

Tecnologia delle basi di dati

Prova di autovalutazione — 22 maggio 2006

Domanda 1

Il check-point, nei vari DBMS, viene realizzato in due modi diversi:

1. in alcuni sistemi si prende nota delle transazioni attive e si rifiutano (momentaneamente) nuovi commit
2. in altri si inibisce l'avvio di nuove transazioni e si attende invece la conclusione (commit o abort) delle transazioni attive

Spiegare, intuitivamente, le differenze che ne conseguono sulla gestione delle riprese a caldo.

Domanda 2

Illustrare, brevemente, ma in modo ordinato, le differenze nelle tecniche di implementazione fra i livelli di isolamento **SERIALIZABLE** e **REPEATABLE READ**, facendo anche riferimento alle diverse strutture fisiche, primarie e secondarie, che possono essere coinvolte. Spiegare, quindi, perché le conseguenti differenze di prestazioni possono essere in alcuni casi enormi e in altri relativamente piccole.

Domanda 3

Nel controllo di concorrenza basato su timestamp una transazione viene uccisa se essa, avendo un timestamp pari a ts , richiede una scrittura su un elemento x tale che $ts < WTM(x)$. Alcuni autori hanno notato che questa uccisione non è in effetti necessaria. Spiegare perché.

Domanda 4

Una tecnica per il controllo di concorrenza diversa da quelle viste nel corso va sotto il nome di “concorrenza basata sulla validazione” ed è basata sulle seguenti ipotesi e principi.

- Ogni transazione è divisa in tre fasi: (i) prima esegue tutte le proprie letture e svolge le proprie elaborazioni “privatamente” (cioè in memoria, senza scrivere sulla base di dati); (ii) poi chiede allo scheduler il permesso di andare in commit (fase di “validazione”) e (iii) se autorizzata, scrive.
 - Lo scheduler ricorda per ogni transazione (1) gli insiemi di dati che essa ha letto $RSET(T)$ e ha scritto (o vuole scrivere) $WSET(T)$, (2) in quale “stato” si trova: **INIZ** (iniziata e non ancora validata), **VAL** (validata, ma con scritture da completare), **CPL** (completata, con l'esecuzione di tutte le scritture) e (3) i corrispondenti istanti di inizio $INIZ(T)$, validazione $VAL(T)$ e completamento $CPL(T)$.
 - Ad una transazione T la validazione viene concessa in tutti i casi esclusi i seguenti, in cui viene negata:
 1. esistono (nell'istante t in cui si esamina la validazione di T):
 - una transazione T' validata o completata, che non era completata quando T è iniziata (e quindi risulterebbe $CPL(T') > INIZ(T)$)
 - un dato $x \in RSET(T) \cap WSET(T')$
 2. esistono (nell'istante t in cui si esamina la validazione di T):
 - una transazione T' validata ma non completata (e quindi risulterebbe $CPL(T') > t = VAL(T)$)
 - un dato $x \in WSET(T) \cap WSET(T')$
1. dimostrare (formalmente o almeno intuitivamente) che la classe di schedule prodotta è propriamente contenuta in CSR.
 2. individuare quando possono essere eliminate le informazioni su una transazione, al fine di evitare la crescita eccessiva dei dati da gestire.