**TRABAJO PRÁCTICO - REPLICACION**

La Dirección de Recursos Naturales y Gestión, junto con la Dirección de Flora de la provincia de Formosa, desean implementar un sistema para el control de material “madera” que circula desde los Bosques Formoseños hasta los Establecimientos para la producción de productos hechos en madera. Debido a la problemática de la Deforestación y tala indiscriminada de árboles, la entidad desea mantener un control y seguimiento total de la materia prima. Para ello, cada Dependencia de la Provincia (una por Cada Departamento) desea implementar un Sistema que registra las altas, bajas y modificaciones de entidades como: Productor, transporte, materia prima, establecimiento de bosque, establecimiento de producción y materia prima. Además, cuenta con un módulo destinado a la Administración de Permisos de Tala, Permisos de Transporte y Autorizaciones de Producción.

Debido a que el Sistema cuenta con información global, las bases de datos de los Sistemas en cada Dependencia, deberán estar sincronizadas, y actualizadas, ya que puede surgir el caso en que un Productor saque un permiso de tala y transporte en Gral. Mosconi para llevar materia prima a la localidad de Pirané, y no puede pasar que este dado de Alta en Gral. Mosconi pero todavía no impacte en el Sistema de Pirané al momento de llegar. De la misma manera, en cada puesto caminero de Gendarmería, se deberá poder realizar consultas al Sistema para ver el estado de un Permiso en circulación.

Se han encontrado con el problema de que, el servicio de Internet en toda la provincia es deficiente, por lo que las Bases de Datos no se encuentran conectadas todo el tiempo, solo a intervalos cortos por día. Por lo tanto se necesita de alguna solución que le permita poder sincronizar las bases de datos de cada Dependencia, sin tener que invertir demasiado en alguna infraestructura de red (como ser un troncal dedicado de Fibra Óptica).

Se lo contrata a usted, en carácter de “Consultor Experto” para que asesore acerca de cuál es el mejor camino a tomar en este caso.

**Actividad**

Teniendo en cuenta el Escenario anterior

1. ¿Qué tipo de solución propone para la problemática planteada? Justifique
2. ¿Qué tipo de Replicación (transaccional, de instantáneas o Mezcla) cree usted es la más apropiada para el escenario propuesto? Justifique

Responda.

1. Explique Brevemente en que consisten los Sistemas de Bases de Datos Distribuidas y los Sistemas de Bases de Datos Paralelos.
2. ¿Que tipo de arquitecturas existen para los Sistemas de Bases de Datos paralelos? Indique ventajas y desventajas de cada uno.
3. ¿Qué diferencia existe entre la Duplicación y la Replicación?
4. ¿En que otros Escenarios utilizaría la Replicación?

**Desarrollo**

1. Para la problemática planteada se ofrece la siguiente solución:

Implementar el uso de un Sistema de Base de Datos distribuidas ya que la misma posee las siguientes características:

* Una Colección de datos compartidos lógicamente relacionados.
* Datos divididos en una serie de fragmentos.
* Los fragmentos pueden ser replicados.
* Los fragmentos se asignan a distintas instalaciones.
* Las instalaciones están enlazadas por redes de comunicaciones.
* Los datos de cada instalación están bajo el control de SGBD.
* El SGBD puede gestionar las aplicaciones locales de manera autónoma.
* Cada SGBD participa al menos en una aplicación global
* La distribución de los datos es transparente al usuario

Y a todo esto sumarle la técnica de Replicación ya que muchas veces esos sistemas deben de tener garantizada la accesibilidad por lo que para evitar problemas es necesario utilizar este tipo de técnicas de replicación de base de datos de forma que un fallo en uno de los servidores de base de datos no impida a los usuarios seguir utilizando la aplicación. Mediante la replicación de base de datos, usuarios de todo el mundo pueden estar accediendo a lo que para ellos son los mismos datos, aunque en realidad, físicamente esos datos pueden estar de forma transparente para el usuario, en diferentes nodos o localidades.

1. Para mi el tipo de replicación más apropiada para el escenario propuesto es la replicación Transaccional ya que primero se envía una copia completa de la base de datos y luego se van enviando de forma periódica (o a veces continua) las actualizaciones de los datos que cambian. Se utiliza cuando necesitas que todos los nodos con todas las instancias de la base de datos tengan los mismos datos a los pocos segundos de realizarse un cambio.

Respuestas:

1. **Sistema de bases de datos distribuidas:** Está compuesta por una única base de datos lógica dividida en una serie de fragmentos. Cada fragmento se almacena en una o más computadoras bajo el control de un SGBD independiente. Las computadoras están conectadas mediante una red de comunicaciones. Poseen capacidad de procesamiento independiente de solicitudes locales y remotos.

* **Sistema de bases de datos paralelo:** Es un SGBD que se ejecuta sobre múltiples procesadores y utilizando múltiples discos y que está diseñado para ejecutar las operaciones en paralelo, siempre que sea posible, con el fin de mejorar prestaciones.

1. Para los Sistemas de Bases de Datos paralelos existen las siguientes arquitecturas:

* **De memoria compartida:** Consiste en una Serie de procesadores que comparten memoria (SMP, multiprocesamiento simétrico)

*Ventaja*

Proporciona alta velocidad

*Desventaja*

Tiene limitaciones de escalabilidad

* **De discos compartidos:** Arquitectura débilmente acoplada donde cada procesador tienen su propia memoria, pero se comparten los discos, por lo tanto, no hay necesidad del particionamiento físico de datos.

Adecuado para aplicaciones inherentemente centralizadas y que requieren alta disponibilidad

* **Procesamiento Paralelo Masivo:** Sin compartición (MMP, Procesamiento Paralelo Masivo). Posee Alta comunicación entre los procesadores y la base de datos se particiona entre todos los discos de cada uno de los sistemas asociados y los datos están disponibles en forma transparente para los usuarios.

Ofrece escalabilidad, pero tiene prestaciones optima si los datos son locales.

Similar SGBDD, pero la distribución de los datos solo es a los efectos de rendimiento, normalmente dentro de la misma computadora.

Es para grandes volúmenes de datos y para respuesta a consultas muy complejas.

1. La diferencia entre replicación y duplicación es que la Replicación puede ser parcial o total y la Duplicación siempre es total, una copia exacta de toda la base de datos.
2. La replicación se podría utilizar en los siguientes escenarios:

* La mayoría de los datos no cambian con frecuencia
* Se replican pocas cantidades de datos
* Se desea que las modificaciones se propaguen a los suscriptores solamente
* La aplicación no permite un periodo de latencia largo
* Los sitios hacen cambios, pero necesitan información de otros sitios
* Exige que se implemente un gestor de conflictos de actualización