



Actividad 3

Conceptos y Comandos básicos
del particionamiento en bases de
datos NoSQL

Yesid Alexander Jiménez Vivas

**Corporación Universitaria
Iberoamericana**

Facultad de Ingeniería
Bases de datos avanzadas

Ingeniería de Software

Mtro. Jorge Castañeda

Bogotá, Colombia

15 de diciembre de 2024

INTRODUCCIÓN

El particionamiento horizontal, o sharding, es una técnica esencial para la gestión de bases de datos distribuidas, especialmente en entornos que manejan grandes volúmenes de datos, como los asociados a eventos deportivos. En este tipo de sistemas, la capacidad de escalar de manera eficiente y garantizar un rendimiento óptimo en tiempo real son factores cruciales para ofrecer una experiencia de usuario fluida y confiable.

Este documento presenta los scripts necesarios para implementar el particionamiento horizontal (sharding) en una base de datos MongoDB, diseñada para gestionar los datos de un evento deportivo. A través de estos scripts, se configura un entorno distribuido que divide los datos en múltiples servidores, lo que permite que el sistema maneje eficientemente tanto la creciente cantidad de datos como el aumento de usuarios que interactúan en tiempo real con la plataforma.

Los comandos aquí presentados incluyen desde la configuración de los servidores de configuración y shards, hasta la definición de la clave de particionamiento y la verificación del estado del clúster. El objetivo de este documento es proporcionar los pasos detallados para establecer un entorno de sharding en MongoDB que cumpla con los requerimientos de escalabilidad, rendimiento y alta disponibilidad establecidos en los requerimientos no funcionales del sistema.

COMANDOS PARA CONFIGURAR EL ENTORNO DE SHARDING EN MONGODB

1. Iniciar el servidor de configuración:

- `mongod --configsvr --replSet configReplSet --port 27019 --dbpath /path/to/configdb`

2. Iniciar los Shards (servidores de datos):

- `mongod --shardsvr --replSet shardReplSet1 --port 27018 --dbpath /path/to/shard1db`
- `mongod --shardsvr --replSet shardReplSet2 --port 27020 --dbpath /path/to/shard2db`

3. Iniciar el enrutador (mongos):

- `mongos --configdb configReplSet/localhost:27019 --port 27017`

4. Conectar a la base de datos de enrutador:

- `mongo --port 27017`

5. Habilitar el sharding para la base de datos:

- `sh.enableSharding("deporteDB")`

6. Definir la clave de particionamiento (shard key):

Suponiendo que tenemos una colección llamada eventos en la base de datos deporteDB, podemos elegir una clave adecuada para el particionamiento, como timestamp o evento_id.

- `sh.shardCollection("deporteDB.eventos", { "evento_id": 1 })`

7. Verificar el estado del sharding:

- `sh.status()`

RESULTADOS ESPERADOS

Al ejecutar los comandos anteriores, el sistema de MongoDB comenzará a distribuir los datos de la colección eventos entre los diferentes shards, basándose en el valor de la clave de particionamiento. Las consultas que utilicen el `evento_id` o el `timestamp` como filtro deberían ser más rápidas, ya que se dirigirán solo a los shards relevantes, mejorando el desempeño y escalabilidad del sistema.

REFERENCIAS

Sarasa, A. (2016). Introducción a las bases de datos NoSQL usando MongoDB. Editorial UOC.

<https://www.editorialuoc.com/introduccion-a-las-bases-de-datos-nosql-usando-mongodb>.

Aramburu Cabo, M. J. y Sanz Blasco, I. (2012). Bases de datos avanzadas. D - Universitat

Jaume I. Servei de Comunicació i Publicacions, <http://hdl.handle.net/10234/48034>.