



Clase 19 - Repaso web services - Introducción NoSQL
Profesora: Erika Gutiérrez Beltrán

Tema 3: Bases de datos NoSQL



Objetivo

Construir bases de datos para datos no estructurados, comprender el funcionamiento y aplicación de estos tipos de bases de datos



Temario

- Arquitecturas de bases de datos manejadoras de grandes volúmenes de datos
- Bases de datos bajo modelos Key - Value
- Bases de datos orientadas a documentos
- Bases de datos columnares
- Bases de datos orientadas a grafos



Logros

Interactuar con grandes volúmenes de datos, realizar consultas e interactuar con los datos



Recordando conceptos

- ¿Cuáles son las diferencias entre un dato estructurado y no estructurado?
- ¿Cuál es la diferencia entre dato e información?
- ¿Cuál es la diferencia entre el modelo conceptual, el modelo lógico y el modelo físico?
- ¿Qué tipos de modelos aprendiste en clase?
- ¿Cuál es la diferencia entre una entidad y una entidad débil?
- ¿Qué es una clave parcial y una clave principal?
- ¿Qué es un atributo derivado?
- ¿Qué cardinalidades pueden tener los modelos?
- ¿Qué es una tupla?
- ¿Qué es una base de datos?
- ¿Tipos de arquitectura que pueden tener las bases de datos?
- ¿Diferencias entre OLTP y OLAP?
- ¿Cuál arquitectura requiere mayor cantidad de recursos computacionales?
- ¿Cómo describirías las propiedades ACID?
- Nombra las diferencias entre un EQUI JOIN y un THETA JOIN
- Nombra las diferencias entre un Natural JOIN y un INNER JOIN



Recordando sobre web services

Vamos a la guía de la clase 14 ubicada en el repositorio y vamos nuevamente a revisar los servicios previamente creados, validaremos cuales nos faltan y vamos a crear 3 de los que se proponen en la guía para recordar como crear un endpoint que luego podrá ser utilizado por un sitio web

<https://github.com/ejisselgb/Bases-de-datos-EAFIT/blob/main/Clase14%20-%20Practica%20SQL.ipynb>

Aprovecha este espacio para resolver problemas y dudas que tengas con respecto al manejo de GIT.



Introducción a las bases de datos NoSQL

Este paradigma de las bases de datos NoSQL nace a partir de la necesidad de almacenar grandes volúmenes de información, los cuales crecen a velocidad.

En el año 2015 se proyectaba un crecimiento en los centros de datos de casi un zetabyte (1.099.511.627.776 gigabytes), lo que corresponde a 4.8 zetabytes para ese año.

El tipo de información a procesar cada día es más compleja de analizar y los datos son más crudos. Aún así las aplicaciones deben seguir siendo rápidas y no afectar su funcionamiento por el procesamiento de grandes cantidades de datos.



Introducción a las bases de datos NoSQL

Características de las bases de datos NoSQL:

- **Escalabilidad horizontal:** facilidad de añadir, eliminar y realizar operaciones sin afectar el rendimiento. (Añadir más nodos a un sistema)
- **Escalabilidad vertical:** añadir más capacidad al nodo (RAM, CPU, Almacenamiento)
- **Habilidad de distribución:** Posibilidad de replicar y distribuir datos entre servidores
- **Libertad de esquema:** posibilidad de modelar datos como según lo requiera la aplicación, además facilita la integración con lenguajes de programación
- **Modelo de concurrencia débil:** no implementa ACID, las normas para determinar una transacción son dadas por quien diseña el sistema



Introducción a las bases de datos NoSQL

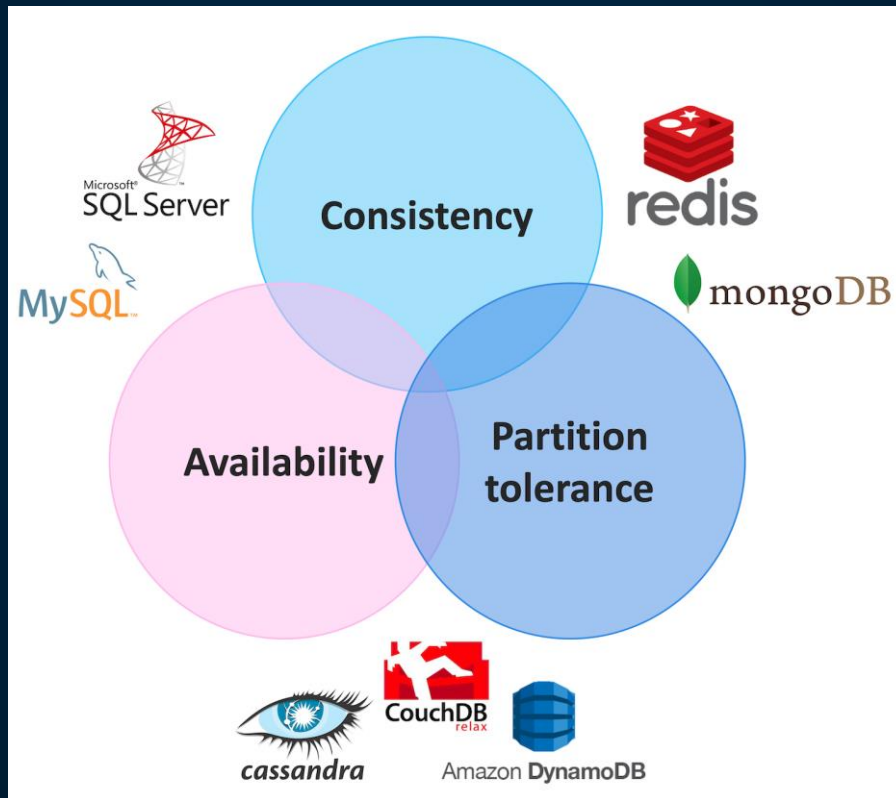
Teorema CAP (Consistency Availability Partition)

Esta propuesta de Eric Brewer parte de la idea de que un entorno distribuido no puede mantener continuamente la consistencia, disponibilidad y tolerancia a la partición de manera simultánea.

- **Consistencia:** posibilidad que tienen los usuarios de acceder a un mismo dato, independientemente del nodo
- **Disponibilidad:** garantía de que cada solicitud recibe una respuesta
- **Tolerancia al particionado:** El Sistema puede seguir en funcionamiento aunque se pierda alguna información



Introducción a las bases de datos NoSQL





Introducción a las bases de datos NoSQL

En el teorema CAP, un sistema distribuido solo puede tener dos de las tres características deseadas



Introducción a las bases de datos NoSQL

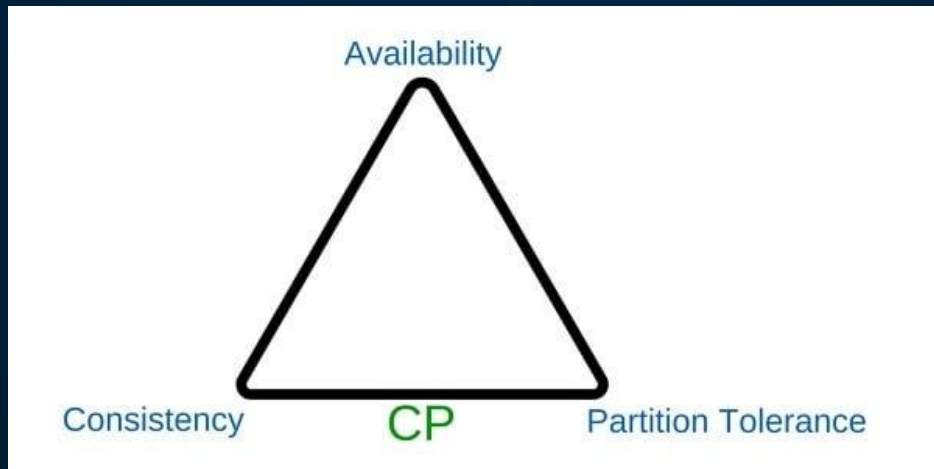
En el teorema CAP, un sistema distribuido solo puede tener dos de las tres características deseadas



Introducción a las bases de datos NoSQL

Tipos de bases de datos del teorema CAP

- **CP**: (Consistencia tolerancia a la partición). El Sistema se encarga de cerrar el nodo no consistente, en caso de existir errores. Queda indisponible hasta solucionar el error.

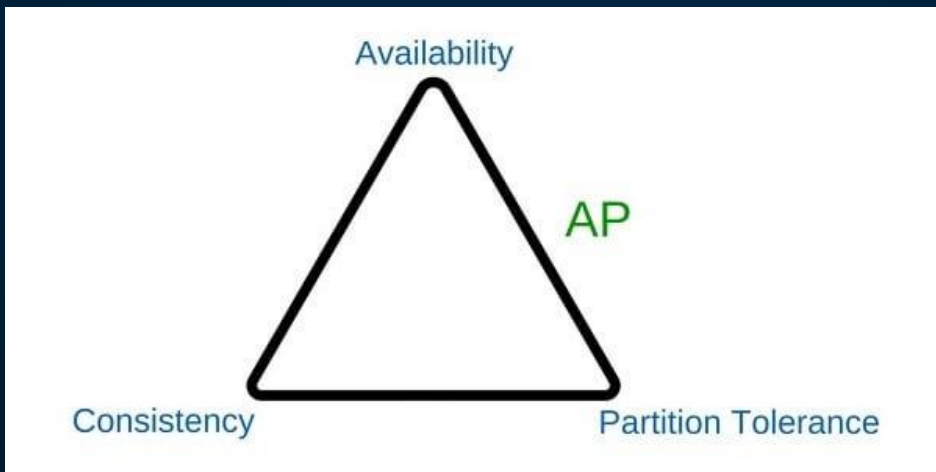




Introducción a las bases de datos NoSQL

Tipos de bases de datos del teorema CAP

- **AP**: (Disponibilidad y tolerancia a la partición) Todos los nodos de una partición estarán disponibles, sin embargo al existir un error una partición puede devolver un version anterior de los datos. Se resincronizan los datos para reparar inconsistencias del Sistema.

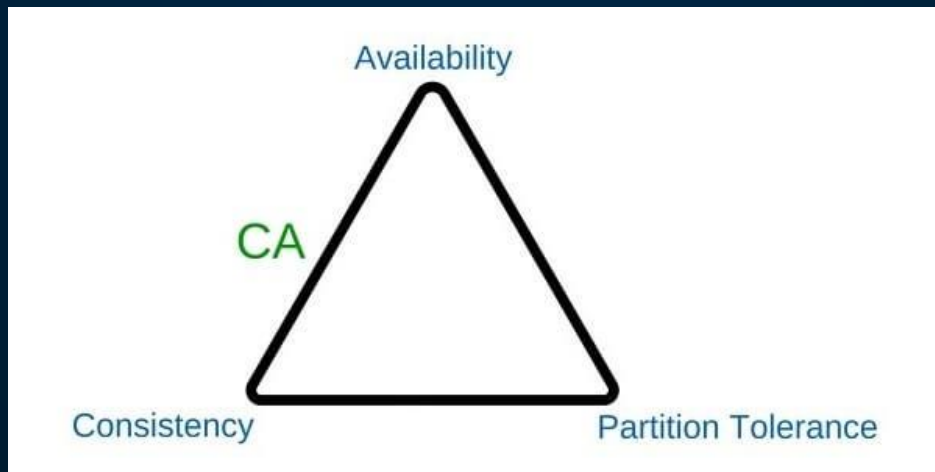




Introducción a las bases de datos NoSQL

Tipos de bases de datos del teorema CAP

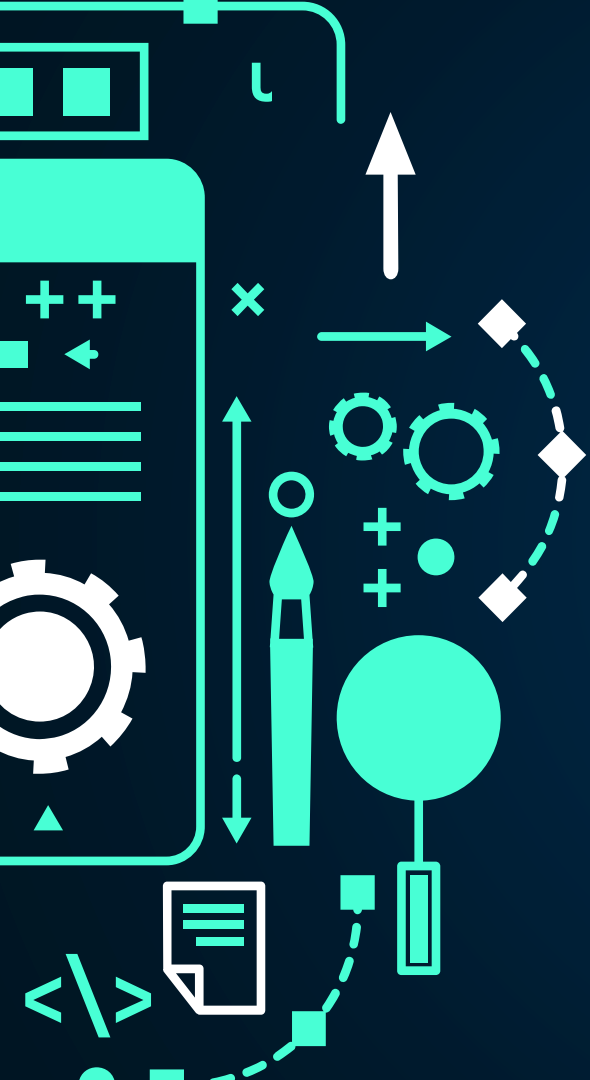
- **CA**: (Consistencia, disponibilidad) No ofrece tolerancia a errores, no se permite la pérdida de comunicación entre nodos.





REFERENCIAS

- Artículo: Utilidad y funcionamiento de las bases de datos NoSQL, Universidad pedagógica y tecnológica de Colombia, 2012
- Teorema CAP: <https://www.ibm.com/co-es/cloud/learn/cap-theorem>



Gracias!