

Atividade de revisão sobre recuperação de falhas (válida para 30% da nota da 2ª. unidade)

1. Para que é usado o log do sistema? Quais são os tipos mais comuns de entradas em um log do sistema?

Os logs do sistema são usados para registrar eventos e atividades em um sistema de computador. Solução de problemas, monitoramento, segurança e conformidade são importantes. Tipos comuns de entradas de syslog incluem sistema, segurança, aplicativo, rede, hardware, auditoria, sistema operacional, autenticação, autorização e eventos de desempenho.

2. O que são entradas de log tipo UNDO, REDO e checkpoints?

O log UNDO registra informações para desfazer ou desfazer alterações de transações incompletas. As entradas de log do tipo REDO registram informações para reverter as alterações feitas por transações confirmadas. Os checkpoints são quando as transações confirmadas são gravadas permanentemente no banco de dados. Juntas, essas entradas de log garantem a integridade dos dados e a recuperação confiável no sistema de gerenciamento de banco de dados.

3. Como o caching é utilizado pelo subsistema de recuperação do SGBD?

Um SGBD utiliza o caching para otimizar a recuperação de dados. O caching envolve armazenar temporariamente dados frequentemente acessados em memória rápida, reduzindo a necessidade de acessos lentos ao disco. Isso é feito por meio do buffer pool, onde páginas de dados são mantidas em memória, e pelo cache de resultados de consultas e índices. O caching ajuda a melhorar o desempenho ao minimizar o acesso ao disco, embora seja necessário equilibrar o tamanho do cache com a disponibilidade de memória.

4. Qual é a diferença entre a atualização no local e a sombra, com relação ao tratamento de imagem do bloco anterior (BFIM) e imagem do bloco posterior (AFIM)? Para responder, revise a seção “23.1.2 Caching (Buffering) de Blocos de Disco” do livro Elsmare e Navathe 6ª. edição.

A "atualização no local" refere-se à modificação direta dos valores de um registro existente, enquanto a "sombra" se refere à prática de manter cópias de segurança de dados para fins de recuperação e auditoria.

5. O que significam os termos steal/no-steal e force/no-force com relação à escrita adiantada do log. Para responder, revise a seção “23.1.3 Logging write-ahead”, ou seja, escrita adiantada de log do livro Elsmare e Navathe 6ª. edição.

Os termos "steal" e "no-steal" referem-se à permissão para escrever páginas modificadas no disco antes ou após a confirmação das transações que as modificaram. "Force" e "no-force" se

relacionam à exigência de escrever páginas no disco imediatamente após a confirmação ou adiar a escrita para melhorar o desempenho. Essas estratégias afetam a persistência dos dados e o desempenho do sistema e são escolhidas com base nos requisitos de cada ambiente de banco de dados.

6. Defina a técnica de recuperação com atualização adiada. Quais são as vantagens e desvantagens dessa técnica? Por que ela é chamada de método NO-UNDO/REDO?

A técnica de recuperação com atualização adiada atrasa a escrita de atualizações de banco de dados no disco até que as transações sejam confirmadas. Vantagens incluem melhor desempenho e menor overhead, enquanto desvantagens envolvem risco de perda de dados, reconstrução complexa e necessidade de controle rigoroso de transações. Essa técnica é chamada de "NO-UNDO/REDO" porque elimina a etapa de UNDO durante a recuperação, mas ainda exige a etapa de REDO para trazer o banco de dados a um estado consistente após falhas.

7. Como a recuperação pode tratar de operações com transação que não afetam o banco de dados, como a impressão de relatórios por uma transação?

Operações de transações que não afetam o banco de dados, como operações de leitura, geralmente não requerem tratamento especial durante a recuperação em sistemas de gerenciamento de banco de dados (SGBDs). Elas são frequentemente ignoradas no processo de recuperação, pelo fato de não alterarem os dados persistentes. Entretanto, para garantir a consistência da saída e manter registros para rastreabilidade e auditoria são considerações importantes para manter a confiabilidade do sistema. Em caso de falha, essas operações podem ser descartadas sem impacto duradouro.

8. Discuta a técnica de recuperação com atualização imediata. Quais são as vantagens e desvantagens da atualização imediata? Qual é a diferença entre os algoritmos de UNDO/REDO e UNDO/ NO-REDO para a recuperação com atualização imediata?

A técnica de recuperação com atualização imediata em sistemas de banco de dados envolve escrever as atualizações no disco imediatamente após ocorrerem, garantindo a durabilidade dos dados.

Vantagens: durabilidade garantida e simplificação da recuperação.

Desvantagens: envolvem impacto no desempenho e overhead de E/S.

A diferença entre os algoritmos de UNDO/REDO e UNDO/NO-REDO nesse contexto reside na necessidade de reexecutar as alterações confirmadas durante a recuperação; o UNDO/REDO requer isso, enquanto o UNDO/NO-REDO não precisa refazer as alterações confirmadas.

9. Sobre a técnica de recuperação ARIES responda: O que são números de sequência de log (LSNs) e como eles são usados? Que informação a Tabela de Páginas Sujas e a Tabela de Transações contêm? Descreva como o checkpoint é usado no ARIES.

A técnica de recuperação ARIES utiliza LSNs para rastrear a ordem das operações no log de transações. A Tabela de Páginas Sujas registra páginas de dados modificadas mas não escritas no disco, enquanto a Tabela de Transações acompanha o estado das transações. O checkpoint é um ponto no tempo que grava o estado do banco de dados no log, incluindo informações das tabelas mencionadas. Isso permite uma recuperação eficiente e consistente após falhas, pois proporciona um ponto de partida conhecido.

10. Considere o log abaixo correspondente a determinado plano de execução no ponto de uma falha do sistema para quatro transações T1, T2, T3 e T4 e responda:

```
[start_transaction, T1]
[read_item, T1, A]
[read_item, T1, D]
[write_item, T1, D, 20, 25]
[commit, T1]
[checkpoint]
[start_transaction, T2]
[read_item, T2, B]
[write_item, T2, B, 12, 18]
[start_transaction, T4]
[read_item, T4, D]
[write_item, T4, D, 25, 15]
[start_transaction, T3]
[write_item, T3, C, 30, 40]
[read_item, T4, A]
[write_item, T4, A, 30, 20]
[commit, T4]
[read_item, T2, D]
[write_item, T2, D, 15, 25]
```

a) Caso usemos o protocolo de atualização imediata com checkpoint, especifique quais transações são revertidas, quais operações no log são refeitas e quais (se houver) são desfeitas, e se ocorre alguma reversão em cascata.

Transações revertidas: Nenhuma transação foi revertida.

Operações refeitas: As operações de T4 e T3 são refeitas após o checkpoint.

Operações desfeitas: Nenhuma operação foi desfeita.

Reversão em cascata: Não ocorre reversão em cascata, já que nenhuma transação é revertida.

b) Caso usemos um protocolo de atualização adiada mostre como o log seria diferente ao remover as entradas desnecessárias; depois especifique quais operações no log são refeitas e quais são ignoradas.

```
[start_transaction, T1]
[read_item, T1, A]
[read_item, T1, D]
[write_item, T1, D, 20, 25]
[commit, T1]
[checkpoint]
[start_transaction, T2]
[read_item, T2, B]
[write_item, T2, B, 12, 18]
[start_transaction, T4]
[read_item, T4, D]
[write_item, T4, D, 25, 15]
[start_transaction, T3]
[write_item, T3, C, 30, 40]
[read_item, T4, A]
[write_item, T4, A, 30, 20]
[commit, T4]
[read_item, T2, D]
[write_item, T2, D, 15, 25]
```

Operações refeitas: Todas as operações a partir do checkpoint (operação 6) seriam refeitas.

Operações ignoradas: As operações de leitura e de escrita que ocorreram antes do checkpoint seriam ignoradas, já que não afetam a recuperação.