



ELEMENTOS DE BASES DE DATOS

Segundo Cuatrimestre de 2015

Trabajo Práctico N° 10

Sistemas Paralelos y Distribuidos

Arquitecturas Paralelas

1. Explique las calidades "aceleración" (speedup) y "escalabilidad" (scaleup).
2. Cuál de las calidades anteriores se busca aplicar en los siguientes casos:
 - (a) Una empresa de telefonía espera aumentar significativamente el número de clientes en los próximos seis meses a partir de promociones que beneficien a sus usuarios. La empresa está interesada en que este proyecto no impacte sobre la calidad del servicio.
 - (b) Una empresa de servicios de televisión, series y películas vía internet piensa mejorar la calidad de imagen de sus productos sin afectar el servicio.
 - (c) El Instituto Nacional de Estadísticas y Censos desea extender las funcionalidades de la aplicación de Consultas y Reportes ofreciendo un generador de consultas con filtros, indicadores y alternativas de visualización para atender los requerimientos de información del Ministerio de Economía.
3. Explique y compare cada una de las alternativas vistas para arquitecturas paralelas.
4. Comparar las arquitecturas paralela y distribuida.

Arquitecturas Distribuidas

5. Explique las ventajas y desventajas de contar con un sistema de bases de datos distribuido.
6. Enumere y explique los principios de transparencia que debiera respetar el nombrado de los datos en un sistema de bases de datos distribuido.
7. Dadas las siguientes situaciones, determine en cada caso si conviene tener repetición de datos, fragmentación horizontal, fragmentación vertical o alguna combinación de ellas:
 - (a) La base de datos de un buscador como *Yahoo!*, *Google* o *Altavista*. El requerimiento principal es el servicio ininterrumpido, de manera que las fallas en el sistema sean transparentes a los usuarios conectados.
 - (b) La base de datos de un supermercado con varias sucursales. El requerimiento principal es agilizar las consultas sobre la tabla

```
stock(articulo_codigo, articulo_cantidad, sucursal_codigo)
```

teniendo en cuenta que la mayoría de las consultas son locales a cada sucursal.
8. Analice las siguientes afirmaciones acerca del manejo de información utilizando bases de datos distribuidas:

- (a) Existe una mejor utilización de los datos: los datos pueden compartirse y el control de los mismos distribuirse entre los distintas sitios.
 - (b) Los sistemas distribuidos son más confiables.
 - (c) El procesamiento de las consultas se ve agilizado.
 - (d) Aumentan las dificultades de comunicación.
9. Considere el siguiente escenario en un sistema distribuido: Hay cuatro sitios S_1, S_2, S_3 y S_4 conectados en red. Del ítem de dato B se tiene una copia en S_1, S_2 y S_3 . Suponga que una transacción T_1 está ejecutándose en el sitio S_1 . Para los protocolos ROWA, Mayoría y 2-de-3 indique:
- (a) Sobre cuantas copias debe obtener un bloqueo la transacción T_1 , para obtener:
 - i. un bloqueo exclusivo sobre el ítem de dato B.
 - ii. un bloqueo compartido sobre el ítem de dato B.
 - (b) Cual es la cantidad mínima (mejor caso) de mensajes de control que se necesitan, (considerando que no es necesario enviar mensajes de un sitio a si mismo) para que la transacción T_1 obtenga:
 - i. un bloqueo exclusivo sobre el ítem de dato B.
 - ii. un bloqueo compartido sobre el ítem de dato B.
10. Considere el escenario planteado en el ejercicio ??.
- (a) Siguiendo el protocolo de *sitio primario* y siendo S_2 el sitio primario para el ítem de dato B, indique cuantos mensajes de control se necesitan para que la transacción T_1 obtenga:
 - i. un bloqueo exclusivo sobre el ítem de dato B.
 - ii. un bloqueo compartido sobre el ítem de dato B.
 - (b) Suponga que los sitios S_3 y S_4 poseen un *read-token* para el dato B. Siguiendo el protocolo de *tokens de copia primaria*, indique cual es la cantidad mínima de mensajes de control que se necesitan para que la transacción T_1 obtenga:
 - i. un bloqueo exclusivo sobre el ítem de dato B.
 - ii. un bloqueo compartido sobre el ítem de dato B.
 - (c) Siguiendo el protocolo de *nodo central* y siendo S_4 el nodo encargado de conceder los bloqueos, indique cual es la cantidad de mensajes de control que se necesitan para que la transacción T_1 obtenga:
 - i. un bloqueo exclusivo sobre el ítem de dato B.
 - ii. un bloqueo compartido sobre el ítem de dato B.
11. Considere el siguiente escenario en un sistema distribuido: Hay cinco sitios S_1, S_2, S_3, S_4 y S_5 conectados en red. Del ítem de dato A existe una copia cada sitio. Tres transacciones T_1, T_2 y T_3 están ejecutándose en S_1, S_2 y S_3 respectivamente. Complete la siguiente tabla indicando si es posible que las situaciones planteadas ocurran en cada uno de los protocolos.

Situación	Rowa	Mayoría	4 de 5
T_2 y T_3 obtuvieron un R-lock sobre el ítem de dato A y T_1 obtuvo un W-lock sobre tres copias del dato A.			
T_2 obtuvo un W-lock sobre el ítem de dato A y T_1 obtuvo un R-lock sobre una copia del dato A.			
Deadlock entre las transacciones T_1 y T_2			

12. Analice la siguiente afirmación: “Que las planificaciones locales para las transacciones que se ejecutan en una localidad sean localmente serializables no garantiza que las planificaciones globales sean serializables”.
13. Supongamos la siguiente situación para una transacción distribuida. Existen cuatro participantes, T_1 (el coordinador), T_2 , T_3 y T_4 sincronizados bajo el protocolo de compromiso de dos fases. Explicar las acciones llevadas adelante por el coordinador, los participantes y las acciones que debe llevar adelante en el sitio que se produce la falla cuando se reestablece:
 - (a) El coordinador T_1 falla después de enviar el mensaje “commit T” a T_2 y T_3 , pero no a T_4 .
 - (b) El participante T_2 falla después de enviar un mensaje “no-T”; todos los otros participantes enviaron “ready-T”.
 - (c) El participante T_2 falla antes de decidir; los otros participantes envían “ready-T”.
 - (d) Todos los participantes enviaron un mensaje “ready-T” y el coordinador T_1 falla antes de comunicarse con los participantes.
14. Explique cuál es el propósito de los siguientes protocolos en un ambiente de bases de datos distribuido:
 - (a) Protocolos para locking distribuido (ROWA, mayoría, etc).
 - (b) Protocolo de dos fases estricto.
 - (c) Protocolo de compromiso de dos fases.
15. ¿Qué problemas se presentan para la gestión de deadlock usando un *servidor de deadlock centralizado*? ¿La *gestión de deadlock distribuida* soluciona estos inconvenientes?