

1. Mějme 200 bloků dat, které chceme třídit pomocí algoritmu dvoufázového třídění vícecestným sléváním (TPMMS). Algoritmus čte data z disku a zapisuje výsledek na disk. Kolik maximálně budeme potřebovat přístupů na disk?
 Správná odpověď: 800
 1. fáze: načtení po blocích do op. paměti tj. sekvenční čtení po blocích = 200 přístupů na disk, utřídění = 0 přístupů na disk, zápis utříděných dat = 200 přístupů na disk
 2. fáze: čtení dat z disku do proudů = 200 přístupů na disk, zápis dat na disk = 200 přístupů na disk
 Celkem 800 přístupů na disk.
2. Jak může vypadat undo-log pro transakci $T: B := A+1, A := B*3$ s počátečními hodnotami $A=1, B=3$?
 Správná odpověď: $\langle T, \text{start} \rangle, \langle T, B, 3 \rangle, \langle T, A, 1 \rangle, \langle T, \text{abort} \rangle$
 Undo-log zapisuje staré hodnoty.
3. Jakým způsobem se provádí index-scan při čtení relace R pomocí indexu ve formě B-stromu?
 Správná odpověď: relace se prochází po záznamech, je potřeba $T(R)$ diskových operací
4. Do jednoho bloku se vejde buď 40 záznamů nebo 100 dvojic (klíč, ukazatel). Kolik bloků zabere dvouúrovňový index tabulky obsahující 1 mil. záznamů v případě použití řídkého (sparse) indexu v obou úrovních. (Do indexu nepočítáme vlastní záznamy tabulky.)
 Správná odpověď: 253 bloků
 1 mil. záznamů zabere $1\,000\,000/40 = 25\,000$ bloků. Takže potřebujeme 25 000 indexů první úrovně a na ně $25\,000/100 = 250$ bloků. Dále potřebujeme 250 indexů druhé úrovně, které zaberou $250/100 = 3$ bloky. Celkem $250 + 3 = 253$ bloků.
5. Předpokládejme, že máme RAID 4 systém disků 3+1. Došlo k výpadku 2. datového disku. Obsah 1. a 3. disku jsou: 10101001, 11001111. Redundantní disk s kontrolními součty obsahuje: 01101001. Jaký byl obsah 2. disku před výpadkem?
 Správná odpověď: 11000000
 Data na paritním disku jsou rovna výrazu $[(\text{data z 1. disku}) \text{ XAND } (\text{data z 2. disku})] \text{ XAND } (\text{data z 3. disku})$.
6. Po provedení operace projekce je (neuvažujeme duplicitu záznamů) je velikost záznamu na výstupu vzhledem k velikosti záznamu na vstupu:
 Správná odpověď: vždy menší nebo stejná
 Projekce je ekvivalentní v SQL příkazu SELECT, který může na výstup dát buď celý záznam (SELECT * ...) nebo jenom některé atributy relace (SELECT A,C,D...).
7. Rozvrhy S_1, S_2 jsou konfliktně ekvivalentní, pokud lze jeden transformovat na druhý pomocí:
 Správná odpověď: posloupnosti výměn nekonfliktních akcí
8. Mějme bloky velikosti 2048 B. Každý blok obsahuje na svém začátku hlavičku obsahující tabulku offsetů, která obsahuje ukazatele velikosti 2 B ukazující na záznamy pevné délky 16 B ukládané na konce bloku. Pro snazené záznamy obsahuje tabulka offsetů "náhrobky" pro případné visící ukazatele. Předpokládejme, že za jeden den jsou pro každý blok průměru provedena dvě vložení nového záznamu a jedno rušení existujícího záznamu. Pokud jsou bloky původně prázdné, za kolik dní budou bloky plné?
 Správná odpověď: 102
 Vložení nového záznamu přidá do bloku $2*(16 + 2)$ B, rušení záznamu odebere z bloku 16 B. Takže denně přibude do bloku $2*(16 + 2) - 16 = 20$ B. Za $2048/20 = 102$ dní se bloky zaplní.
9. Mějme rozšířené (extensible) hešování, které momentálně používá adresář se 128 položkami. Kolik maximálně ukazatelů adresáře může ukazovat na jednu stejnou oblast (bucket)?
 Správná odpověď: 64
 Položky adresáře ukazují na oblasti s poloviční kapacitou a ty na oblasti o úroveň níže opět s poloviční kapacitou...
10. Mějme relace $X(a,b)$ a $Y(b,c)$. Kdy je výsledkem přirozeného spojení $X \bowtie Y$ prázdná relace?
 Správná odpověď: Když atributy b relací X a Y jsou disjunktní množiny.
 Přirozené spojení "spojí" řádky relací X, Y , které mají stejnou hodnotu atributu b .