

Base de Datos (75.15 / 75.28 / 95.05)

Evaluación Integradora - 26 de julio de 2017

TEMA 20171C4						Padrón: _____
ACR		SQL		Proc.		Apellido: _____
Proc.II		CyT		NoSQL		Nombre: _____
Corrigió:						Cantidad de hojas: _____
Nota:						<input type="checkbox"/> Aprobado <input type="checkbox"/> Insuficiente

Criterio de aprobación: El examen está compuesto por 6 ítems, cada uno de los cuales se corrige como B/B-/Reg/Reg-/M. Se aprueba con nota mayor o igual a 4(cuatro), equivalente a desarrollar el 60% del examen correctamente.

1. (*Álgebra y Cálculo Relacionales*) Considere los siguientes esquemas de relación:

- **Gimnastas**(cod_gimn, nombre, nacionalidad)
- **Exhibiciones**(cod_exhib, fecha, hora, estadio)
- **Puntajes**(cod_gimn, cod_exhib, ijur, puntos)

Los mismos describen el desempeño de los gimnastas participantes en el Campeonato Mundial de Gimnasia Artística. En este mundial, cada gimnasta participa en una o más exhibiciones, y en cada una de ellas es puntuado por una serie de 3 jurados (estos últimos se distinguen a través del índice *ijur* en la tabla *Puntajes*). A partir de la siguiente consulta escrita en el lenguaje del Álgebra Relacional:

$$\text{TEMP} = \pi_{\text{cod_gimn}}(\pi_{\text{cod_gimn}, \text{cod_exhib}}(\text{Puntajes}) - \pi_{\text{cod_gimn}, \text{cod_exhib}}(\sigma_{\text{puntos} \geq 8}(\text{Puntajes})))$$

$$\text{Gimnastas} * (\pi_{\text{cod_gimn}}(\text{Gimnastas}) - \text{TEMP})$$

Se pide:

- a) Explique en palabras qué hace esta consulta.
- b) Para las instancias de relación de *Gimnastas* y *Puntajes* que se muestran en formato de tabla en el *Cuadro 1*, muestre –también en formato de tabla– cuál será la relación resultante al ejecutar la consulta.
- c) Traduzca la consulta al lenguaje del Cálculo Relacional de Tuplas (C.R.T).

GINNASTAS

cod_gimn	nombre	nacionalidad
126	Ralph Trunks	USA
170	Emma Stepherson	REINO UNIDO
191	Kerstin Hawelka	AUSTRIA
202	Lucas Bonarriga	URUGUAY
215	Mariángeles Tejeda	ESPAÑA
396	Koniga Kaya	JAPÓN

PUNTAJES

(...continuación)

cod_gimn	cod_exhib	ijur	puntos	cod_gimn	cod_exhib	ijur	puntos
126	ANILLAS	1	8.33	202	BARRA FIJA	1	9.12
126	ANILLAS	2	8.70	202	BARRA FIJA	2	9.15
126	ANILLAS	3	8.58	202	BARRA FIJA	3	9.18
126	PARALELAS	1	9.10	202	CABALLO c/ ARCOS	1	6.50
126	PARALELAS	2	9.15	202	CABALLO c/ ARCOS	2	6.40
126	PARALELAS	3	9.12	202	CABALLO c/ ARCOS	3	6.85
170	PARALELAS	1	7.80	215	ANILLAS	1	8.15
170	PARALELAS	2	7.70	215	ANILLAS	2	8.20
170	PARALELAS	3	8.05	215	ANILLAS	3	7.90
170	BARRA FIJA	1	9.50	215	BARRA FIJA	1	9.12
170	BARRA FIJA	2	9.45	215	BARRA FIJA	2	9.20
170	BARRA FIJA	3	9.10	215	BARRA FIJA	3	9.18
191	SUELO	1	6.60	396	SUELO	1	8.30
191	SUELO	2	6.80	396	SUELO	2	8.35
191	SUELO	3	6.80	396	SUELO	3	8.20
191	ANILLAS	1	8.05	396	PARALELAS	2	7.95
191	ANILLAS	2	7.80	396	PARALELAS	1	7.80
191	ANILLAS	3	7.70	396	PARALELAS	3	7.60
191	PARALELAS	1	7.01	396	CABALLO c/ ARCOS	1	6.35
191	PARALELAS	2	7.15	396	CABALLO c/ ARCOS	2	6.30
191	PARALELAS	3	7.80	396	CABALLO c/ ARCOS	3	6.20

Cuadro 1: Instancias de relación para el Ejercicio 1 (Álgebra y Cálculo Relacionales).

2. (*SQL*) Se acaba el Campeonato Mundial de Gimnasia Artística y los organizadores desean encontrar a los 3 gimnastas que quedarán en el podio. Para ello, es necesario calcular para cada gimnasta un único promedio que combine todos los puntajes que ha obtenido en todas las exhibiciones en que participó. El podio quedará conformado por aquellos 3 gimnastas que tengan el promedio más alto y que hayan participado en al menos 3 exhibiciones distintas. Escriba una consulta en SQL que devuelva el código de gimnasta, nombre, nacionalidad, puntaje promedio y cantidad de exhibiciones de los gimnastas que obtendrán el oro, plata y bronce, en ese orden.

3. (*Procesamiento de Consultas*) Dados los siguientes esquemas de relación sobre los clientes de una empresa y las facturas de las ventas realizadas:

- Clientes(cod_cliente, nombre, apellido)
- Facturas(numero_factura, cod_cliente, fecha_factura, monto)

se propone construir un índice de árbol por código de cliente (*cod_cliente*) en la tabla *Facturas* para mejorar el costo de resolución de la junta natural en términos de accesos a disco. Se pide:

- a) Calcule el nuevo costo de la junta natural suponiendo que el índice a construir fuera de agrupamiento (*clustering*).
- b) Calcule el nuevo costo de la junta natural en caso que el índice a construir no fuera de agrupamiento.

Asuma en ambos casos que un índice de árbol tendría una altura de 3, y considere la siguiente información de catálogo. Muestre todos los cálculos involucrados.

CLIENTES	FACTURAS
$n(\text{Clientes}) = 1500$	$n(\text{Facturas}) = 45000$
$B(\text{Clientes}) = 150$	$B(\text{Facturas}) = 9000$
	$V(\text{cod_cliente}, \text{Facturas}) = 1500$

4. (*Procesamiento de Consultas II*) Mencione las heurísticas de optimización de consultas estudiadas para bases de datos relacionales.
5. (*Concurrencia y Transacciones*) Un SGBD de dudosa procedencia utiliza el siguiente protocolo de *locking*:

- Inmediatamente antes de leer un ítem se adquiere un *lock* sobre el mismo, que se libera inmediatamente después de la lectura.
- Inmediatamente antes de escribir un ítem se adquiere un *lock* sobre el mismo, que se libera inmediatamente antes del *commit* de la transacción.

Dado el siguiente par de transacciones con los *locks* ya insertados de acuerdo con el protocolo anterior:

$T_1 : b_{T_1}; L_{T_1}(Y); R_{T_1}(Y); U_{T_1}(Y); L_{T_1}(X); R_{T_1}(X); U_{T_1}(X); L_{T_1}(Y); W_{T_1}(Y); U_{T_1}(Y); c_{T_1};$

$T_2 : b_{T_2}; L_{T_2}(X); R_{T_2}(X); U_{T_2}(X); L_{T_2}(X); W_{T_2}(X); L_{T_2}(Y); R_{T_2}(Y); U_{T_2}(Y); L_{T_2}(Y); W_{T_2}(Y); U_{T_2}(X); U_{T_2}(Y); c_{T_2};$

Pruebe que este protocolo no es efectivo para la prevención de anomalías mostrando un solapamiento no serializable de estas transacciones.

Nota: Todos los *locks* son aquí genéricos, no existen *locks* de lectura o de escritura. Sólo cambia la acción del protocolo de *locking* según se trate de una lectura o de una escritura.

6. (*NoSQL*) Indique si las siguientes afirmaciones referentes a MongoDB son verdaderas o falsas, justificando su respuesta.

- a) En un documento JSON pueden existir atributos multivaluados.
- b) La operación **group** del pipeline de agregación puede aparecer varias veces dentro de un mismo pipeline.
- c) La operación **unwind** del pipeline de agregación invierte el orden de presentación de los resultados.
- d) Si un atributo se escoge como *shard key* de una colección, entonces será obligatorio definirlo desde ese momento en todos los documentos de la colección, aunque puede no existir en los documentos previamente existentes de dicha colección.
- e) El mecanismo de *sharding* realiza una fragmentación vertical de los documentos de una colección, asignando subconjuntos de atributos de los documentos a cada *shard*.