Instituto FOC

Módulo de desarrollo de aplicaciones web

BRUNO MARENCO CERQUEIRA

Tarea Individual 1: Conceptos Básicos sobre Bases de Datos

Octubre/2017

Índice

Pregunta 1

Cuestión

Comente en cuál de los siguientes casos sería más conveniente el uso de una base de datos distribuida y por el contrario, en cuál sería más adecuado el uso de una base de datos centralizada justificando su respuesta.

- Base de datos de gestión de un sistema sanitario público.
- Base de datos de un pequeño videoclub.

Respuesta

Para saber si es necesaria la fragmentación de la información en algún proyecto, debemos entender la estructura lógica que va a tener el proyecto a corto y medio plazo, para verificar si las ventajas que ofrece son necesarias en el caso evaluado. En este planteamiento, estamos antes dos casos de uso de base de datos, por un lado, un sistema sanitario público, por otro lado, un pequeño videoclub.

En el caso del pequeño videoclub, la base de datos que vamos a tener será local, con un acceso limitado y con un presupuesto mucho menor para elaborar cualquier sistema informático. Por eso, **lo más conveniente sería una base de datos centralizada**, que posee menores costes de creación, implementación y manutención. La información estaría centralizada en una unidad física, lo que conlleva algunos inconvenientes, pero debido a que se necesita menor capacidad de acceso y, siendo una empresa pequeña, se deben mantener los costes bajos, sería la mejor alternativa para implementar.

En el sistema sanitario público, donde la base de datos es mucho más basta, la necesidad de un sistema más flexible para crecer, con mayor número de accesos y que probablemente, tendrá una inversión inicial mucho superior al primer sistema, podremos plantear una fragmentación de la base de datos para **crear un sistema de**

base de datos distribuido. Eso implica mayores costes en el desarrollo de la plataforma, pero trae consigo todas las ventajas de un sistema así. Mejora el rendimiento en el uso de la aplicación al tener los datos dispersos en diversas máquinas físicas, el acceso a los datos se realizará a una máquina en cuestión, con el sistema conectado en diversas máquinas, se permitirá una mayor eficiencia y velocidad en los accesos. La fiabilidad es mucho mayor, dado que ya no se depende de una única máquina, ante la caída de un servidor, el resto puede mantener el funcionamiento del sistema. Tiene una mejor flexibilidad para el crecimiento del sistema, pudiendo alocar los datos por región según necesidad de información específica por las diversas partes que integran el sistema.

Conclusión

A partir de la respuesta obtenida, se puede deducir que sistemas mayores, aunque exigen más inversión, pueden beneficiarse de todas las ventajas que conllevan los sistemas distribuidos, muy especialmente la fiabilidad y eficiencia de los accesos del sistema y la flexibilidad para el futuro crecimiento de la base de datos según necesidad. Los sistemas menores pueden centralizar las bases de datos por el menor capital disponible de inversión y no ver problemas operacionales en el sistema, debido a una menor demanda de accesos y a una menor complejidad del sistema.

Pregunta 2

Cuestión

Supongamos que accedemos a la página web de ebay, y que la base de datos de dicha página web está construida sobre la arquitectura ANSI/SPARC en tres niveles. ¿Cuándo se visualizan los distintos productos en la pantalla, en qué nivel de la arquitectura ANSI/SPARC se encuentra?

Respuesta

La arquitectura ANSI/SPARC es un diseño de arquitectura abstracta para los sistemas de gestión de bases de datos, dispone de tres niveles: interno, conceptual y externo.

Cada nivel sirve para separar las actividades del gestor de base de datos y los usuarios sin afectar al resto. Debido a esto, el gestor de base de datos puede alterar los modelos conceptuales de la base de datos sin afectar al nivel de los usuarios.

Los dos primeros niveles son el interno y el conceptual, que son los más alejados de los usuarios. En el nivel interno se describen los archivos que tienen la información, es el más cercano al almacenamiento físico y el nivel conceptual es la representación de los datos que se desean guardar en forma de esquema, con las relaciones existentes entre los diversos datos.

El nivel que permite la visualización de los productos por parte del usuario es el nivel externo y es la visión final de los datos por parte de los usuarios, lo que permite a los usuarios ver y manipular la información de forma superficial, sin afectar a los modelos de esquema internos, relaciones entre bases de datos, o sistema de archivos de la información. Con este método, los usuarios pueden usar y modificar las informaciones de las bases de datos sin llegar a modificar la estructura interna del sistema.

Conclusión

El sistema ANSI/SPARC es ampliamente utilizado en los sistemas de gestión de bases de datos ya que permite una rápida abstracción de cada nivel de acceso y visualización de la base de datos en su totalidad. Los niveles externos son utilizados por los usuarios, permitiendo el uso de la aplicación por su parte, pero sin afectar la estructura interna.

Pregunta 3

Cuestión

Realizar una pequeña investigación en Internet, en la que se indague sobre los principales sistemas gestores de bases de datos que se encuentran hoy día en el mercado, clasificándolos según el lenguaje de consulta que utilizan SQL o NoSQL.

Respuesta

Los sistemas de gestión de datos más utilizados históricamente son los que utilizan SQL como lenguaje de consulta, se basan en modelos relacionales, a diferencia de estos los lenguajes NoSQL no utilizan SQL como lenguaje de consulta y no tienen bases de datos relacionales.

Definidas las principales diferencias entre SQL y NoSQL, podemos definir las ventajas y desventajas de cada uno. SQL posee una sólida y gran comunidad y su estructura está muy padronizada, lo que permite encontrar soluciones a problemas enfrentados fácilmente. Por su parte, los lenguajes NoSQL son mucho más novedosos, por lo que su comunidad es muy reciente y dispersa y todavía no se están estandarizando los puntos clave del lenguaje, lo que puede llevar a estancar y dificultar su evolución y también dificulta el aprendizado por parte de novatos y las migraciones entre bases de datos.

Por su parte, las bases de datos NoSQL se están popularizando debido a grandes ventajas respecto a SQL, como son una mayor eficiencia en la consulta de informaciones, por lo que grandes empresas como Google, Amazon o Facebook utilizan esta tecnología. También son más flexibles que los SQL tradicionales debido a que no se basan en modelos relacionales, por lo que su arquitectura puede ser desarrollada dependiendo de las necesidades poco a poco, a diferencia de si se trabaja con SQL, que el modelo deberá estar definido previamente o será mucho más difícil hacer cambios posteriormente.

Existen diversos lenguajes siendo desarrollados siguiendo SQL o NoSQL. Dentro de los lenguajes SQL más utilizados podemos encontrar:

MYSQL es uno de los sistemas de gerenciamiento de banco de datos más populares que existe, puede ser optimizado para Web y está disponible para prácticamente cualquier sistema operacional, además de ser un software libre. Es optimizado para la agilidad de las operaciones en banco de datos.

PostgreSQL es un sistema de gerenciamiento de bases de datos versátil, gratuito y de código abierto que funciona con instrucciones SQL. Es optimizado para aplicaciones complejas y que muevan grandes volúmenes de datos o que traten informaciones críticas.

Entre las opciones de pago más extendidas en el mercado podemos encontrar a **Microsoft SQL Server,** y a **Oracle,** un software de gran potencia y elevado precio, que sólo las grandes compañías son capaces de adquirir.

Dentro de los lenguajes NoSQL podemos encontrar diversos tipos como orientada a documento, a columnas, de clave valor y grafo.

MongoDB es probablemente el lenguaje NoSQL más famoso y orientada a documentos. En vez de tablas, guarda los datos en documentos que son almacenado en BSON (representación binaria de JSON). No tiene las normas de las tablas relacionales, los documentos de la misma concepción pueden tener esquemas diferentes. Està escrito en C++ y las consultas se hacen a través de objetos JSON.