

**Actividad 3 - Conceptos y Comandos básicos del particionamiento en bases de
datos NoSQL**

Alberto Mario Lechuga Alean

CORPORACIÓN UNIVERSITARIA IBEROAMERICANA

Ingeniería de software

Bases de datos avanzadas

JORGE CASTAÑEDA

2024

Documento de Requerimientos No Funcionales

Escenario

Para el torneo deportivo de fútbol planteado en la primera actividad, se espera un alto volumen de datos debido a la gran cantidad de equipos, partidos y espectadores. El sistema debe ser capaz de manejar grandes cantidades de datos en tiempo real, asegurando un rendimiento óptimo y una alta disponibilidad.

Criterios de Calidad

- 1. Desempeño:** El sistema debe ser capaz de manejar hasta 10,000 transacciones por segundo.
- 2. Escalabilidad:** La base de datos debe poder escalar horizontalmente para manejar el aumento de datos y usuarios.
- 3. Disponibilidad:** El sistema debe tener una disponibilidad del 99.99%, minimizando el tiempo de inactividad.
- 4. Consistencia:** A pesar de la partición de datos, el sistema debe garantizar la consistencia de los datos en todas las particiones.

Casos de Prueba

1. Escalabilidad

- **Caso de Prueba 1:** Insertar un número significativo de nuevos deportistas y verificar que el rendimiento no se degrade.
- **Criterio de Aceptación:** El tiempo de respuesta de las consultas no debe aumentar significativamente con el crecimiento de datos.

2. Rendimiento

- **Caso de Prueba 2:** Realizar consultas frecuentes sobre los encuentros deportivos y verificar que la carga se distribuye equitativamente entre los nodos.

- **Criterio de Aceptación:** La distribución de carga debe ser uniforme en todos los nodos.

3. Disponibilidad

- **Caso de Prueba 3:** Detener uno de los nodos del clúster y verificar que las consultas continúan ejecutándose sin interrupciones.

- **Criterio de Aceptación:** El sistema debe ser capaz de manejar la pérdida de un nodo sin afectar la disponibilidad.

Comandos de Particionamiento (Sharding)

Estrategia de Particionamiento

Para cumplir con los requerimientos planteados, se utilizará una estrategia de particionamiento basada en el campo `team_id`, que es único para cada equipo. Esto permitirá distribuir uniformemente los datos entre los diferentes shards.

Comandos para Crear el Entorno de Particionamiento

1. Habilitar el Sharding en la Base de Datos:

```
use admin
```

```
db.runCommand({ enableSharding: "torneo_deportivo" })
```

Este comando habilita el sharding en la base de datos llamada `torneo_deportivo`. Esto prepara la base de datos para que pueda ser particionada en diferentes shards.

2. Crear un Índice en el Campo `team_id` :

```
use torneo_deportivo
```

```
db.equipos.createIndex({ team_id: 1 })
```

Este comando crea un índice en el campo `team_id` de la colección `equipos`. El índice es necesario para shardear la colección, ya que MongoDB utiliza este índice para distribuir los datos entre los shards.

3. Shardear la Colección `equipos` :

```
use admin

db.runCommand({
  shardCollection: "torneo_deportivo.equipos",
  key: { team_id: 1 }
})
```

Este comando configura la colección `equipos` para que sea shardeada utilizando el campo `team_id` como clave de particionamiento. Esto significa que los documentos en la colección `equipos` serán distribuidos entre los diferentes shards basándose en el valor de `team_id`.

4. Agregar Shards al Cluster:

```
sh.addShard("shard1/localhost:27018")
sh.addShard("shard2/localhost:27019")
sh.addShard("shard3/localhost:27020")
```

Estos comandos agregan tres shards al clúster de MongoDB. Cada shard es un nodo independiente que almacena una parte de los datos de la colección `equipos`. Esto permite que la carga de trabajo se distribuya entre múltiples servidores, mejorando el rendimiento y la escalabilidad del sistema.

Bibliografía

- Sarasa, A. (2016). Introducción a las bases de datos NoSQL usando MogoDB. Editorial UOC. <https://elibro.net/es/lc/biblioibero/titulos/58524>
- Para el desarrollo de las actividades dentro de la estrategia de aprendizaje # 3 (Particionamiento en Bases de datos NoSQL) se debe realizar como mínimo la lectura de los siguientes recursos bibliográficos:
- Para la actividad 5 (Conceptos y comandos básicos para el Particionamiento en Bases de Datos NoSQL) se invita al estudiante a comprender el Particionamiento de bases datos no relacionales (NoSQL) su importancia a nivel tecnológico en el contexto de Big Data y su diferencia con el modelo de Base de Datos Relacionales. Por lo que se sugiere la lectura de capítulo 3(Bases de datos distribuidas e integración de información distribuida) del libro Aramburu Cabo, M. J. y Sanz Blasco, I. (2012). Bases de datos avanzadas. D - Universitat Jaume I. Servei de Comunicació i Publicacions, <https://elibro.net/es/lc/biblioibero/titulos/51741>
- Adicionalmente, el capítulo VIII (Sharding) del libro de Sarasa, A. (2016) Introducción a las bases de datos NoSQL usando MongoDB. Editorial UOC, <https://elibro.net/es/lc/biblioibero/titulos/58524>