**TRABAJO PRÁCTICO Nº 3 – TIPOS DE BASES DE DATOS**

**BASES DE DATOS NoSQL**

1. **¿Qué son?**

Las bases de datos NoSQL son sistemas de almacenamiento y consulta de datos que se diferencian de las bases de datos SQL tradicionales en su enfoque y estructura. Permiten almacenar y recuperar datos de forma no relacional, diseñadas para manejar grandes volúmenes de datos no estructurados o semiestructurados, como documentos, pares clave-valor, datos de gráficos o columnas. Son ideales para aplicaciones con grandes cantidades de datos distribuidos y escalabilidad dinámica.

1. **Características de las bases de datos NoSQL.**
   1. Modelo de datos flexible: No requieren un esquema rígido
   2. Escalabilidad horizontal: manejo de clusters para manejar grandes volúmenes de datos y tráfico.
   3. Alto rendimiento y baja latencia: manejo de grandes cantidades de datos y operaciones de alta velocidad, ideal para aplicaciones en tiempo real como redes sociales y análisis de big data.
   4. Variedad de tipos de almacenamiento: clave-valor, documentales, columnares y de grafos, etc.
   5. Alta disponibilidad y tolerancia a fallos: ofrece mecanismos de replicación, garantizando la disponibilidad de datos incluso ante fallos.
   6. Consistencia eventual: Muchas bases de datos NoSQL priorizan la disponibilidad sobre la consistencia inmediata, lo que significa que puede haber un retraso en la sincronización de datos entre nodos.
2. **Diferencias con las bases de datos relacionales.**

|  |  |  |
| --- | --- | --- |
|  | Bases de datos relacionales | Bases de datos NoSQL |
| Cargas de trabajo óptimas | diseñadas para el procesamiento de transacciones en línea altamente coherentes y transaccionales. | Diseñadas para varios patrones de acceso a datos que incluyen aplicaciones de baja latencia. Las bases de datos de búsqueda NoSQL están diseñadas para hacer análisis sobre datos semiestructurados. |
| Modelo de datos | Normalización de los datos en tablas conformadas por filas y columnas. | variedad de modelos de datos, como clave-valor, documentos, gráficos y columnas, que están optimizados para el rendimiento y la escala. |
| Propiedades ACID | Ofrecen propiedades de atomicidad, coherencia, aislamiento y durabilidad (ACID) | Puede hacer concesiones al flexibilizar algunas de las propiedades ACID a favor de un modelo de datos más flexible. |
| Rendimiento | Normalmente, el rendimiento depende del subsistema de disco y de la optimización de las consultas | Depende del tamaño del clúster de hardware subyacente, la latencia de red y la aplicación que efectúa la llamada. |
| Escalado | Escalado vertical a las capacidades de computación del hardware | Las bases de datos NoSQL suelen ser particionables. Los patrones de acceso pueden escalar horizontalmente. |
| API | Solicita almacenar y recuperar datos que están comunicados mediante consultas que se ajustan a un lenguaje de consulta estructurado (SQL). | Las API basadas en objetos permiten a los desarrolladores almacenar y recuperar fácilmente estructuras de datos. |

1. **Tipos de bases de datos NoSQL**.
   1. **Base de datos de clave-valor**

Las bases de datos de clave-valor son altamente particionables y permiten un escalado horizontal en niveles que otros tipos de bases de datos NoSQL no pueden lograr. Una base de datos de clave-valor almacena datos como un conjunto de pares clave-valor en los que una clave sirve como un identificador único. Las claves y los valores pueden ser cualquier cosa, desde objetos simples hasta objetos compuestos complejos. Los casos de uso como juegos o tecnología publicitaria se prestan particularmente bien con el diseño de datos de almacén clave-valor. Se suelen usar para almacenar sesiones de usuarios, datos en caché o información en tiempo real.

**Ejemplos**:

* **Redis**: Se usa mucho para almacenar datos temporales, como caché o sesiones.
* **Amazon DynamoDB**: Ideal para manejar cargas masivas con baja latencia, como en aplicaciones móviles y juegos.
  1. **Bases de datos de documentos**

Las bases de datos de documentos tienen el mismo formato de modelo de documento que los desarrolladores utilizan en el código de sus aplicaciones. Almacenan los datos como objetos JSON que son flexibles, semiestructurados y de naturaleza jerárquica. La naturaleza flexible, semiestructurada y jerárquica de los documentos y las bases de datos de documentos permite que evolucionen según las necesidades de las aplicaciones. Los documentos pueden contener datos semiestructurados con diferentes niveles de anidamiento. El modelo de base de datos de documentos funciona bien con catálogos, perfiles de usuario y sistemas de administración de contenido en los que cada documento es único y evoluciona con el tiempo.

**Ejemplos**:

* **MongoDB**: Popular en el desarrollo de aplicaciones web.
* **Couchbase**: Utilizada para aplicaciones móviles y web con alta demanda de sincronización en tiempo real.
  1. **Bases de datos de grafos**

Su objetivo es facilitar la creación y ejecución de aplicaciones que funcionen con conjuntos de datos con un nivel alto de conexión. Usan nodos para almacenar entidades de datos y periferias para almacenar relaciones entre entidades. Permiten realizar consultas sobre las relaciones entre los nodos de manera rápida. Puede usar una base de datos de grafos para crear y ejecutar aplicaciones que funcionen con conjuntos de datos con un nivel alto de conexión. Los casos de uso típicos para una base de datos de grafos incluyen redes sociales, motores de recomendaciones, detección de fraude y gráficos de conocimiento.

**Ejemplos**:

* **Neo4j**: Se utiliza para análisis de redes sociales y detección de fraudes en tiempo real.
* **Amazon Neptune**: Utilizada para construir gráficos de conocimiento o motores de recomendaciones basados en conexiones entre datos.
  1. **Bases de datos en memoria**

Mientras que otras bases de datos no relacionales almacenan datos en discos o SSD, los almacenes de datos en memoria RAM están diseñados para eliminar la necesidad de acceder a los discos. Son ideales para aplicaciones que requieren tiempos de respuesta de microsegundos o que tienen grandes picos de tráfico. Puede usarlas en aplicaciones de juegos y tecnología publicitaria para funciones como tablas de clasificación, tiendas de sesiones y análisis en tiempo real.

**Ejemplos**:

* **Memcached**: Utilizada para el almacenamiento en caché en aplicaciones web y reducir la carga en bases de datos más lentas.
* **Redis**: Además de su uso como base de datos clave-valor, Redis también es una base de datos en memoria, ideal para operaciones con tiempos de respuesta extremadamente bajos.
  1. **Buscar en bases de datos**

Una base de datos de motores de búsqueda es un tipo de base de datos no relacional que se dedica a la búsqueda de contenido de datos. Utilizan índices para categorizar características similares entre los datos y facilitar la capacidad de búsqueda. Las bases de datos de los motores de búsqueda están optimizadas para clasificar datos no estructurados, como imágenes y videos.

**Ejemplos**:

* **Elasticsearch**: Utilizada para búsquedas en tiempo real y análisis de grandes volúmenes de datos en registros, sitios web y motores de búsqueda.
* **Apache Solr**: Usado en motores de búsqueda y análisis de datos empresariales para realizar consultas rápidas y eficaces.

1. **¿Para qué son más usadas?**

**Bases de datos de clave-valor**

**Redis**: Almacenamiento en caché, manejo de sesiones de usuarios, mensajería en tiempo real y contadores.

**Amazon DynamoDB**: Gestión de grandes volúmenes de datos en tiempo real, con baja latencia, comercio electrónico, sistemas de pago.

**Bases de datos de documentos**

**MongoDB:** Aplicaciones web, catálogos de productos, almacenamiento de datos de usuarios, Plataformas de comercio

**Couchbase:** Aplicaciones móviles y web que necesitan sincronización en tiempo real y alta disponibilidad.

**Bases de datos de grafos**

**Neo4j:** Análisis de redes sociales, motores de recomendación, análisis de fraudes.

**Amazon Neptune:** Motores de búsqueda semántica, sistemas de recomendaciones basados en conexiones (por ejemplo, "clientes que compraron esto también compraron..."),

**Bases de datos en memoria**

**Memcached:** Caché en aplicaciones web, reducción de la carga de bases de datos relacionales, almacenamiento de datos temporales.

**Redis:** Almacenamiento en caché, administración de sesiones, análisis en tiempo real, colas de mensajes, tablas de clasificación en juegos.

**Bases de datos de búsqueda**

**Elasticsearch:** Búsquedas en tiempo real, análisis de grandes volúmenes de datos, procesamiento de logs, motores de búsqueda de sitios web.

**Apache Solr:** Motores de búsqueda empresarial, análisis de datos masivos, indexación de contenido para sitios web.

1. **¿Qué tipo de base de datos llevarían a cabo el almacenamiento en la propuesta BIG DATA?**

Para un sistema de Big Data en CompraGamer usaremos Elasticsearch.

Razones para elegir Elasticsearch:

* Búsqueda rápida y eficiente: Permite realizar búsquedas complejas y análisis de grandes volúmenes de datos no estructurados, ideal para el catálogo de productos y para mejorar la experiencia del usuario en el sitio.
* Análisis en tiempo real: Puede manejar y analizar datos de logs y comportamiento de usuarios en tiempo real, ayudando a tomar decisiones rápidas y optimizar el rendimiento del sitio.
* Escalabilidad: Es altamente escalable y puede adaptarse al crecimiento del negocio sin comprometer la velocidad de búsqueda o el rendimiento.
* Integración fácil: Se integra bien con otras tecnologías, lo que permite construir un ecosistema completo para la gestión de datos.

Esta opción proporciona una base sólida para mejorar la experiencia de usuario y optimizar las operaciones en el e-commerce de CompraGamer.