

# Introducción a Cassandra con Node.js

# Bases de datos

Se pueden clasificar según el nivel de estructura de sus datos.

- Relacionales
  - Altamente estructurados
  - Requiere determinar primero la estructura
  - Si se cambia la estructura de uno de los datos, puede perjudicar todo el sistema, ya que las tablas están relacionadas.



# Bases de datos

Se pueden clasificar según el nivel de estructura de sus datos.

## No Relacionales

- Semi estructurados
- No es necesario definir por completo la estructura
- La información se puede guardar en
  - Documentos
  - Graficos
  - Clave-Valor
  - En-Memoria
  - Búsqueda



Cassandra



redis



CouchDB

# NoSQL



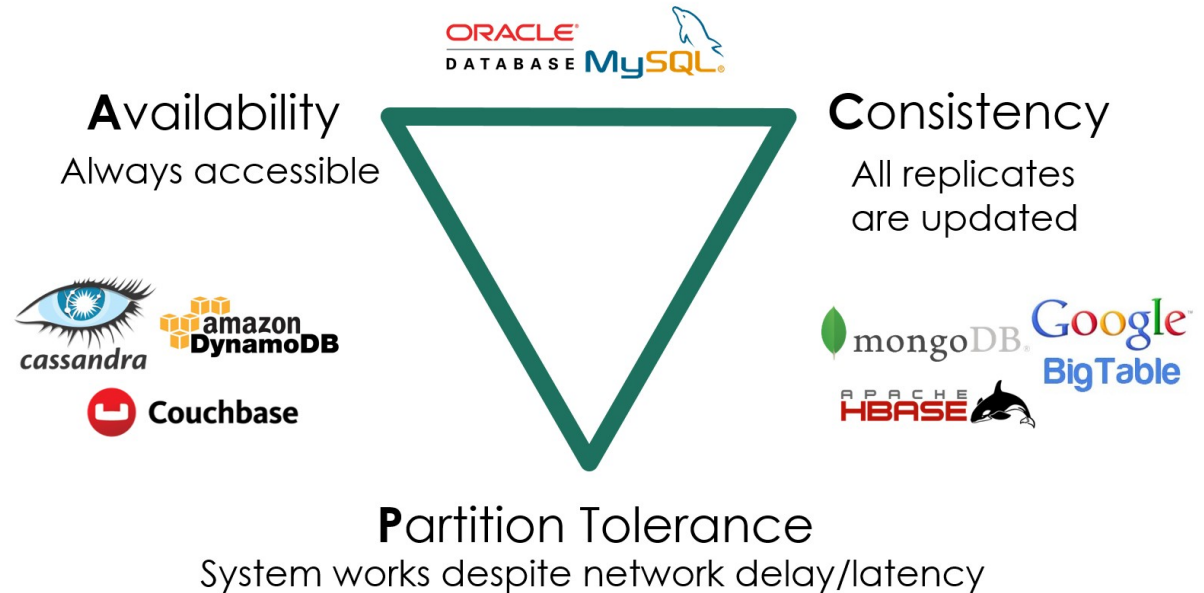
# Escala de Bases de datos según la estructura de su información

## The data continuum



# CAP: Teorema para bases distribuida

- Consistencia
- Disponibilidad
- Tolerancia de partición



# Características Cassandra

- **Descentralizada**

- No utiliza el paradigma maestro-esclavo. Cada nodo es capaz de presentarse al cliente como una replica completa o parcial de la base de datos.

- **Distribuida**

- La información es distribuida a través de muchos nodos e incluso centros de datos.

- **Escalable**

- Se puede configurar para escalar horizontalmente creando nuevos nodos y agregándolos al clúster

# Características Cassandra

- **Altamente disponible**

- Es tolerante a fallas, la información permanecen disponible incluso si uno o muchos nodos y centros de datos caen.

- **Simular consistencia**

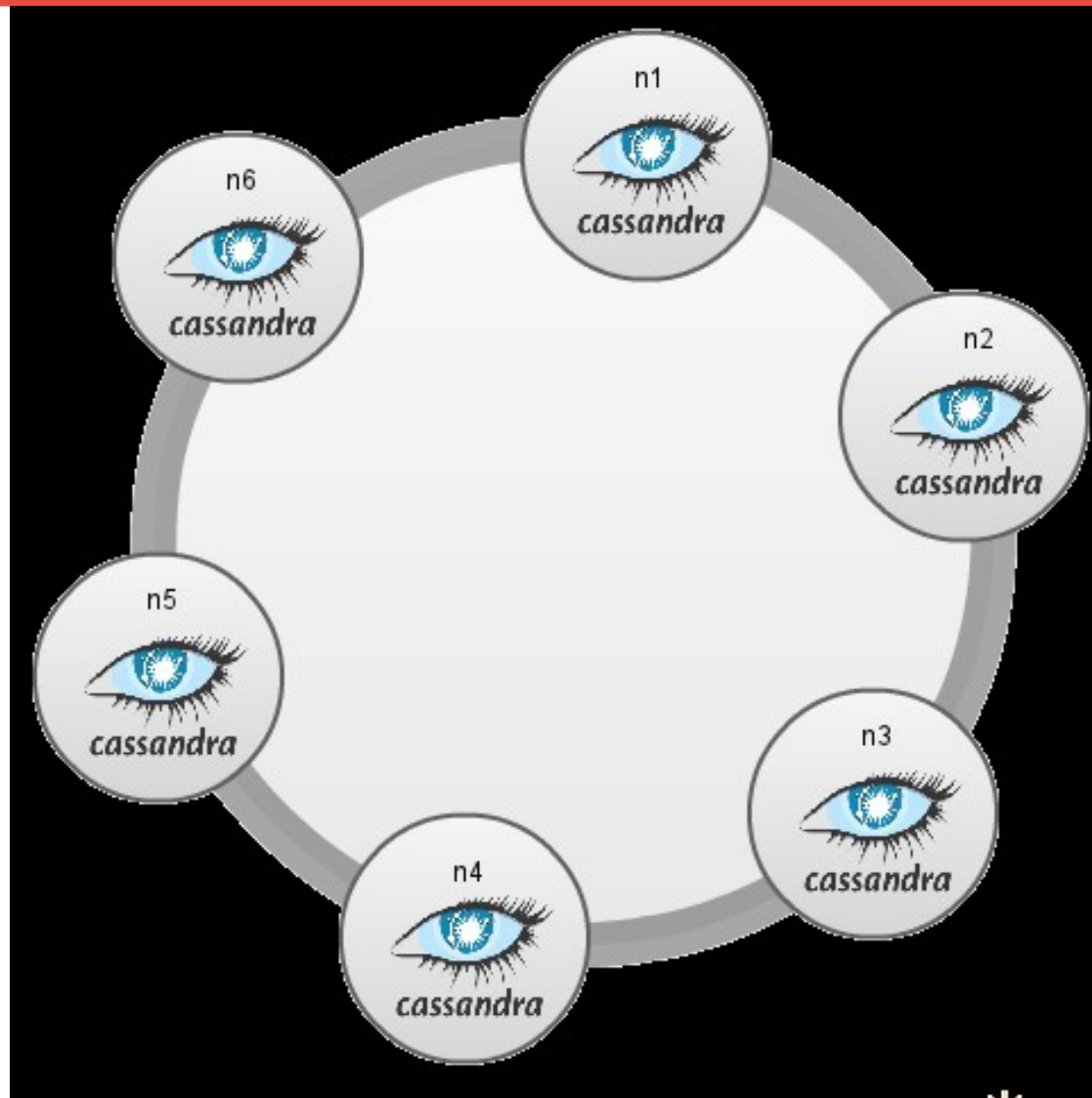
- Se puede ajustar la consistencia mediante el equilibrio entra la disponibilidad y la consistencia de los datos en los nodos de Cassandra, configurando:
  - Factor de replicación: En cuantos nodos replicara la data
  - Nivel de consistencia: En cuantos nodos verificara la consistencia

# Como se distribuye la informacion en Cassandra

- **Utilizan el modelo peer-to-peer (Descentralizada)**
  - Es decir cada nodo actua como servidor de otros, de manera que no necesitan un servidor central.
- **La información se almacena en filas de longitud variable, estructuradas por claves de partición.**
- **Se almacena en diferentes zonas disponibles o múltiples centros de datos.**
- **Ofrece AP**
  - Disponibilidad
  - tolerancia de partición



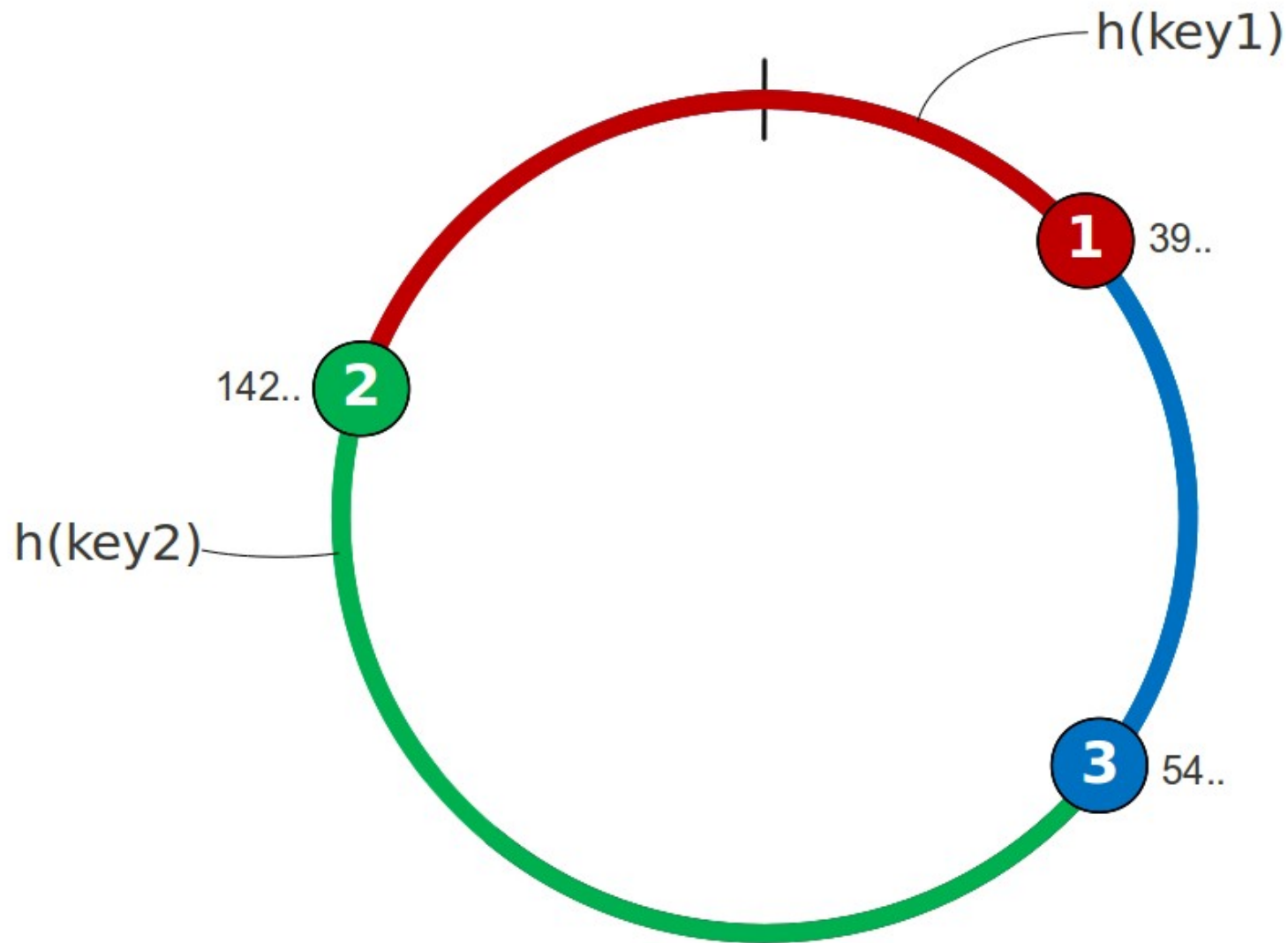
# Modelo peer-to-peer



# Como se distribuye la informacion en Cassandra

- Casandra usa tokens para determinar que nodo almacena que data.
- Un token es un entero de 64 bits
- Le asigna un rango de tokens a cada nodo, de manera que cada token le pertenece a un nodo.
- Por cada nodo que se agrega o quita se redistribuyen los tokens.
- La clave de partición de una fila se usa para calcular el token utilizando una función Hash para determinar el nodo en el cual esta almacenada esa información

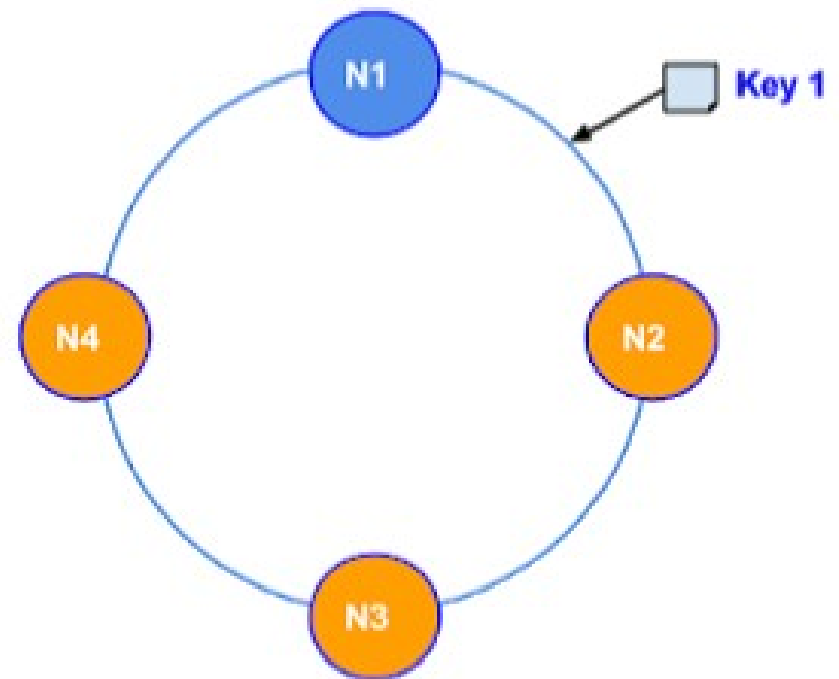
# Partitioning of Keys in the Cluster



# Como se replica la data: Estrategia simple

Los nodos sucesores en sentido del reloj se convierten en las replicas.

SimpleStrategy with RF = 3



Key 1 replicas: {N2, N3, N4}

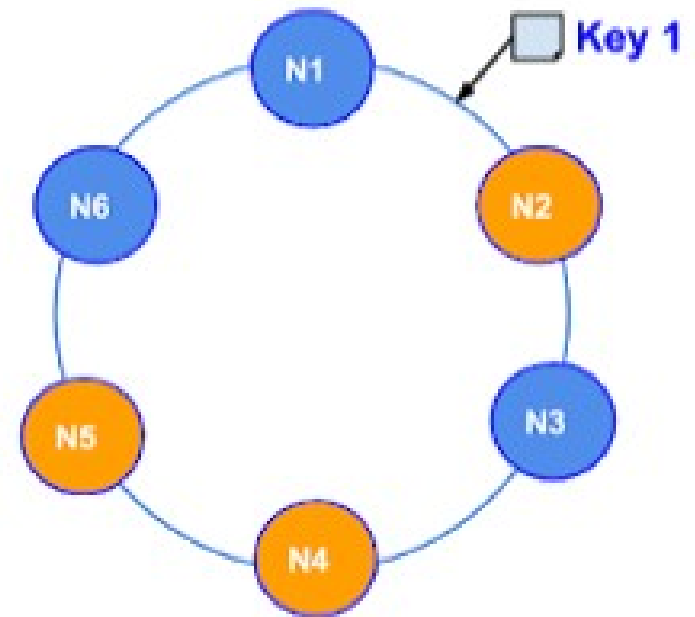
# Como se replica la data:

## Estrategia de Topologia de red

Se escogen como replicas nodos que estan en diferente rack disponible en cada centro de datos.

NetworkTopologyStrategy with Replication factor  
{ DC1: 2, DC2: 2 }

Node	DC	RACK
N1	DC1	RACK2
N3	DC1	RACK1
N6	DC1	RACK1
N2	DC2	RACK1
N4	DC2	RACK1
N5	DC2	RACK2



Key 1 Replicas DC1 : {N3, N1} DC2: {N2, N5}

# Lecturas Recomendadas

- <https://www.pragma.com.co/academia/lecciones/bases-de-datos-relacionales-vs.-no-relacionales>
- <https://dzone.com/articles/introduction-apache-cassandras>
- <https://dzone.com/articles/an-introduction-to-apache-cassandra>
- <https://www.datafluent.org/account>
- <https://www.infoq.com/articles/cassandra-mythology/>
- <https://towardsdatascience.com/cap-theorem-and-distributed-database-management-systems-5c2be977950e>
- <https://www.infoworld.com/article/3298539/datastax-review-cassandra-made-faster-and-easier.html>
- <https://docs.datastax.com/en/articles/cassandra/cassandrathenandnow.html>

## Lecturas Recomendadas

<https://www.ibm.com/developerworks/library/os-apache-cassandra/index.html>

<https://www.slideshare.net/Stratio/nodejs-and-cassandra>