

# Si quieres saber la hora, necesitas un reloj. Si quieres tener dudas, dos



## Madrid Cassandra Users - by DataStax -

Javier Gómez Santos  
SK-N2 Big Data Architect



an **NTT DATA** Company

An aerial photograph of a dense urban skyline, likely New York City, with numerous skyscrapers. A large green circle is overlaid on the left side of the image, containing the text "About Everis" in a white, sans-serif font.

## About Everis

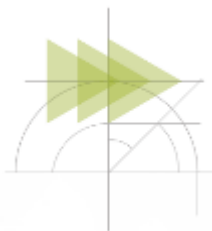


## Información corporativa

**everis an NTT DATA Company** es una consultora multinacional que ofrece soluciones de negocio, estrategia, transformación digital, desarrollo y mantenimiento de aplicaciones tecnológicas y outsourcing.



Más de  
**17.000**  
profesionales



Operamos en  
**15**  
países



Facturación  
**816**  
millones de euros



Argentina - Bélgica - Brasil - Chile - Colombia - EE.UU. - España - Holanda - Italia - Luxemburgo - Marruecos - México - Perú - Portugal - Reino Unido

Cierre a 31 de marzo de 2016 - Facturación grupo **everis** (millones de €)



an **NTT DATA** Company

A low-angle, upward-looking photograph of several modern skyscrapers with glass facades, reaching towards a bright, hazy sky. The perspective creates a sense of height and architectural grandeur.

Sobre mi

## Sobre mi

### Javier Gómez Santos



#### Actualmente:

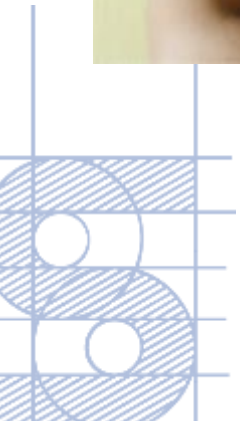
- Big Data Architect @ Everis
- Profesor asociado @ UCM
- Presidente de la asociación Codernauts
- Colaborador @ Pronoide

#### En el pasado:

- Formador técnico @ Pronoide
- Investigador @ UCM
- Backend developer @ Santander, CECA, Orange

#### Advertencia:

- Soy informático, no físico





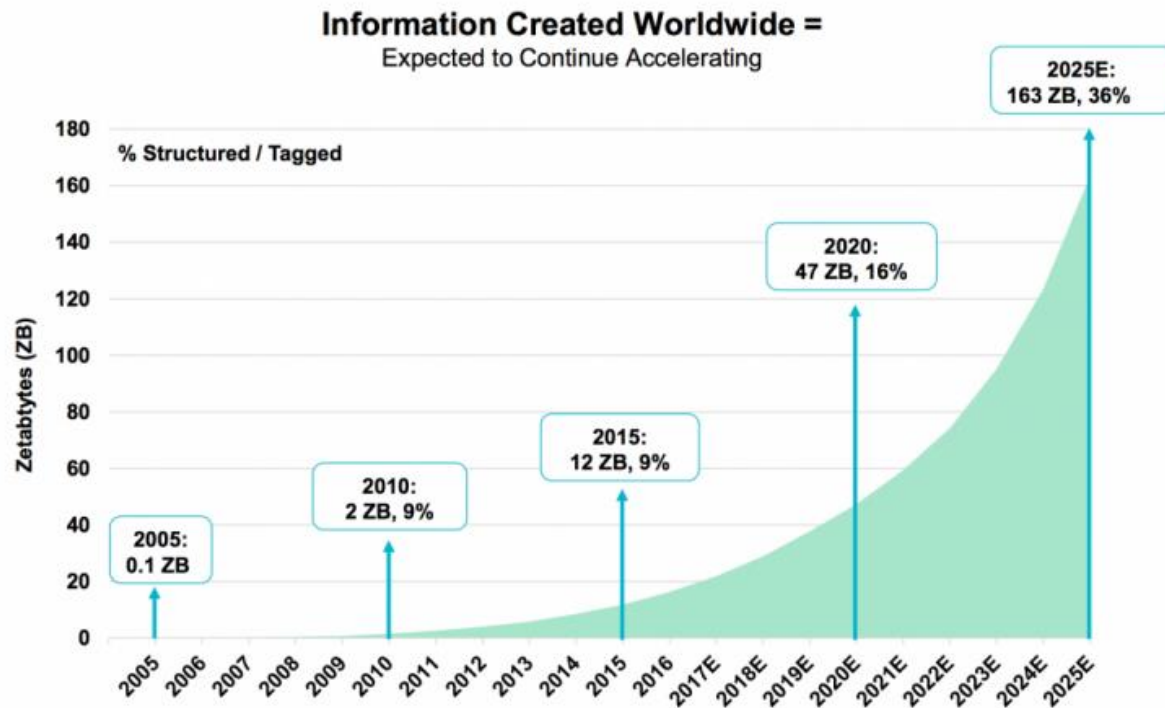
an NTT DATA Company

A low-angle, upward-looking photograph of several modern skyscrapers with glass facades, reaching towards a bright, hazy sky. The perspective creates a sense of height and architectural grandeur.

## Big Data, los orígenes

## Crecimiento de datos en los últimos años

Data Volume Growth Continues @ Rapid Clip...  
% Structured / Tagged (~10%) Rising Fast...





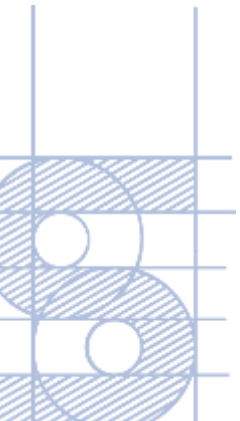
## Imposible para gestionar por un sistema tradicional





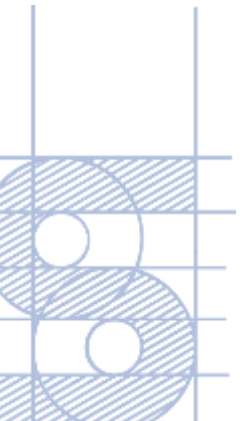
## ¿Qué enfoques aportan las BBDD NoSQL?

- Distribuye la información a través de múltiples nodos
- Es más laxo con la consistencia
- Es más laxo con los esquemas
- Busca optimizar la información para adaptarla a las necesidades actuales



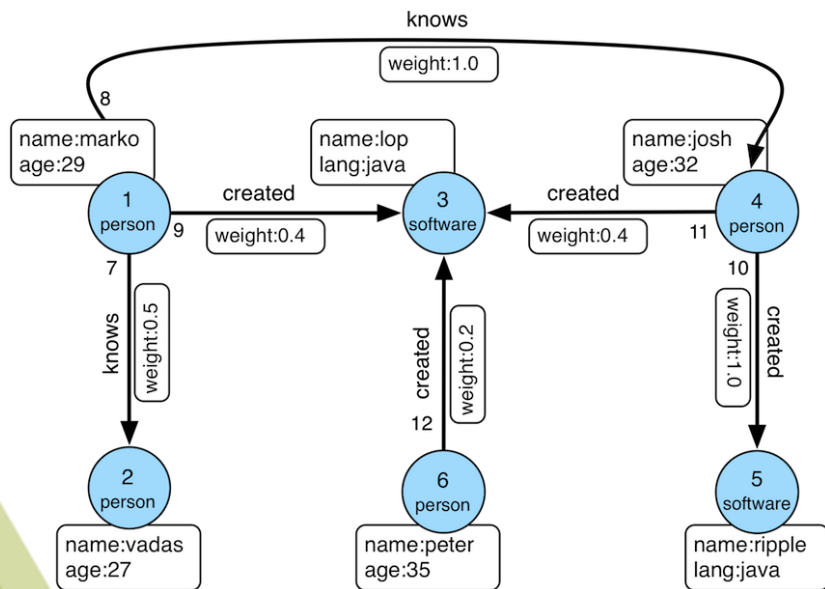
### Hay 4 grandes enfoques a las bases de datos no-relacionales:

- Grafos
- Clave-Valor
- Documento
- Column Family



## Grafos

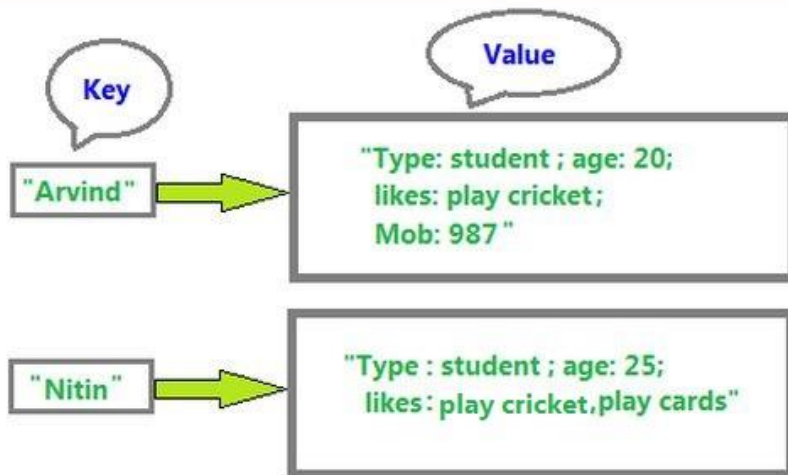
Los elementos están distribuidos como vértices de un grafo, siendo las uniones entre vértices las relaciones de ambos (con propiedades).



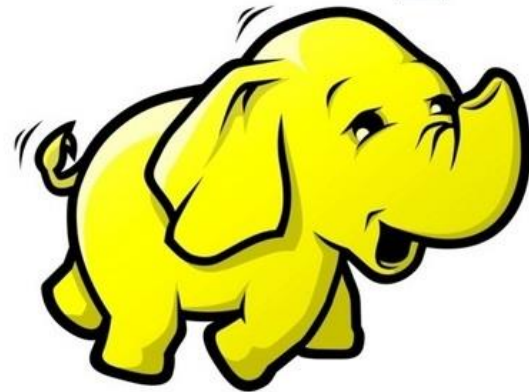
**Gremlin**  
 $G = (V, E)$

## Clave-Valor

Existen claves que enlazan a un conjunto de valores asociados a la misma de cualquier tipo



**hadoop**





## Documento

Se almacenan conjuntos de documentos (JSON) que pueden ser consultados en su totalidad o en parte.

### Students Document

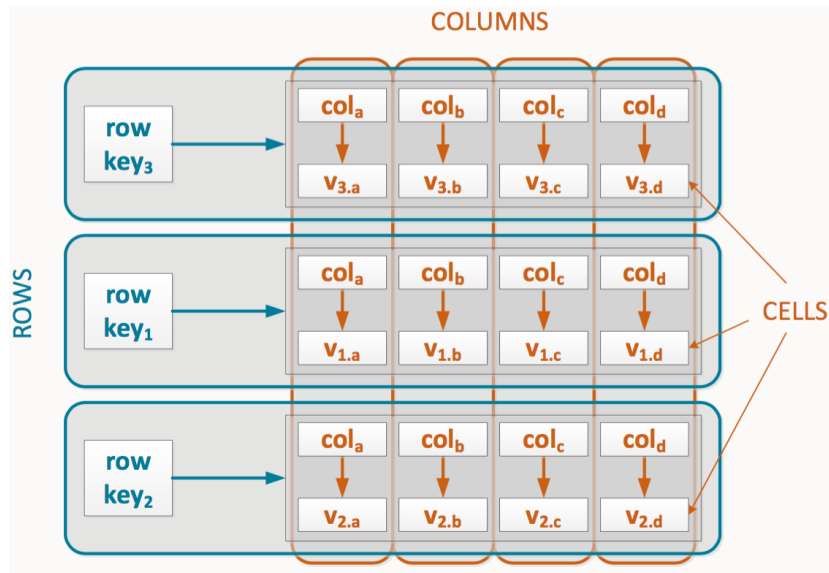
```
{
  _id : "5373aadadcac133aad5b6660",
  name : "kaushal patel",
  email : "kp@example.com",
  courses : {
    course_name : "java",
    fees : "5000",
    duration : "3",
    professor : "g.r."
  }
}
```



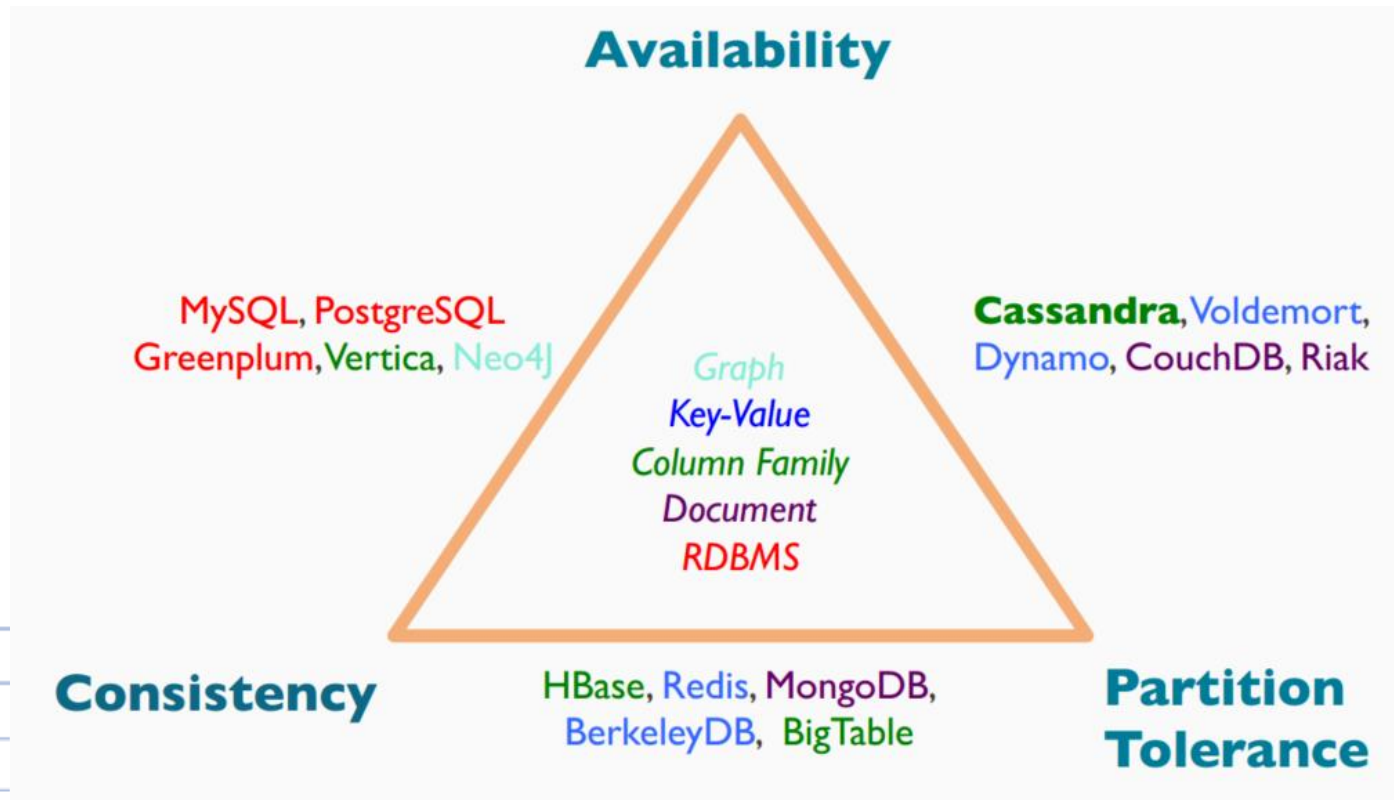
# mongoDB®

## Column Family

Las claves enlazan con un conjunto de columnas



## El teorema CAP





Apache  
Cassandra



## ¿Qué es Cassandra?

### Características

- Una base de datos **NoSQL** basada en **Column Families**.
- Posee un **esquema definido**, pero **flexible**, que permite su alteración.
- Usa un modelo **totalmente distribuido**, sin nodos maestros.

# Cassandra

## ¿Qué es Cassandra?

Posee CQL, un lenguaje de consultas similar al SQL

```
-- Inserts or updates
INSERT INTO Standard1 (KEY, col0, col1)
VALUES (key, value0, value1)
```

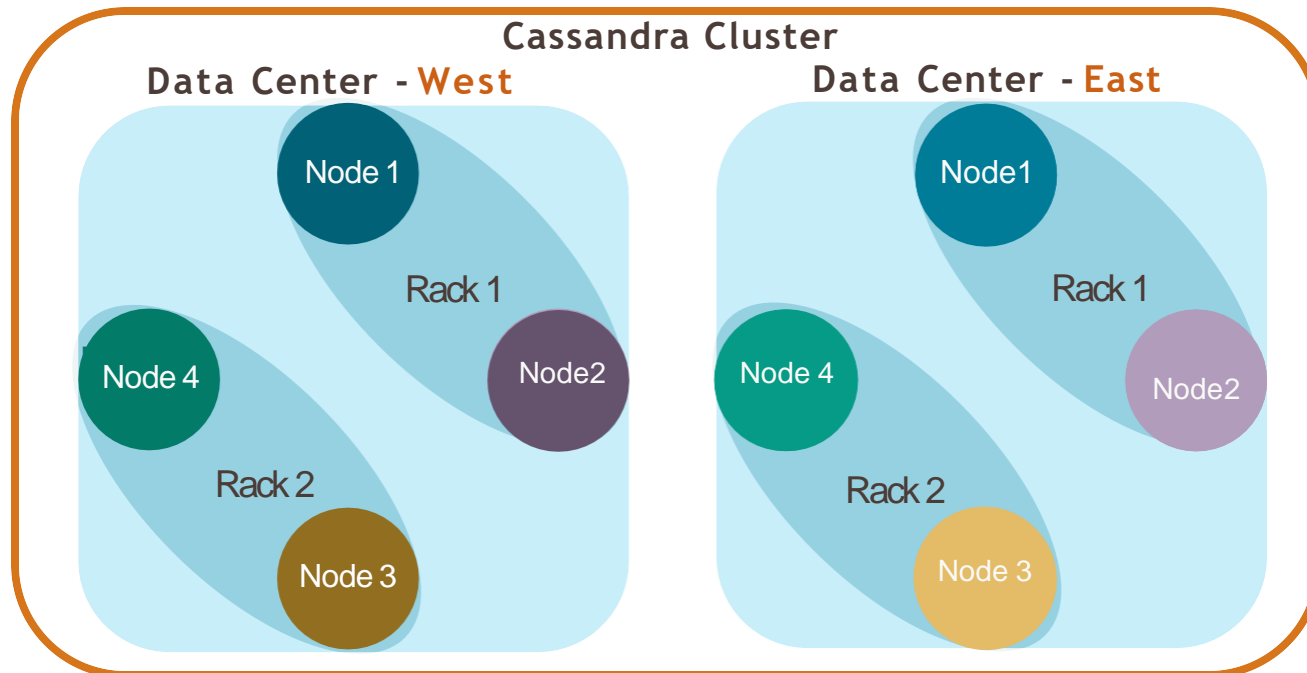
vs.

```
-- Inserts or updates
UPDATE Standard1
SET col0=value0, col1=value1 WHERE KEY=key
```

# Cassandra

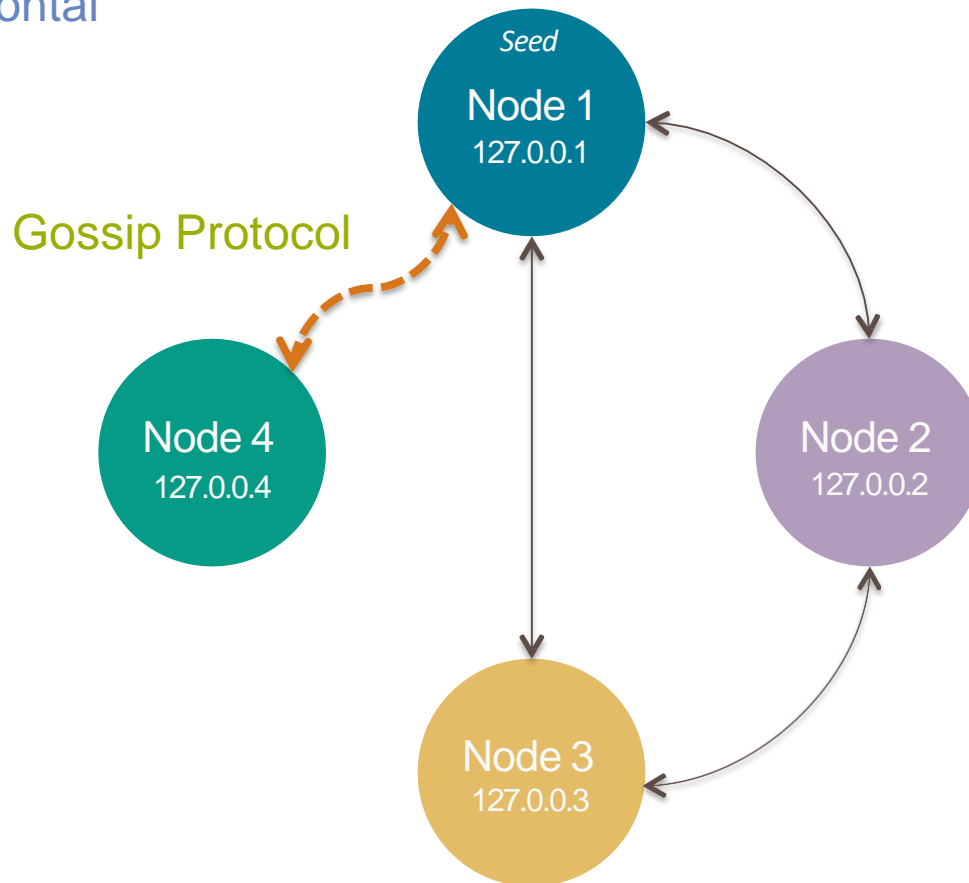
## ¿Qué es Cassandra?

### Clúster Cassandra



## ¿Qué es Cassandra?

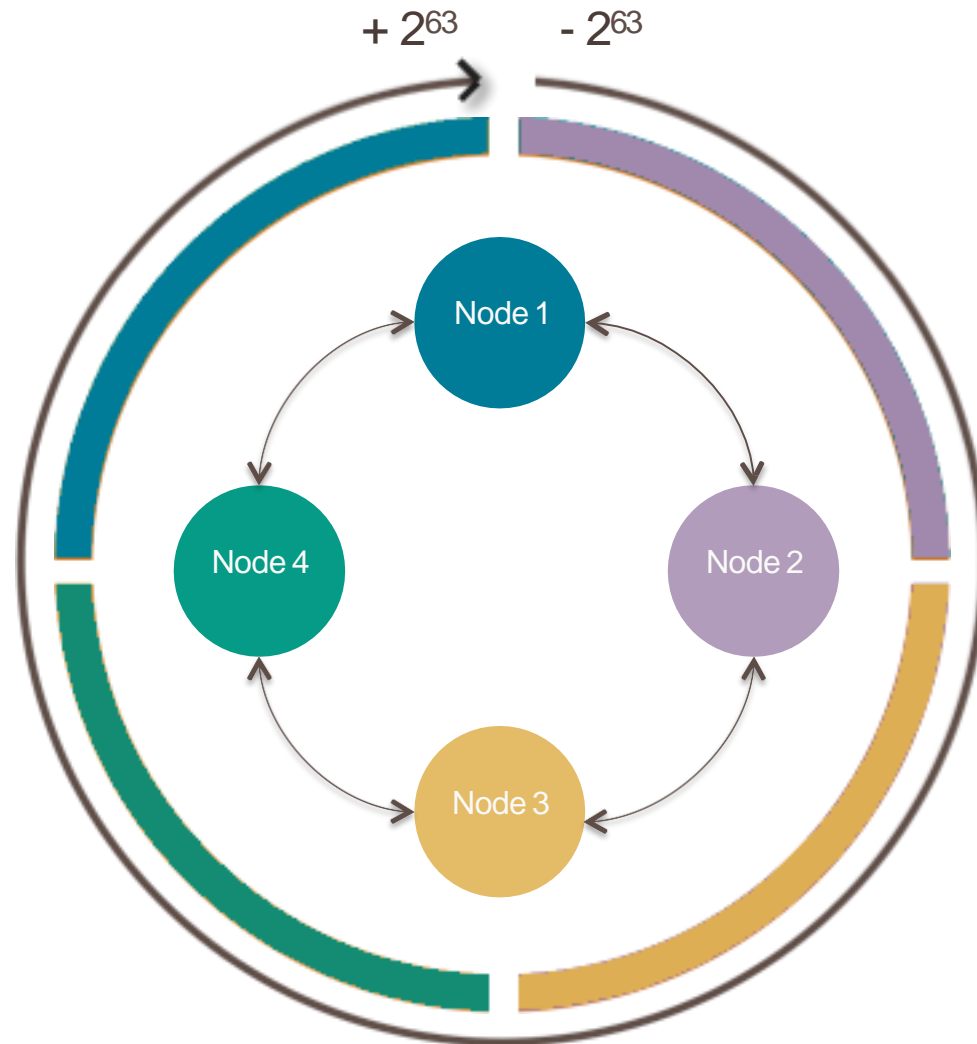
Escalado horizontal





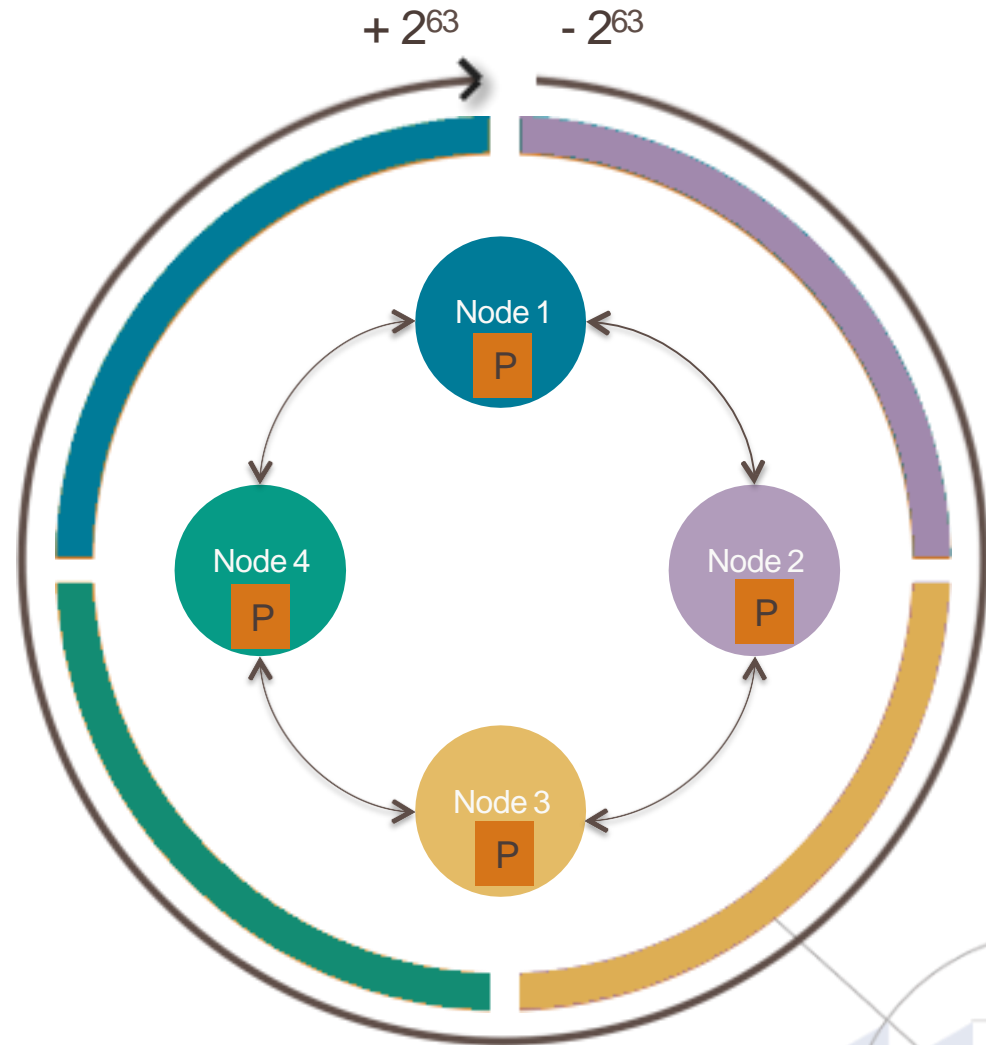
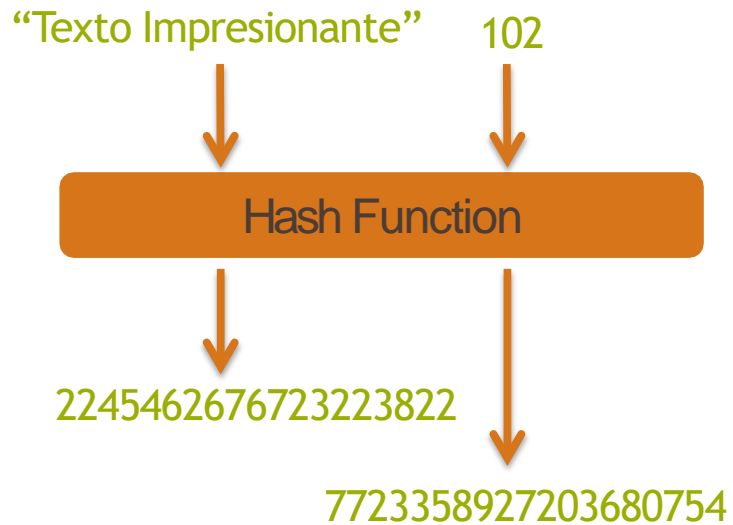
## ¿Qué es Cassandra?

Data sharding



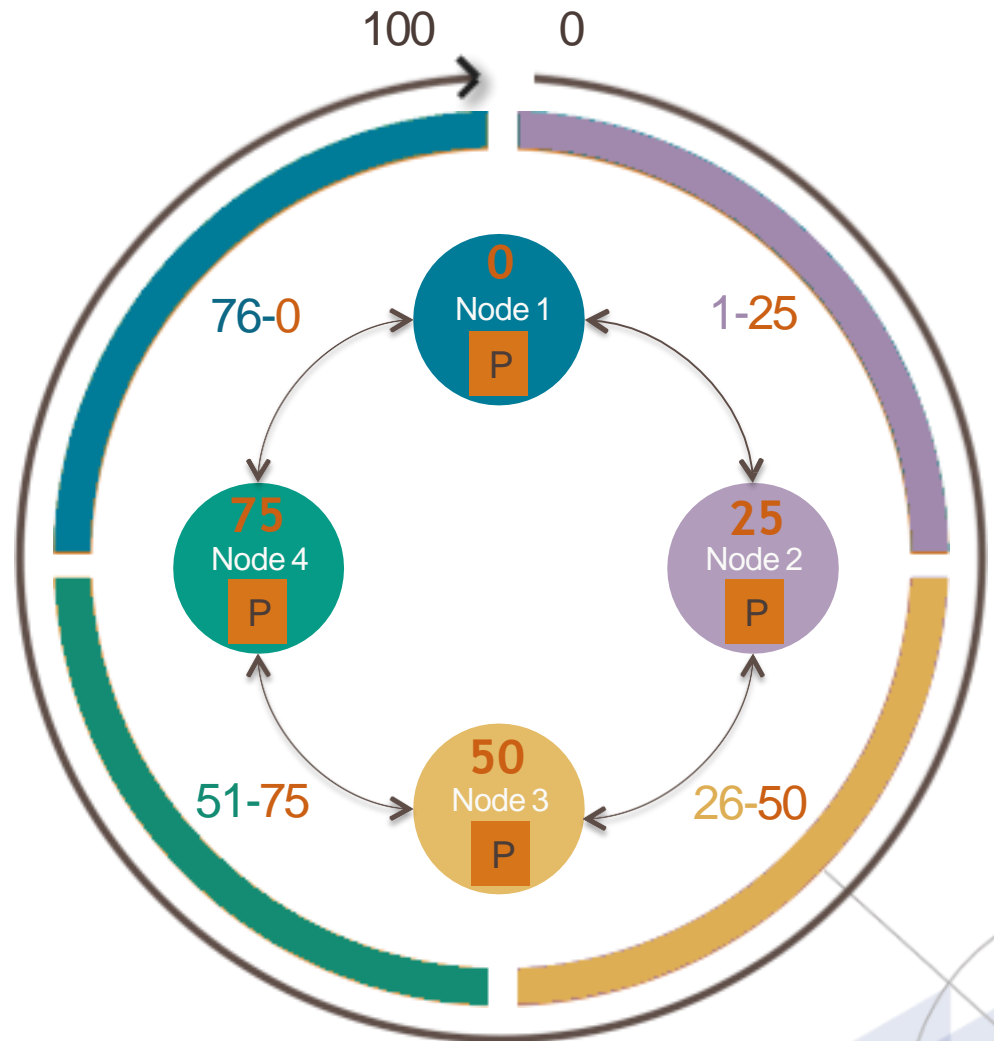
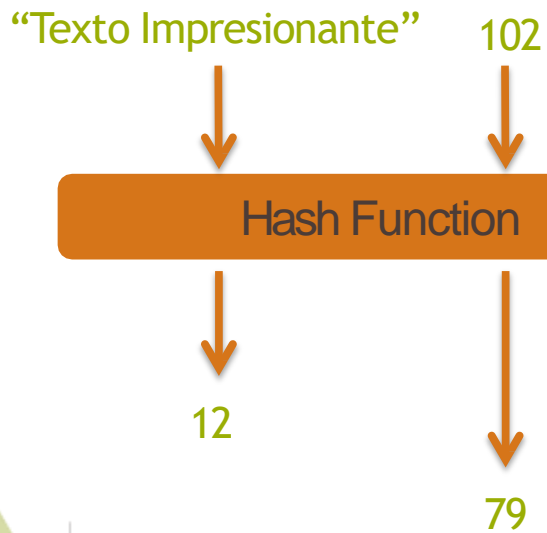
## ¿Qué es Cassandra?

### Data sharding



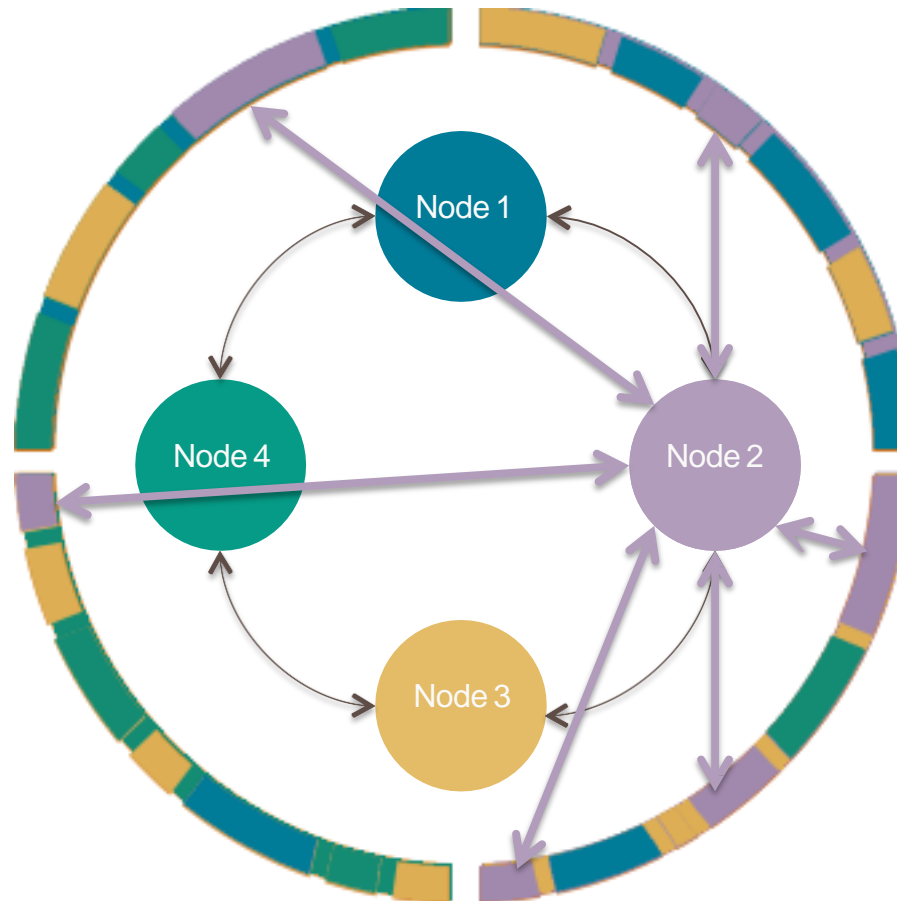
## ¿Qué es Cassandra?

Data sharding



## ¿Qué es Cassandra?

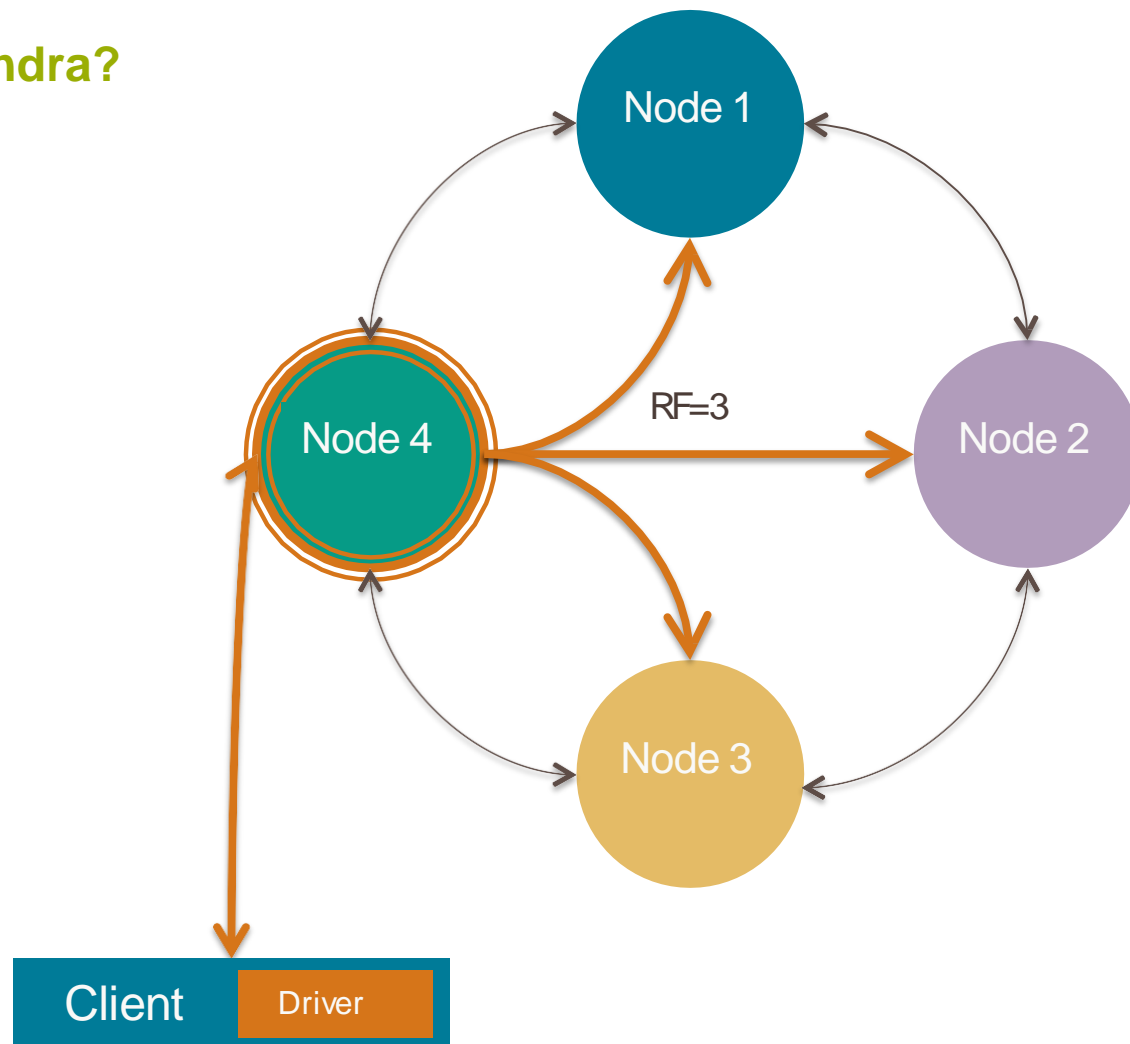
### Data sharding





## ¿Qué es Cassandra?

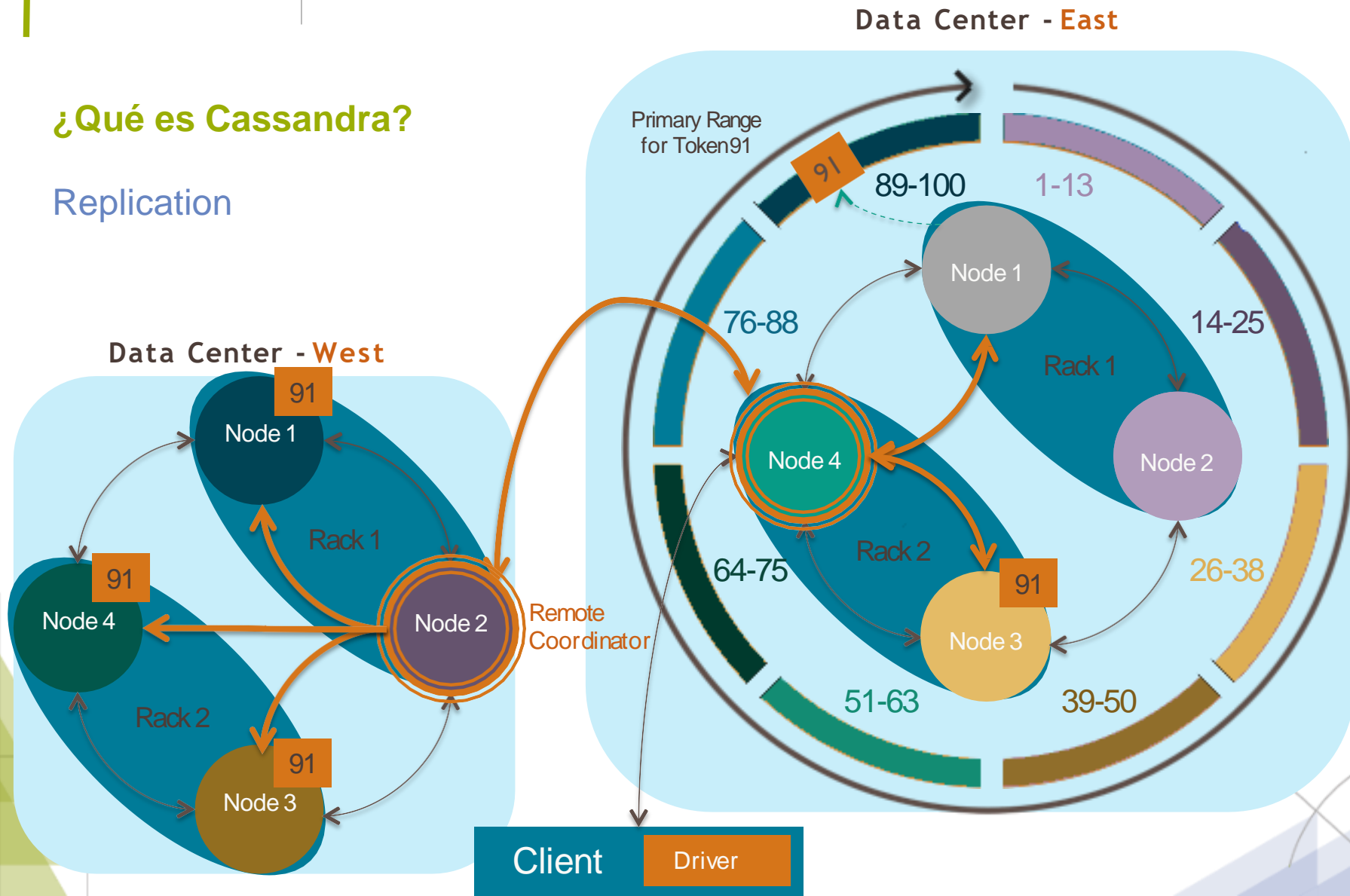
### Coordination



# Cassandra

## ¿Qué es Cassandra?

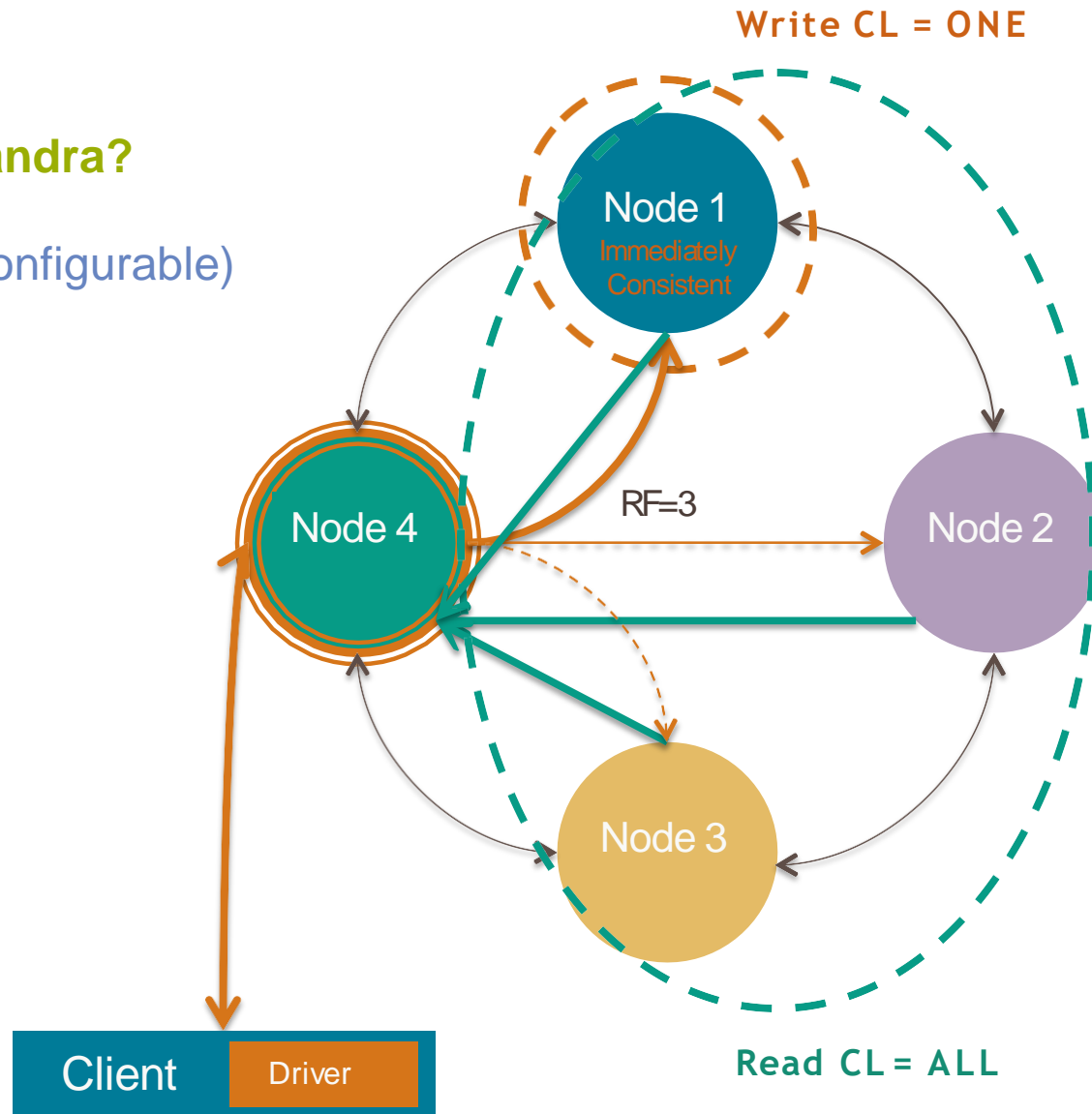
### Replication



# Cassandra

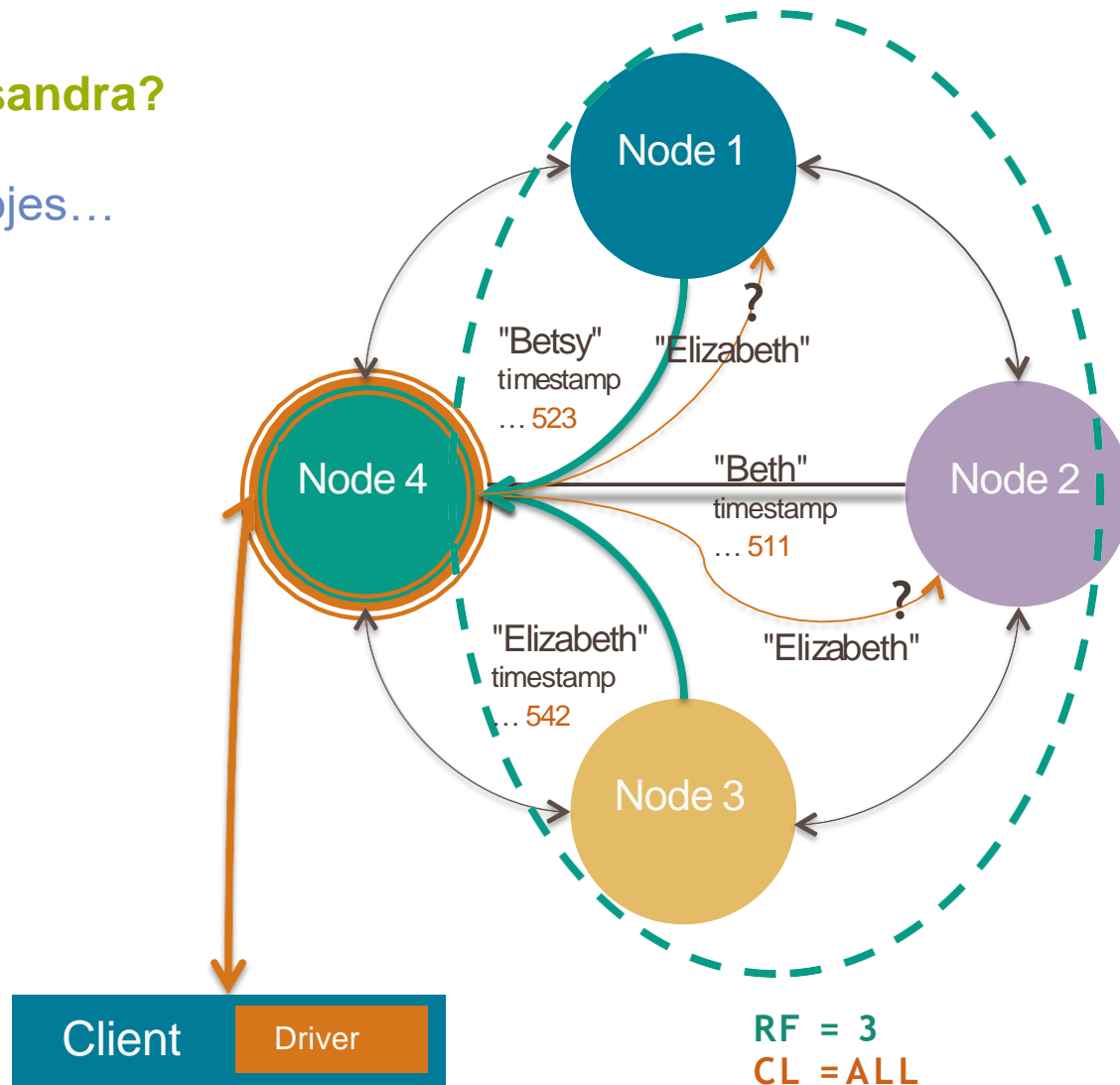
## ¿Qué es Cassandra?

Consistencia (configurable)



## ¿Qué es Cassandra?

Si tienes 2 relojes...



## Hintered handoff

Es un mecanismo de recuperación para escrituras contra nodos caídos:

El coordinador almacena la escritura que queda pendiente sobre el nodo destino si el nodo destino está caído o no responde a la operación

Dicha escritura, se almacena en la tabla `system.hints` (previo a Cassandra 3.0) o en disco (Cassandra 3.0+)

Cuando el nodo afectado se recupera, la escritura se vuelve a solicitar desde el coordinador que había almacenado el hinted handoff

Esto puede ser configurado desde el fichero `cassandra.yaml`

- **hinted\_handoff\_enabled**: Para activar o desactivar el hinted handoff (true por defecto)
- **max\_hint\_window\_in\_ms**: Cuanto tiempo almacenamos hinted handoffs de un nodo caído, después de este tiempo ya no se generan hinted handoffs para el nodo caído (en estos casos, se debe usar alguna herramienta como `repair` para hacer consistentes dichos nodos al volver)

## Zombie Data

Existe una operación de reparación, `nodetool repair`, que repara los datos de un nodo caído, debe ejecutarse si:

- Se ha recuperado un nodo caído (especialmente si ha superado el tiempo indicado por `max_hint_window_in_ms`)
- Un nodo ahogado vuelve a estar activo
- De manera periódica sobre nodos en los que se realizan pocas lecturas (read-repair no va a solucionarlo)
- Sobre nodos con actividad de escritura o borrado

El tema del borrado es delicado en Cassandra, ya que no se borra el dato directamente, se marca como borrado (este marcador se llama `tombstone`), y no se elimina definitivamente hasta pasado un tiempo (llamado `gc_grace_seconds`, por defecto 864000). Es importante hacer reparaciones con mayor frecuencia que dicho borrado, para evitar datos zombies (Por ejemplo, un nodo vuelve con una partición que no ha marcado como eliminada, pero que se marcó como eliminada y se borró del resto de nodos, como no tiene nada con lo que comparar, replica la partición y el dato ha vuelto a la vida).

Se puede modificar el `gc_grace_seconds` en cada tabla con el comando:

```
ALTER TABLE tabla WITH gc_grace_seconds=350000;
```



FIN

Gracias por venir

**Si quieres saber la hora, necesitas un reloj. Si quieres tener dudas, necesitas dos.**

Cassandra, escalabilidad y consistencia



Javier Gómez Santos



[javier.gomez.santos@ucm.es](mailto:javier.gomez.santos@ucm.es)



@Javi\_Pronoide



<https://www.linkedin.com/in/javier-g%C3%B3mez-santos-7b182939/>



<https://github.com/gatchan00>