

# Prueba Corta 3 y 4

Gabriela Gutiérrez Valverde - 2019024089

## Explique el concepto de Write Concern en MongoDB.

- El **Write Concern** funciona para dar un nivel de durabilidad a los datos. En MongoDB consiste en garantizar que la operación de escritura sea propagada por todos los nodos. Tiene la consecuencia de que aumenta el tiempo de espera.

## Explique diferencias entre bases de datos NoSQL y SQL.

- La primera diferencia es que las Bases de Datos SQL almacenan los datos de manera estructurada (como tablas, con todos sus atributos definidos), mientras que las Bases de Datos NoSQL no tienen una estructura definida o puede ser variada (como un documento json).
- Las Bases de Datos SQL proporcionan mucha integridad, por la consistencia en los datos, sin embargo en las Bases de Datos NoSQL no es un fuerte, pero estas últimas proporcionan lecturas rápidas. Esto es debido a la atomicidad de las operaciones.
- Las Bases de Datos NoSQL tienen una alta capacidad de escalabilidad además que se adaptan fácilmente a los cambios. En diferencia, las Bases de Datos SQL tienen una capacidad de escalabilidad baja y no son fácilmente adaptadas a los cambios, es un proceso que se puede volver muy complejo.
- Las Bases de Datos SQL se recomiendan para cuando no hay una cantidad de datos extremadamente grande, mientras que las Bases de Datos NoSQL sí están pensadas para grandes volúmenes de información (como por ejemplo, su uso para el Big Data), ya que son distribuidas.

## ¿Porqué la localidad de datos es relevante para la escogencia del hardware a utilizar?

La localidad de datos es relevante por diferentes razones. Primero, algunas veces los datos deben estar en una región específica, por ejemplo si es servicio del gobierno afecta el dónde se almacenan los datos y el tipo de manejo. Luego, regiones como US East (N. Virginia) son más viejas, pero también son considerablemente más baratas y suficientemente estables, lo cual puede ser un punto considerable para un proyecto. Ahora, esto también tiene que ver con los data tiers ya que por ejemplo para Hot, Warm, Cold y Content se es necesario modificar los datos y se debe tener un hardware que de una buena respuesta, pero para Warm, no se necesita tanto poder de procesamiento pero si bastante capacidad de memoria. Por último, para Cold y Frozen no es necesario tener réplicas ya que se conservan en un lugar de almacenamiento barato que es lento pero muy confiable porque los respaldos están en disco.

## Explique el concepto de Federated Queries y el impacto que tienen estas en el rendimiento de bases de datos.

- **Federated Queries:** es la forma en la que se puede enviar una consulta para unir tablas de diferentes "Datasets". Es decir, mientras que tradicionalmente las consultas se hacen a una sola Base de Datos específica, en el caso de los Federated Queries se puede dirigir hacia bases diferentes al mismo tiempo para

obtener una sola tabla que combina los resultados de ambas. En cuanto al impacto de redimiento, estas consultas si pueden tener un tiempo de respuesta más grande, ya que se debe esperar a que se haga la consulta a la base de datos externa y los datos se muevan de forma temporal.