



TEMA-11-ACCESO-A-DATOS.pdf



user_4383848



Acceso a datos



1º Desarrollo de Aplicaciones Multiplataforma



Estudios España

antes



Descarga sin publi
con 1 coin



Después



WUOLAH

TODOS LOS LADOS DE LA CAMA

14 NOVIEMBRE
SOLO EN CINES

CINEFOX NOWPLAY MOLADES MOLADES MOLADES MOLADES MOLADES MOLADES MOLADES MOLADES

14
NOVIEMBRE
SOLO EN CINES

TODOS LOS LADOS DE LA CAMA

TEMA 11 - ACCESO A DATOS INTRODUCCIÓN A LAS BD NO-SQL

1. CARACTERÍSTICAS DE LAS BASES DE DATOS NO-SQL:

No-SQL se refiere a todas las bases de datos y almacenes de datos que no están basadas en el sistema tradicional de bases de datos relacionales en sí, o no están basadas en sus principios.

Suele relacionarse con grandes conjuntos de datos a los que se accede y se manipulan a escala web. El concepto No-SQL no representa un producto único o una única tecnología, representa un grupo de productos y un conjunto relacionado de conceptos para el almacenamiento y su administración.

Sus **características** son:

1. Modelo de datos no relacional.
2. Funcionan bien en clusters.
3. Suelen ser en su mayoría de código abierto.
4. Su construcción está orientada para las aplicaciones web de nueva generación.
5. Son bases de datos con ausencia de esquema (**schemaless**)

Dentro de las diferentes características mencionadas:

- **Funciona bien en modo cluster:** Podemos definir básicamente este concepto como un conjunto de máquinas implementadas como servidores de procesamiento paralelo. Funcionan entre sí como un único recurso. Y como bien sabemos a cada elemento o máquina lo llamamos nodo.
- **Bases de datos con ausencia de esquema o schemaless:** Es una característica muy flexible, se puede almacenar información no uniforme y se facilita así la evolución.

2. FUNDAMENTOS DE LAS BASES DE DATOS NO-SQL:

A esta **cantidad ingente de datos** que produce el social-media y contenido manejado por el usuario, hay que sumarle las aportaciones de las nuevas fuentes de información como sensores, sistemas globales de posicionamiento o GPS, rastreadores etc...

Estos grandes paquetes de información, han introducido nuevos retos y oportunidades para el almacenamiento de información. Además, **los datos son cada vez más semiestructurados y escasos**. Esto quiere decir que las bases de datos relacionales son examinadas y requieren una definición de esquema inicial y referencias relacionales. Para resolver el problema relacionado con los grandes volúmenes de información y los datos semi-estructurados, surge nuevas clases de bases de datos dentro de la familia **No-SQL**. Este tipo de base de datos consisten en el **almacenamiento basado en columnas, clave/valor y de documentos**.

BASES DE DATOS RELACIONALES	BASES DE DATOS NOSQL
<p>La información está almacenada en un modelo relacional con filas y columnas.</p> <p>Una fila contiene información sobre un elemento mientras que las columnas contienen información específica.</p>	<p>La información está almacenada en un cliente o en una base de datos diferente con modelos de datos distintos.</p>
<p>Sigue un esquema fijo. Las columnas son definidas y establecidas antes de la entrada de datos. Además, cada fila contiene información de cada columna.</p>	<p>Sigue un esquema dinámico, puedes añadir columnas en cualquier momento</p>
<p>A favor del escalado vertical.</p>	<p>Facilita el escalado horizontal. Se puede escalar a través de servidores múltiples. Los servidores múltiples tienen la ventaja del precio en relación al Hardware que se va añadiendo comparado con el escalado vertical.</p>
<p>Atomicidad, consistencia, aislamiento y durabilidad. (ACID)</p>	<p>No está a favor de los principios ACID.</p>

3. BENEFICIOS DE LAS BASES DE DATOS NO-SQL I:

a) Capacidad de fuente de datos primaria y de análisis

El primer criterio de una clase empresarial NoSQL es que debe servir como fuente principal de datos que recibe información de distintas aplicaciones de negocio. También debe actuar como segunda fuente de información o análisis las aplicaciones de “business intelligence”. Este tipo de bases de datos deben de ser capaces de integrarse de una forma rápida a todos los tipos de datos estructurados, semiestructurados o sin estructura. Además, deben ser capaces de ejecutar consultas de alto rendimiento.

b) Capacidad Big Data

No se limitan a trabajar con Big Data. Una base de datos de este tipo de clase empresarial, puede escalar para administrar grandes volúmenes de datos desde Terabytes hasta Petabytes. Además de almacenar grandes volúmenes de datos, ofrece un alto rendimiento para la velocidad, variedad y complejidad de los datos

c) Disponibilidad continua

Para que una base de datos sea considerada de clase empresarial, debe ofrecer disponibilidad continua, sin un solo punto de fallo.

Además, en lugar de proporcionar la función de disponibilidad continua fuera del software, la solución NoSQL ofrece una disponibilidad continua integrada, que debe incluir las siguientes características clave:

1. Todos los nodos de un clúster deben poder atender solicitudes de lectura incluso si algunas máquinas no funcionan.
2. Debe ser capaz de replicar y segregar datos fácilmente entre diferentes partes físicas en un centro de datos. Esto evitará cortes de hardware.
3. Debe poder admitir diseños de distribución de datos que sean centros de datos múltiples ya sea en instalaciones físicas o en la nube

4. BENEFICIOS DE LAS BASES DE DATOS NO-SQL II:

a) Capacidad de tener múltiples centros de datos

Por lo general, las empresas poseen bases de datos altamente distribuidas que se encuentran en varios centros de datos y ubicaciones geográficas distintas.

La replicación de datos es una característica que ofrecen todas las bases de datos relacionales. Sin embargo, ninguna puede ofrecer un modelo simple de distribución sin causar problemas de rendimiento.

Una buena solución empresarial **NoSQL** debe admitir la implementación de varios centros de datos y debe proporcionar una opción configurable para mantener un equilibrio entre el rendimiento y la coherencia.

b) Fácil replicación independientemente de la ubicación

Para evitar que la pérdida de datos afecte a una aplicación, una buena solución NoSQL proporciona una gran capacidad de replicación. Estos incluyen una capacidad de lectura y escritura en cualquier lugar con compatibilidad total y dependencia de ubicación.

Esto significa que **se pueden escribir datos en cualquier nodo de un clúster**, hacer que **se repliquen** en otros nodos, y ponerlos a disposición de todos los usuarios. Además, la capacidad de escritura en cualquier nodo debe garantizar que los datos estén seguros en caso de un corte de suministro eléctrico o cualquier otro tipo de incidente.

5. BENEFICIOS DE LAS BASES DE DATOS NO-SQL III:

a) Sin capa de almacenamiento en caché separada

Una buena solución NoSQL es capaz de utilizar múltiples nodos y distribuir datos entre todos los nodos adjuntos. **No requiere una capa de almacenamiento en caché específica para almacenar datos**. Las memorias caché de todos los nodos pueden almacenar datos para una entrada y salida rápida o para un acceso entrada/salida. La base de datos NoSQL elimina el problema de sincronizar los datos de la caché con la base de datos persistente.

Por lo tanto, admite una escalabilidad simple con menos problemas de administración.

b) Lista para la nube

Una solución NoSQL de nivel empresarial **debe estar preparada para la nube**. Un clúster de base de datos NoSQL **debe** poder funcionar en una configuración de la nube, como **Amazon EC2**, y también debe poder extender y reducir un clúster cuando sea necesario.

También debe admitir una solución híbrida en la que parte de la base de datos se aloje dentro de las instalaciones de la empresa y otra parte se aloje en la nube.

c) Alto rendimiento con escalabilidad lineal

Pueden mejorar el rendimiento agregando nodos a un clúster. Normalmente, el rendimiento de otros sistemas de bases de datos puede disminuir cuando se agregan los nodos adicionales, sin embargo, una buena solución NoSQL aumenta el rendimiento de las operaciones de lectura y escritura cuando se agregan nodos adicionales. Estas ganancias de rendimiento son de **naturaleza lineal**.

[Beneficios](#) más de las bases de datos No-SQL:

- Soporte de esquema flexible.
- Admite lenguajes y plataformas clave para desarrolladores.
- Fácil de implementar, mantener y extender.
- Comunidad de código abierto.



14
NOVIEMBRE
SOLO EN CINES

TODOS LOS LADOS DE LA CAMA

6. TIPOS DE BASES DE DATOS NO-SQL:

- 1. Bases de datos clave/valor:** La base de datos clave valor, es realmente una tabla hash de claves y valores. Algunos de los ejemplos son: Riak, Tokyo Cabinet, **servidor Redis**, Memcached y Scalaris
 - 2. Bases de datos basadas en documentos:** Almacena documentos compuestos por elementos etiquetados. Algunos ejemplos son: **MongoDB**, CouchDB, OrientDB y RavenDB
 - 3. Bases de datos basadas en columnas:** Cada bloque de almacenamiento contiene datos de una sola columna. Ejemplos: BibTable, **Cassandra**, Hbase e Hypertable.
 - 4. Bases de datos basadas en gráficos:** Es una base de datos de red que utiliza nodos para representar y almacenar datos. Algunos ejemplos son: **Neo4J**, **InfoGrid**, Infinite Graph y FlockDB. La disponibilidad de opciones en las bases de datos NoSQL tiene sus propias **ventajas** (permite elegir un diseño de acuerdo a los requisitos de su sistema) e **inconvenientes** (aun ajustando el producto a los recursos del sistema no siempre funcionará correctamente).

7. INTRODUCCIÓN AL BIG DATA:

(DATA) → Son las **cantidades, los caracteres o símbolos con los que se realizan operaciones mediante una máquina**. Pueden almacenarse y transmitirse en forma de señales eléctricas y registrarse en soportes de almacenamiento magnéticos, ópticos o mecánicos.

Big Data son realmente datos, pero con la diferencia que su tamaño es muy grande. Con ello definimos en realidad una colección de datos que es gigante y crece cada vez más con el tiempo.

¿Por qué las bases de datos Big Data son sistemas diferentes?

Hay algunas características como la escalabilidad a gran volumen, la rapidez de mover información y de procesarla, y las propiedades de algunos de los datos que se tratarán en el proceso, presentan nuevos desafíos importantes a la hora de diseñar soluciones. El objetivo de la mayoría de los sistemas de Big Data es sacar a la luz conocimientos y conexiones de **grandes volúmenes de datos heterogéneos** que no serían posibles con métodos convencionales.

8. TIPOS DE BIG DATA:

a) Estructurado

Definimos datos estructurados como toda aquella información que pueda ser almacenada, se pueda acceder y se accede, de forma fija.. En la actualidad existe un problema real, los datos crecen en tal medida que alcanzan tamaños como el [zettabyte](#).

b) Datos no estructurados

Cualquier dato cuya forma o estructura sea desconocida se clasificará como datos no estructurados. Además de que el tamaño es enorme, los datos no estructurados plantean múltiples desafíos si nos referimos al procesamiento de los mismos. Hoy en día muchas compañías disponen de sistemas no estructurados donde mezclan colecciones que contienen imágenes de texto, números, direcciones, ficheros media, etc. Esto puede suponer un problema real si estos datos no son procesados y si no se sabe sacar partido de ellos.

WUOLAH

c) Datos semi-estructurados

Aparentemente los datos semi-estructurados tienen la similitud de ser en parte, datos estructurados en cuyo interior poseen datos no estructurados. La realidad es que no están planteados para ser plenamente estructurados como una tabla. Un ejemplo sería los **datos representados como en el formato XML**.