

UEA 1115114: Base de Datos

Notas al Diseño de BD,
integridad, seguridad
Procedimientos Catalogados y
Disparadores
Bases de Datos Distribuidas
Cliente-servidor

Carlos Barrón Romero

Departamento de Ciencias Básicas,
División Ciencias Básicas e
Ingeniería

Recapitulación y Recordatorios

1. Tomen en cuenta los aspectos de normalización hasta 3era forma normal.
2. Reflexione en su diseño con consultas en álgebra y SQL (como parte de la prueba de escritorio del diseño), como en los ejercicios de los capítulos 6 y 7 del Date.

Objetivos y actividades de la clase

- Por favor preguntas de su proyecto
- Hoy ampliaremos los conceptos de diseño Integridad, restricciones y la regla de oro.
- Entender y conocer de los manejadores y de las bases de datos distribuidas para analizar su funcionamiento y perspectivas de desarrollo e utilización
- Ligar con Bases de Datos Distribuidas los conceptos: Diseño de BD, Procedimientos Catalogados y Disparadores, integridad, seguridad, modelo Cliente-servidor, problemática para el administrador

Objetivos y actividades de la clase

- Los alumnos realizarán cálculos de estrategias de manejo de Bases de Datos Distribuidas para que objetivamente analicen y definan el funcionamiento óptimo de una BD Distribuida

Integridad y regla de oro

- A la implementación de un diseño de BD se le agregan: Restricciones que se deben cumplir. Por ejemplo cuando un dato puede tomar un valor por omisión.
- Muchos MBD permiten definir tales restricciones, como por ejemplo la clave de un proveedor no puede ser nula:

CONSTRAINT BVPY V (V# NOT Null)

Donde BVPY es la base que hemos visto en clase con las relaciones P, Y,
V, VPY

Note que la llaves de VPY no pueden ser invalidas!

Integridad y regla de oro

- Regla de ORO: no introducir instancias a una BD que no cumpla con el diseño y restricciones
- ESTO ES EJEMPLO DE VIOLACION DE LA REGLA DE ORO, ¿Porqué?

V, V# → (PROVEEDOR, STATUS, CIUDAD)

V#	PROVEEDOR	STATUS	CIUDAD
	SUPER COMER S.A.	30	ATENAS

¿Cuál es la violación de esta instancia?

Integridad y regla de oro

- Regla de ORO: no introducir instancias a una BD que no cumpla con el diseño y restricciones
- ESTO ES EJEMPLO DE VIOLACION DE LA REGLA DE ORO

V (del diseño partes, proveedores y proyectos, contraportada Date)

V#	PROVEEDOR	STATUS	CIUDAD
	SUPER COMER S.A.	30	ATENAS

Viola nuestra convención de 2da forma normal

V# → (PROVEEDOR, STATUS, CIUDAD)

Anomalías de borrado e inserción

- No respetar la integridad puede ocasionar pérdida de información
- Por la integridad y restricción del diseño del diagrama Entidad-vínculo, se deben insertar primero las instancias de V, P, Y antes que las de VPY.
- Ejemplos:
- No se puede insertar en vpy (3,10,21, 100)
- Solo se puede borrar (2,11,22,50) o (2,11,22,30), pero no ambos

V		
V#	PROVEEDOR	CIUDAD
1	Gómez hermanos	DF
2	Comex	Puebla

P		
P#	PARTE	PESO
10	Puerta	10.0
11	Ventana	5.0
12	Pintura	8.0

Y		
Y#	PROYECTO	CIUDAD
21	VIADUCTO	DF
22	BICENTENARIO TEATRO UAM	Toluca

VPY			
V#	P#	Y#	Cantidad
1	10	21	100
2	11	22	50
1	11	21	25
1	12	21	75
2	11	22	30

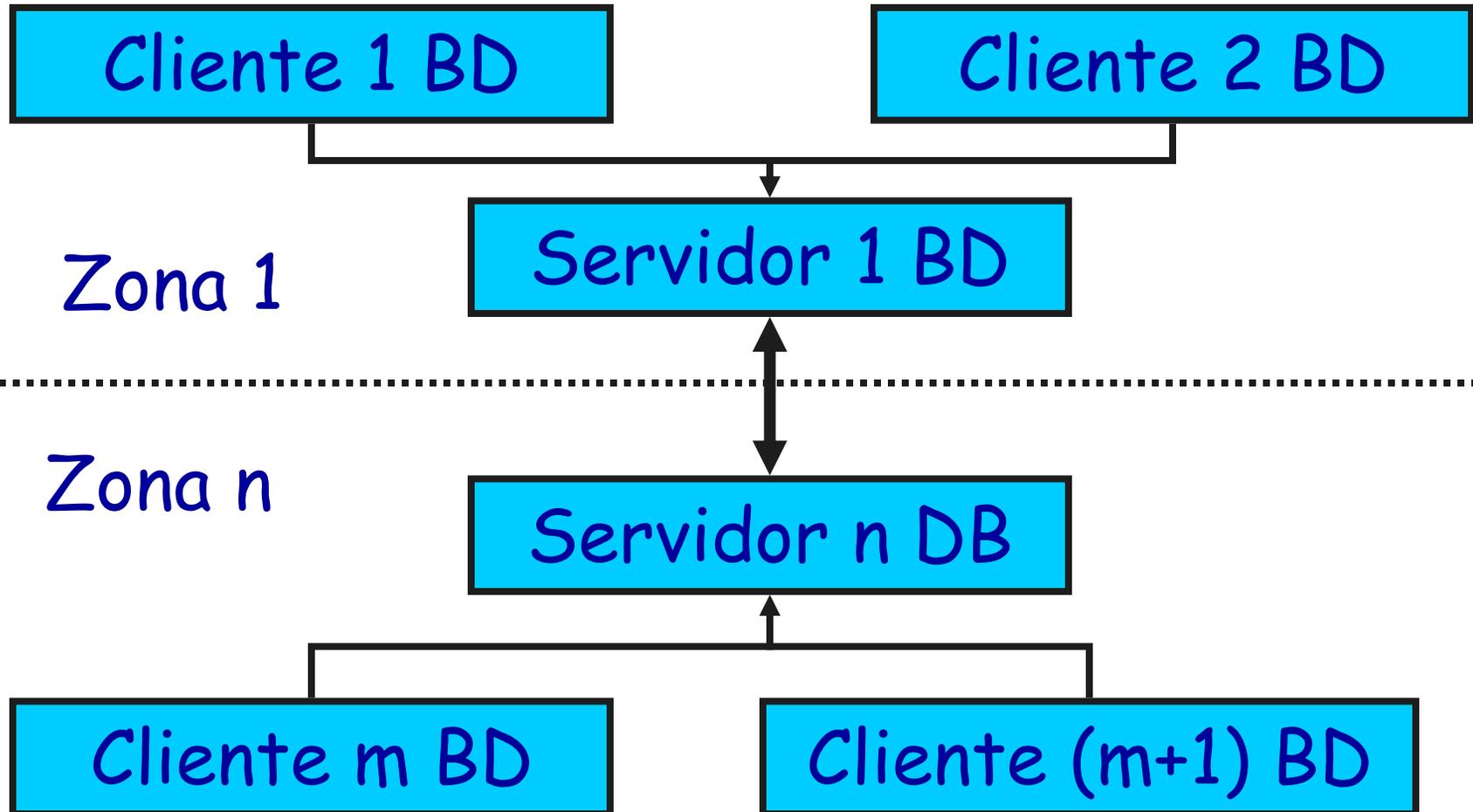
Seguridad

- El Administrador de BD es responsable de
 - Relacionar datos, usuarios y aplicaciones y puntos de acceso
 - Control por Prioridades y tablas de acceso como en los SO
 - Borrado y respaldo de datos.
 - Borrado permanente y manejo de información histórica

Procedimientos Catalogados y Disparadores

- Los modernos manejadores de base de datos permite definir datos que usan procedimientos catalogados y que se disparan en forma "automatica".
- Por ejemplo un punto de reorden:
 - Artículo: Jugo de frutas del monte
 - Max_existencia (100).
 - Existencia (2)
 - Punto_de_reorden (4): Si Existencia \leq Punto_de_reorden, Entonces envía un EMAIL: solicito (Max_existencia-Existencia) unidades del artículo (\bar{X}).
 - En este caso el EMAIL: **solicito 98 unidades del artículo Jugo de frutas del Monte.**

Cliente-Servidor



Bases de Datos Distribuidas

- DOCE OBJETIVOS
 1. Autonomía Local
 2. No dependencia de un sitio central
 3. Operación Continua
 4. Independencia de ubicación
 5. Independencia de Fragmentación (¿Qué es fragmentación?)

Bases de Datos Distribuidas

DOCE OBJETIVOS (cont.)

- Fragmentación: Almacenar una relación en varios sitios o archivos
 - Replicación: Copia de una relación en otro sitio o archivo
6. Independencia de Replicación
 7. Procesamiento de Consultas Distribuidas
 8. Administración de transacciones distribuidas

Bases de Datos Distribuidas

DOCE OBJETIVOS (cont.)

9. Independencia de hardware
10. Independencia del Sistema Operativo
11. Independencia de la red de comunicaciones
12. Independencia del DBMS: Misma VISTA, misma interfase.

Procesamiento de BD Distribuidas

- El Ing. en Computación de saber calcular los tiempos de servicio de una Base de datos Distribuidas para efecto de sugerir estrategias y formas de acomodar los recursos de la base de Datos Distribuidas (**Date, pág. 664-666**).
- La formula básica es:
$$TC = R + VD / VC$$

Donde R es Retardo
VD Volumen Datos
VC Velocidad de Transmisión

Procesamiento de Consultas, ejemplo

Objetivo: Comparación de 6 estrategias

Suposiciones de Datos y consulta:

1. Base de Datos (Proveedores y partes):

V {V#, Ciudad } 10,000 en sitio A

P {P#, Color } 100,000 en sitio B

VP {V#, P# } 1,000,000 en sitio A

Cada tupla de 25 bytes (200 bits)

Procesamiento de Consultas, ejemplo

Objetivo: Comparación de 6 estrategias

- Suposiciones de Datos y consulta:

2. Consulta

```
(( V JOIN VP JOIN P) WHERE  
  CIUDAD = 'Londres' AND  
  COLOR = 'Rojo' )) {V#}
```

Cantidad de partes rojas = 10

Cant. de envíos Proveedores de Londres = 100,000

Procesamiento de Consultas, ejemplo (cont.)

Objetivo: Comparación de 6 estrategias

- Suposiciones de comunicación:
 1. Velocidad de Transmisión de Datos (VT): 50,000 bits/seg
 2. Retardo acceso (R): 0.1 seg

Procesamiento de Consultas, ejemplo (cont.)

Objetivo: Comparación de 6 estrategias

Suposiciones de comunicación:

1. Formula Tiempo de transmisión de datos:

$$TD = R \text{ 'total' } + (\text{volumen datos}) / VT$$

2. Formula Tiempo de envío de mensajes:

$$TM = \text{Cantidad de mensajes} / 10 + (\text{cantidad bits}) / VT$$

Procesamiento de Consultas

Resumen de datos del problema

V {V#, Ciudad } 10,000 en sitio A

P {P#, Color } 100,000 en sitio B

VP {V#, P# } 1,000,000 en sitio A

Cada tupla de 25 bytes (200 bits)

partes rojas = 10

envíos proveedores de Londres = 100,000

$TD = 0.1 * no_solicitudes + (volumen\ datos) / VT$

$TM = Cantidad\ de\ mensajes / 10 + (volumen\ datos\ mensajes) / VT$

$VT = 50,000\ bits/seg$

$R = 0.1\ seg$

Los mensajes son de a lo mas 10 bits (pequeños).

Procesamiento de Consultas, ejemplo (cont.)

Estrategias

1. Mover la partes hacia el sitio A y procesar la consulta en A:

Datos

P {P#, Color } 100,000 en sitio B

Cálculo

$$T[1] = TD = 0.1 + (100,000 * 200) / 50,000$$

~ 400 seg (6.67 min.)

Procesamiento de Consultas, ejemplo (cont.)

Estrategias

2. Mover Proveedores y envíos hacia el sitio B y procesar la consulta en B:

Datos

V {V#, Ciudad } 10,000 en sitio A

VP {V#, P# } 1,000,000 en sitio A

2 mensajes Retardo = $0.1 * 2 = 0.2$

Calculo

$$T[2] = 0.2 + (10,000 + 1,000,000) * 200 / 50,000$$

~ 4040 seg (1.12 horas!)

Procesamiento de Consultas, ejemplo (cont.)

Estrategias

3. Juntar Proveedores y envíos en A, restringir el resultado a los proveedores de Londres y luego para cada uno de estos revisar en B para ver la parte que les corresponde (2 mensajes).

Datos

V {V#, Ciudad} 10,000 en sitio A

envíos proveedores de Londres = 100,000

2 para ver lo que les corresponde

Calculo (2. Formula Tiempo de envío de mensajes)

$$\begin{aligned} T[3] &= 2(100,000)/10 + (10*100,000) / 50,000 \\ &= 20,020 \text{ seg} \\ &\sim 20,000 \text{ seg (5.56 horas!)} \text{ (Resultado del libro)} \end{aligned}$$

Otro problema

Dados los datos de una BD distribuida en tres sitios:

Velocidad de Transmisión = 50,000 bits/seg.

Retardo = 0.1 seg.

V {V#, Ciudad} 10,000 en sitio A

P {P#, Color} 100,000 en sitio B

VP {V#, P# } 1,000,000 en sitio C

Cada instancia de 20 bytes (160 bits)

a) Hay una replica en C de P de todas sus instancias, las de color azul son 10.

b) No hay replicas. Pero hay 1,000 instancias de color azul en P.

VP asociadas a las de color azul = 100 (se supone que solo hay 50 con V# distintos en VP)

Se requiere la consulta

```
(( V JOIN VP JOIN P) WHERE COLOR = 'azul' )) {V#, P.Color}
```

Para los incisos a) y b) ¿De que forma y donde se realiza la consulta para el menor tiempo de transmisión?

Cierre de clase

- Las Base de Datos Distribuidas están en el límite de aplicaciones actuales modernas (gracias a los avances en redes) y su problemática da temas de investigación
- Restricciones de diseño, procedimientos catalogados y disparadores facilitan la automatización de la BD.
- El problema de la seguridad y de una vista homogénea independiente de los manejadores es un tema actual muy importante
- **LA PRACTICA HACE AL MAESTRO Y LES QUITA LAS DUDAS, por tanto es aconsejable: Revisar el Date: Capítulos 8, 16, 20.**

Conclusiones



Contacto: Carlos Barrón R
cbarron@correo.cua.uam.mx
cbarron99@hotmail.com