

DI- FCT/UNL

31 de maio de 2016

Sistemas de Bases de Dados**2º teste 2015/16****Duração: 2 horas (sem consulta)****Grupo 1**

Considere parte duma base de dados para uma empresa de autocarros para uma rede europeia englobando 32 países (onde os atributos que constituem a chave primária estão sublinhados):

Locais({codLocal,nome,país,população,...})
 Viagens({número,data,cMotorista,matrícula})

Carreiras({número,localOrigem,localDestino})
 Motoristas(cMotorista,nome,sexo,dataNasc,...)

Para cada uma destas tabelas existe um índice secundário de árvore B+ sobre o(s) atributo(s) da chave primária, criado com a ordem das colunas indicada nas tabelas.

O SGBD adoptado usa blocos de 8KB (8192 bytes). Os registos de todas as tabelas têm dimensão variável, sendo que, em média, um registo da tabela de locais ocupa 0,5KB, motoristas ocupa 1KB, um registo de carreiras e de viagens ocupa 128 bytes. Num dado momento a tabela de locais tem 3.200 tuplos, a de carreiras 12.000 tuplos, a de viagens 3.000.000 tuplos e a de motoristas 1.000 tuplos. Um nó da árvore B⁺ pode conter cerca de 100 chaves de pesquisa e sabe-se que o tempo de um seek é dez vezes superior ao da transferência de um bloco ($t_s = 10 * t_T$), e a memória comporta apenas 100 blocos.

Nota: Neste grupo, sempre que se solicitarem exemplos, estes devem ser **exclusivamente** sobre esta base de dados. Além disso, **todas** as respostas deverão conter uma **breve justificação**.

- 1 a) Estime o número de tuplos devolvidos pela seguinte query, assumindo que os números das carreiras são sequenciais e seguidos, iniciando-se em 1.

**SELECT * FROM Viagens NATURAL INNER JOIN Carreiras
 WHERE número <= 10 AND data = '2016-05-31'**

- 1 b) As transações no escalonamento abaixo são executadas em modo READ UNCOMMITTED. Assumindo que as tabelas da base de dados se encontram vazias, indique que valores são escritos pelas instruções DBMS_OUTPUT sabendo que n1 e n2 são variáveis numéricas. Justifique se as transações são serializáveis.

Passo	Transação T1:	Transação T2:
1		SELECT COUNT(*) INTO n2 FROM Carreiras;
2		DBMS_OUTPUT.PUT_LINE(n2);
3		INSERT INTO Carreiras VALUES(12001,...);
4	INSERT INTO Carreiras VALUES(12002,...);	
5		SELECT COUNT(*) INTO n2 FROM Carreiras;
6		DBMS_OUTPUT.PUT_LINE(n2);
7		COMMIT;
8	SELECT COUNT(*) INTO n1 FROM Carreiras;	
9	DBMS_OUTPUT.PUT_LINE(n1);	
10	COMMIT;	

- 1 c) Considere a seguinte lista de eventos envolvendo três transações (as únicas a executar no sistema). Apresente o registo de log de acordo com o algoritmo estudado sabendo que no início o localOrigem é sempre 0.

	T1	T2	T3	sistema
1	begin transaction UPDATE carreiras SET localOrigem = 1 WHERE número = 1;			
2		begin transaction UPDATE carreiras SET localOrigem = 2 WHERE número = 2;		
			begin transaction UPDATE carreiras SET localOrigem = 2 WHERE número = 3;	
3	COMMIT;			
4				CHECKPOINT
5			ROLLBACK;	
6				CRASH!
7				RECOVER

- 1 d) Apresente um escalonamento com apenas duas transações que origine uma exceção quando executado num sistema em modo de isolamento SNAPSHOT ISOLATION.
- 1 e) Para melhorar a disponibilidade do sistema, optou-se por replicar a informação em 3 locais distintos com a variante do protocolo de controlo de concorrência baseado em maioria mas que permite alguns dos locais estarem indisponíveis. Complete toda a tabela seguinte onde cada réplica tem o formato **Carreiras(1,localOrigem,localDestino):Versão**. Um X indica indisponibilidade do local.

Operação	Local 1	Local 2	Local 3	Resultado
read C(1,X,Y)	C(1,1,2):1	C(1,2,1):2	C(1,1,2):1	X= , Y=
write C(1,3,1)		X		Sucesso/Insucesso
write C(1,3,2)	X	X		Sucesso/Insucesso

Grupo 2

Nota: A resposta a cada uma das alíneas deste grupo **não pode em caso algum exceder uma página**.

- 2 a) Explique porque motivos alguns sistemas de gestão de bases-dados só consideram árvores profundas à esquerda (left-deep join trees) em vez de árvores arbitrárias.
- 2 b) Indique o propósito de cada uma das fases do algoritmo de confirmação em duas fases (2 *Phase Commit* – 2PC).
- 2 c) Alguns algoritmos de recuperação baseados em *logs* realizam operações de pontos de controlo (*checkpoints*). Descreva esta operação e explique as vantagens de as realizar pelo SGBD.