

**UNIVERSIDADE DE BRASÍLIA**  
**Faculdade do Gama**

**Sistemas de Banco de Dados 1**

**Estudo sobre recuperação a falhas em banco de dados**

**Guilherme Soares Rocha**  
**211039789**

Brasília, DF  
2023

### **A) Existem falhas no bancos de dados? Quais?**

Existem diversas falhas que podem ocorrer em bancos de dados. Algumas delas são:

**Perda de dados:** Isso pode ocorrer devido a erros humanos, falhas de hardware ou software, desastres naturais, ataques cibernéticos ou corrupção de dados. A perda de dados pode resultar em informações irre recuperáveis e afetar a integridade e confiabilidade do banco de dados.

**Inconsistência de dados:** Isso ocorre quando os dados em um banco de dados estão desatualizados, duplicados ou não correspondem às regras de integridade definidas. Pode acontecer devido a operações mal executadas, problemas de sincronização entre sistemas ou falta de validações adequadas.

**Conflitos de concorrência:** Quando múltiplos usuários acessam e modificam simultaneamente o mesmo conjunto de dados, podem ocorrer conflitos de concorrência. Isso pode resultar em problemas como perda de atualizações, leituras incorretas ou bloqueios prolongados, afetando a consistência e o desempenho do banco de dados.

**Problemas de segurança:** Bancos de dados podem ser alvo de ataques cibernéticos, como invasões, injeção de SQL, roubo de dados sensíveis ou acesso não autorizado. Falhas na implementação de medidas de segurança adequadas podem expor o banco de dados a riscos de segurança.

**Problemas de desempenho:** Bancos de dados podem enfrentar problemas de desempenho devido a consultas mal otimizadas, falta de índices adequados, configurações inadequadas ou aumento no volume de dados. Isso pode resultar em tempos de resposta lentos, degradação do desempenho ou até mesmo falhas no sistema.

É importante implementar práticas adequadas de gerenciamento de banco de dados, manutenção regular, backups e medidas de segurança para evitar essas falhas e garantir a integridade, disponibilidade e confiabilidade dos dados.

### **B) Existe como recuperar as falhas no banco de dados (bd)? Quais? (explique cada uma delas)**

Sim, existem mecanismos e técnicas para recuperar falhas no banco de dados. Abaixo estão algumas das principais técnicas de recuperação de falhas:

- **Backup e restauração:** Essa técnica envolve a criação regular de cópias de segurança dos dados do banco de dados e a capacidade de restaurar esses backups em caso de falhas. Os backups podem ser completos ou incrementais, e a restauração é feita para trazer o banco de dados de volta a um estado consistente antes da ocorrência da falha.

- **Log de transações e recuperação baseada em log:** O log de transações registra todas as operações de modificação feitas no banco de dados. Com base no log, é possível reconstruir o estado do banco de dados até o ponto de falha e aplicar as operações de recuperação necessárias para restaurar a consistência dos dados.

- **Ponto de verificação (checkpoint):** O ponto de verificação é um mecanismo para garantir que todas as alterações feitas até certo momento no banco de dados tenham sido gravadas em disco. Em caso de falha, o banco de dados pode ser restaurado para o último ponto de verificação conhecido, reduzindo a quantidade de registros de log que precisam ser aplicados durante a recuperação.

- **Replicação e alta disponibilidade:** A replicação envolve a criação de cópias dos dados do banco de dados em servidores adicionais. Se ocorrer uma falha em um servidor, um dos servidores de réplica pode assumir o controle e continuar fornecendo acesso aos dados. Isso melhora a disponibilidade dos dados e permite uma recuperação mais rápida em caso de falha.

- **Clustering e failover:** Em ambientes de cluster, vários nós de banco de dados estão interconectados para fornecer alta disponibilidade e tolerância a falhas. Se um nó falhar, outro nó assume as operações e mantém o serviço em funcionamento sem interrupção.

- **Recuperação de transações distribuídas:** Em sistemas distribuídos, onde várias instâncias de banco de dados estão envolvidas, a recuperação de falhas pode envolver protocolos de coordenação e consenso para garantir que as transações sejam recuperadas de forma consistente em todos os nós.

A escolha da técnica apropriada depende das características do sistema, dos requisitos de disponibilidade e da criticidade dos dados armazenados.

### **C) O que significam as siglas ou expressões relacionadas abaixo (em bd)?**

- **log em banco de dados**
- **undo**
- **redo**
- **checkpoint em banco de dados**
- **data lake**

- **Log:** O log é um registro sequencial de todas as transações que ocorrem no sistema. Ele registra todas as alterações feitas nos dados, incluindo inserções, atualizações e exclusões. O log é usado para garantir a consistência e a integridade dos dados, permitindo a recuperação em caso de falhas ou problemas no sistema.

- **Undo:** O undo é uma funcionalidade do banco de dados que permite reverter ou desfazer transações já concluídas. Ele armazena informações necessárias para desfazer ou cancelar operações, como a reversão de atualizações em caso de falhas ou erros. O undo é usado principalmente em transações de banco de dados para garantir a consistência e a atomicidade.

- **Redo:** O redo é uma funcionalidade do banco de dados que permite repetir ou refazer transações já concluídas. Ele armazena informações necessárias para refazer operações que podem ter sido perdidas devido a falhas ou interrupções do sistema. O redo é usado para garantir a durabilidade e a recuperação de dados em caso de problemas.

- **Checkpoint:** O checkpoint é um processo de gravação de todas as transações pendentes do log no armazenamento permanente. Ele marca um ponto de verificação no log, indicando que todas as transações até aquele momento foram gravadas em disco. O checkpoint é usado para fins de recuperação e recuperação de falhas, garantindo que as alterações estejam persistentes e disponíveis em caso de falhas de sistema.

- **Data lake:** Um data lake é um repositório centralizado de grande volume e variedade de dados brutos ou não processados. É um armazenamento de dados que permite armazenar dados estruturados e não estruturados em sua forma original, sem a necessidade de uma estrutura rígida ou pré-definida. O data lake é usado como uma fonte de dados para análises avançadas, processamento de big data e descoberta de insights a partir de dados em sua forma bruta.