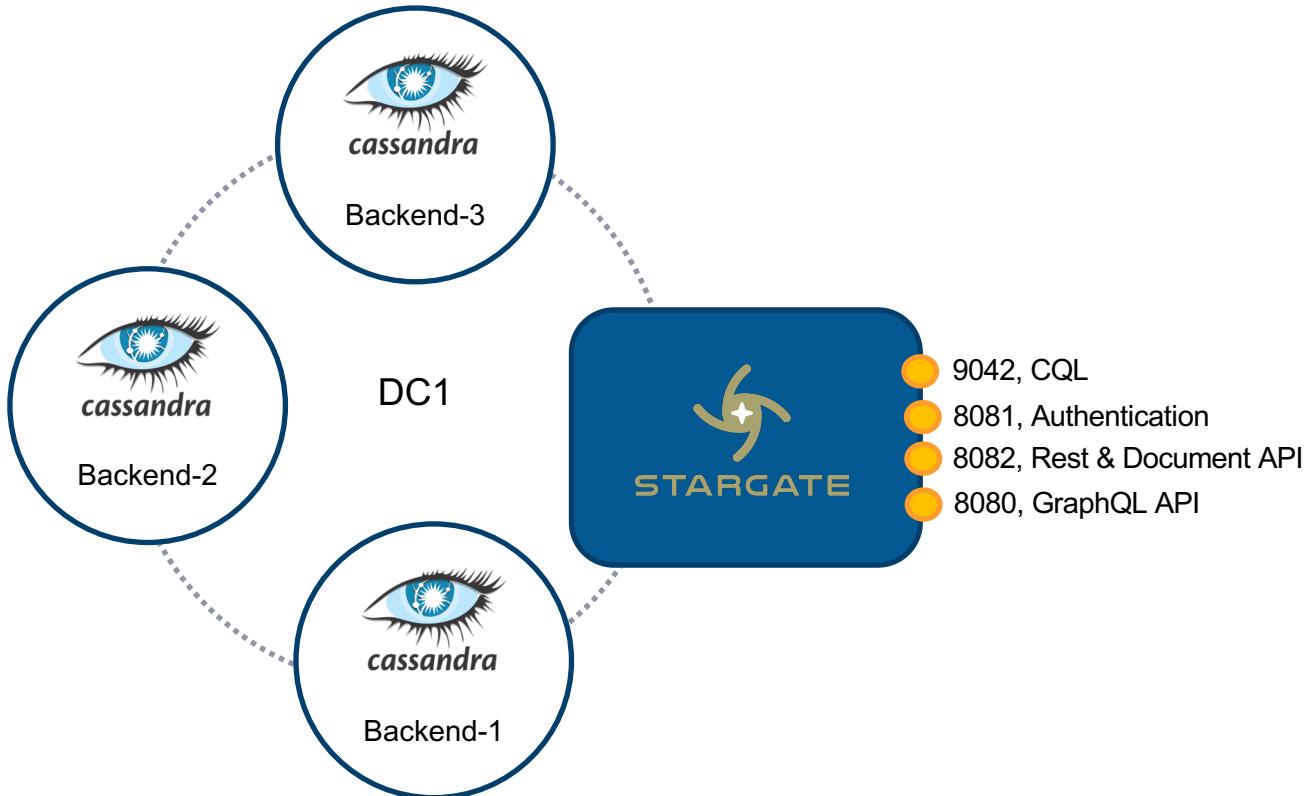
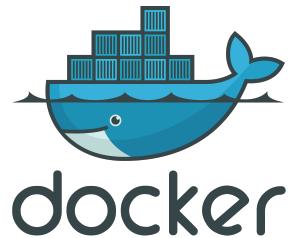
*Expert*

Trying Stuff  
Until it Works

O RLY?

The Practical Developer  
@ThePracticalDev

@elbruno



# Agenda



Why: The rise of Data Gateways

What: Architecture and Implementation

How: Apache Cassandra as Multi-model DB

Demo, talk is cheap



**Astra, NEXT Gen DBaaS.. with Stargate**

DATASTAX

# ASTRA



## CASSANDRA-AS-A-SERVICE

Cloud-native Database-as-a-Service built on Apache Cassandra



## CLOUD NATIVE

Powered by our open-source Kubernetes Operator for Cassandra.



## NO OPERATIONS

Eliminate the overhead to install, operate, and scale Cassandra.



## ZERO LOCK-IN

Deploy on AWS or GCP and keep compatibility with open-source Cassandra.



## POWERFUL APIs

Out-of-the-box REST and GraphQL endpoints and browser CQL shell.



## 5 GB FREE TIER

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

# Simplify Application Development

# Familiar Language

```
INSERT INTO mytable  
(id,name,address) VALUES  
(1,'Bob Smith','1 Main  
Street')  
SELECT * FROM mytable WHERE  
id=1  
UPDATE mytable SET  
name='Tom Smith' WHERE id=1  
DELETE FROM mytable WHERE  
id=1
```

# Easy Dev Tools

127.0.0.1 (9092) localhost/27040&#039;-39&#039;-4e2c-8e8c-8c239b6cc572

DataStudio

Working with CQL v6.0.0

Resolution Desktop Screen

Horizon

## Viewing Results

Studio provides different ways to display results, including a table view and different styles of charts. Below you can see the table view and a pie chart.

Table results, containing only basic types (nested types coming up a few cells below)

SELECT \* FROM system.local;

index	videntity	username	event	event_timestamp	
1	"900014d6-4612-0400-3f00-7fb7853a500"	root	"900014d6-4612-0400-3f00-7fb7853a500"	stop	20000
2	"900014d6-4612-0400-3f00-7fb7853a500"	root	"900014d6-4612-0400-3f00-7fb7853a500"	start	30000
3	"900014d6-4612-0400-3f00-7fb7853a500"	root	"900014d6-4612-0400-3f00-7fb7853a500"	start	0
4	"900014d6-4612-0400-3f00-7fb7853a500"	root	"900014d6-4612-0400-3f00-7fb7853a500"	stop	20000
5	"900014d6-4612-0400-3f00-7fb7853a500"	root	"900014d6-4612-0400-3f00-7fb7853a500"	start	3000

Displaying 1 - 16 of 36 results for the last statement

Sources: 36 entries (including 3 hidden) Duration: 0.012s

Table results, containing only basic types (nested types coming up a few cells below)

SELECT \* FROM system.local;

index	videntity	username	event	event_timestamp	
1	"900014d6-4612-0400-3f00-7fb7853a500"	root	"900014d6-4612-0400-3f00-7fb7853a500"	stop	20000
2	"900014d6-4612-0400-3f00-7fb7853a500"	root	"900014d6-4612-0400-3f00-7fb7853a500"	start	30000
3	"900014d6-4612-0400-3f00-7fb7853a500"	root	"900014d6-4612-0400-3f00-7fb7853a500"	start	0
4	"900014d6-4612-0400-3f00-7fb7853a500"	root	"900014d6-4612-0400-3f00-7fb7853a500"	stop	20000
5	"900014d6-4612-0400-3f00-7fb7853a500"	root	"900014d6-4612-0400-3f00-7fb7853a500"	start	3000

Displaying 1 - 16 of 36 results for the last statement

Sources: 36 entries (including 3 hidden) Duration: 0.012s

Table results, containing only basic types (nested types coming up a few cells below)

SELECT \* FROM system.local;

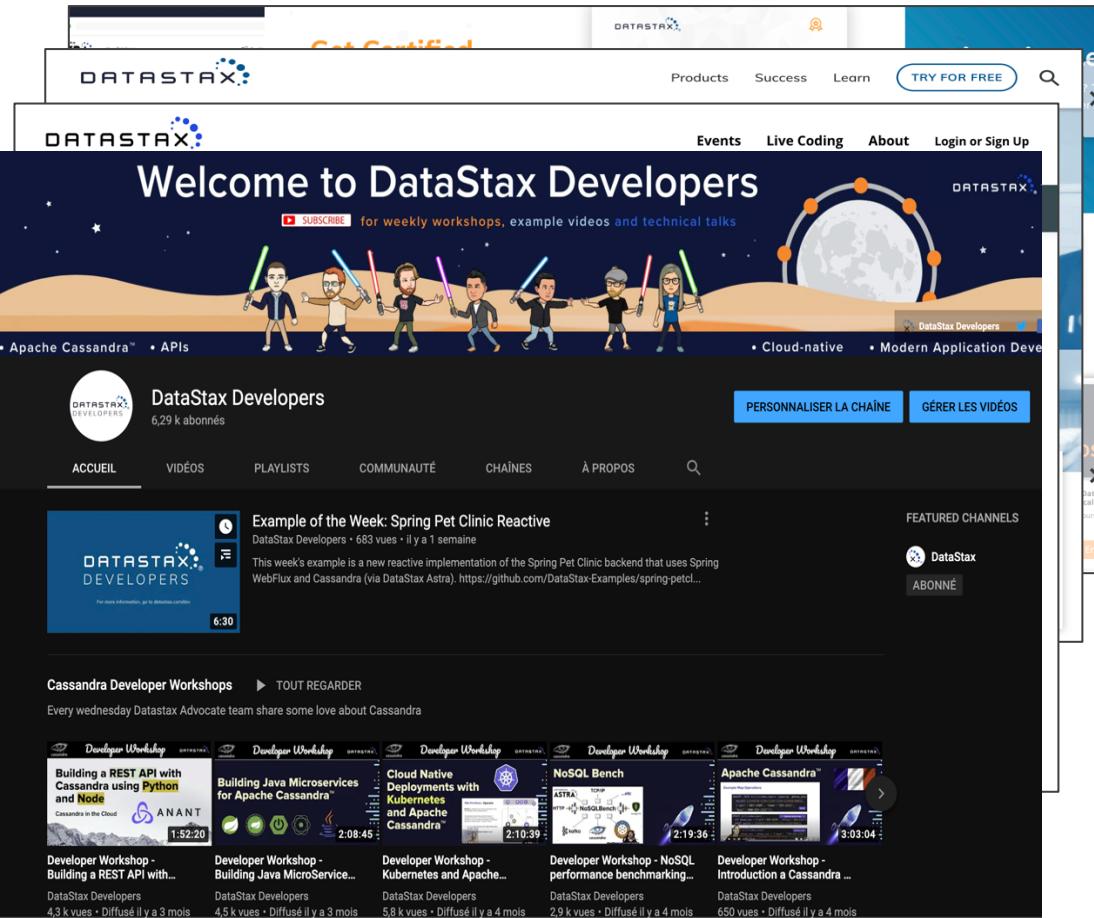
index	videntity	username	event	event_timestamp	
1	"900014d6-4612-0400-3f00-7fb7853a500"	root	"900014d6-4612-0400-3f00-7fb7853a500"	stop	20000
2	"900014d6-4612-0400-3f00-7fb7853a500"	root	"900014d6-4612-0400-3f00-7fb7853a500"	start	30000
3	"900014d6-4612-0400-3f00-7fb7853a500"	root	"900014d6-4612-0400-3f00-7fb7853a500"	start	0
4	"900014d6-4612-0400-3f00-7fb7853a500"	root	"900014d6-4612-0400-3f00-7fb7853a500"	stop	20000
5	"900014d6-4612-0400-3f00-7fb7853a500"	root	"900014d6-4612-0400-3f00-7fb7853a500"	start	3000

Displaying 1 - 16 of 36 results for the last statement

Sources: 36 entries (including 3 hidden) Duration: 0.012s

# Great Drivers





- ✓ Academy.datastax.com
- ✓ datastax.com/dev
- ✓ community.datastax.com
- ✓ Datastax Developers  
**YouTube Channel**



# Thank You!