

OTIMIZAÇÃO, MONITORAMENTO E OPERAÇÃO EM BANCOS DE DADOS





PUC Minas
Virtual

UNIDADE II – TÉCNICAS DE OTIMIZAÇÃO DE BANCOS DE DADOS



PUC Minas
Virtual

INDEXAÇÃO

COMO O BANCO ARMAZENA OS DADOS?

- Os princípios que iremos mostrar nesta conversa se aplicam aos bancos de dados relacionais.
- Quando utilizamos outros tipos de bancos de dados precisaremos fazer uma pequena adaptação conceitual mas podemos verificar que todos eles são plenamente aplicáveis.
- Vamos falar da organização do armazenamento e do mecanismo de recuperação índice!
- Registros são armazenados em páginas de 8Kb, que tem um *header* que contém links para as outras páginas e *hash*.
- São agrupados de 8 em 8 formando este conjunto de extensão.

| | | | |
|---|-------|---|-------|
| H | DADOS | H | DADOS |
| H | DADOS | H | DADOS |
| H | DADOS | H | DADOS |
| H | DADOS | H | DADOS |

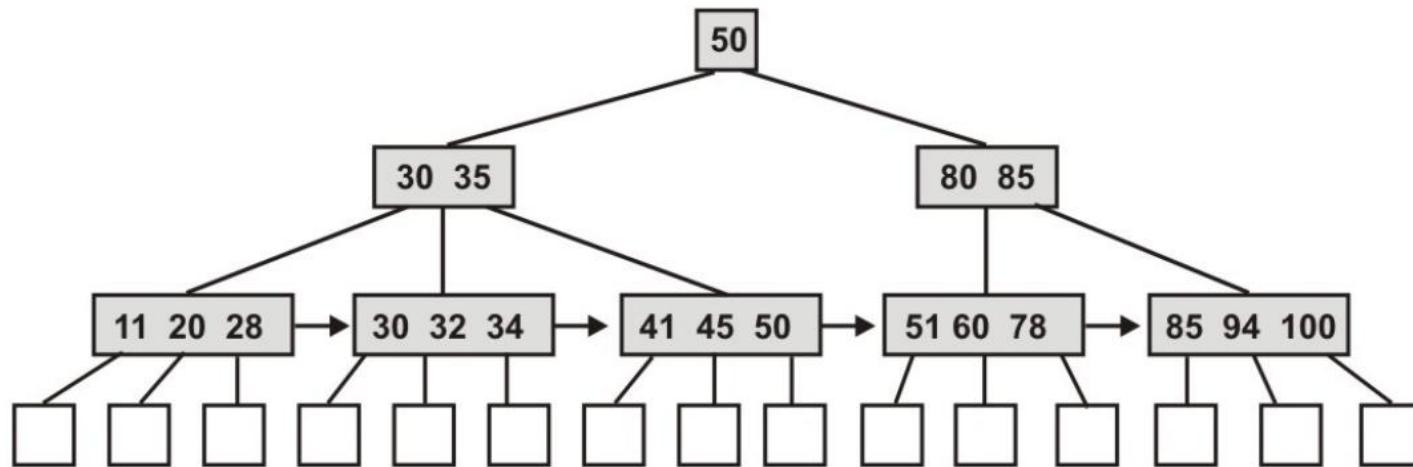
Fonte: DEVMEDIA, 2007.

UM RECURSO MUITO PODEROSO

- Os registros de dados não são armazenados em uma ordem específica, e não existe uma ordenação sequente para as páginas de dados.
- É utilizado um conceito de organização que se chama **pilha**, e quando a capacidade das páginas de dados se esgota estabelece-se um link que as liga.
- **Vamos a duas formas básicas de método de acesso para recuperação de dados:**
- **Exame de tabela:** examina todas as páginas de dados das tabelas, começando do início da tabela passando por todos os registros, extraíndo aqueles que satisfazem aos critérios da consulta.
- **Usando índices:** percorrendo a estrutura da árvore do índice para localizar os registros.

UM RECURSO MUITO PODEROSO

- Um índice para uma tabela é uma organização de dados que permite que permita acelerar o processamento de consultas, por comparação, extraíndo somente aqueles registros necessários para satisfazerem os critérios passados pela consulta.

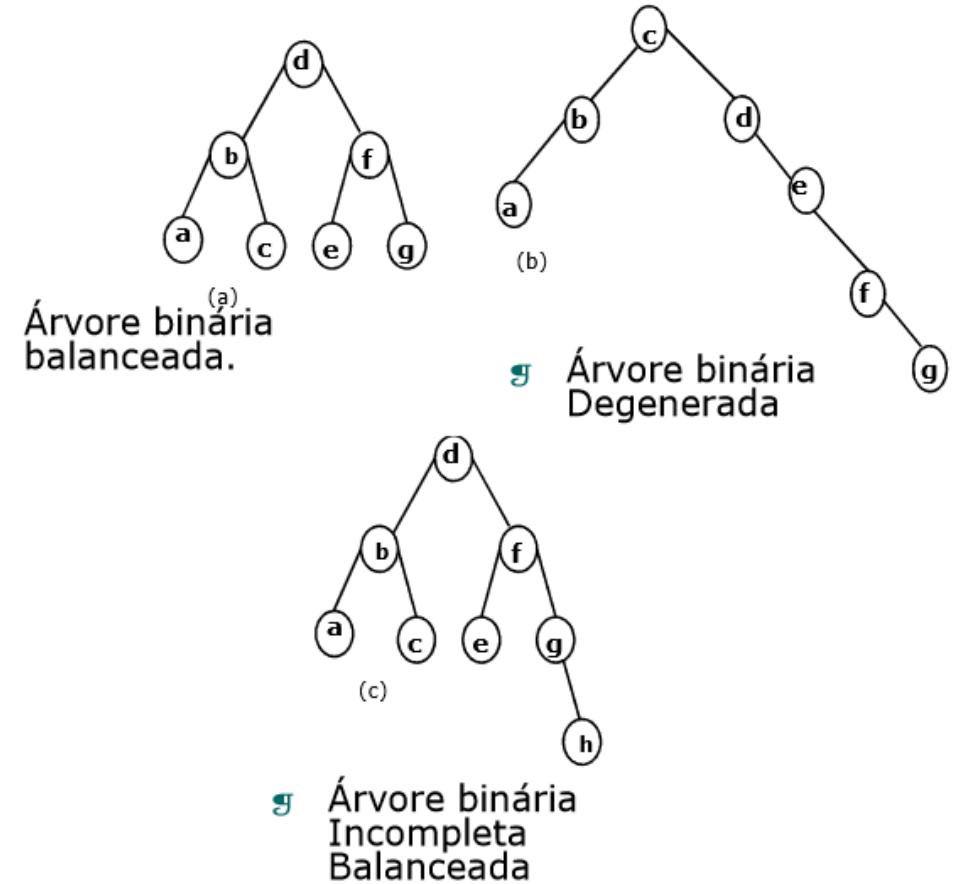


Fonte: PEDROZO & VAZ, 2007.

- Uma árvore B+ é uma árvore balanceada cujas folhas contêm uma sequência de pares chaves-ponteiro.
- As chaves são ordenadas pelo seu valor.

UM RECURSO MUITO PODEROSO

- Esta árvore é do tipo balanceada e projetada para facilitar a interação com dispositivos de armazenamento secundário como disco magnético.
- A ideia é tentar melhorar a questão do I/O nestes dispositivos diminuindo o tempo de acesso.
- Se o braço puder apontar para o lugar certo de recuperação do registro, agiliza-se a leitura.



Fonte: GHISLENI, 2021.

UM OLHO NAS MÉTRICAS E OUTRO NO OTIMIZADOR

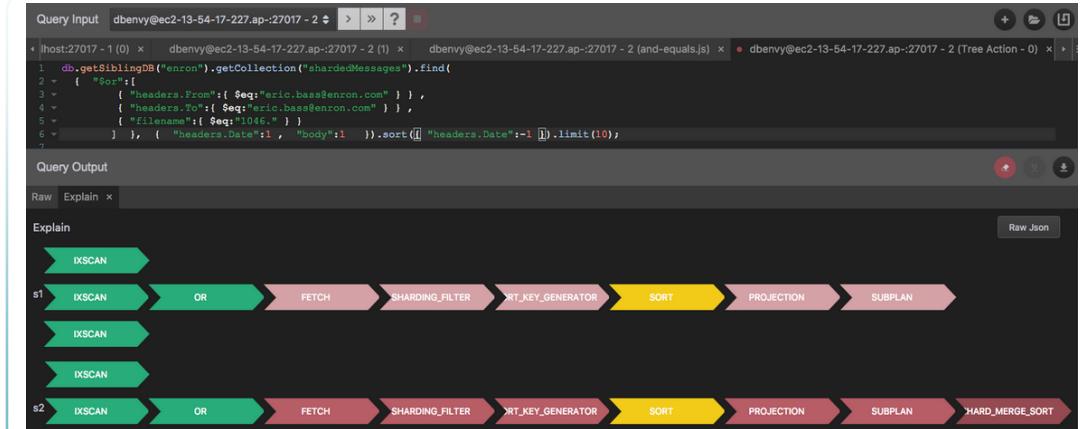
- Então vamos supor que eu seja chamado para diagnosticar um problema, observe o cenário:
- Olho para as métricas críticas de desempenho do sistema operacional, duas coisas se destacam: a CPU e o I/O no disco primário. Tanto a média de carga da CPU quanto as latências de E/S do disco sugerem que o sistema precisa de mais capacidade de CPU e I/O.



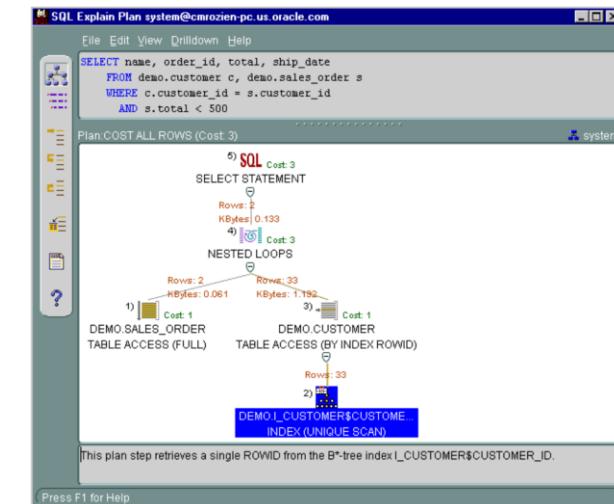
Fonte: <https://livecoins.com.br/permissoa-para-bancos-guardarem-bitcoin-pode-ser-sinal-de-desespero-do-sistema-financeiro/>.

UM OLHO NAS MÉTRICAS E OUTRO NO OTIMIZADOR

- Um componente de subsistema muito importante é o otimizador de consultas que é quem define qual o método será mais eficiente para recuperar os dados.
- Entra aqui um recurso que quase todos os bancos tem que é o **EXPLAIN** ainda não vamos entrar neste assunto.
- Vamos apenas nos ater aos índices e seus tipos.



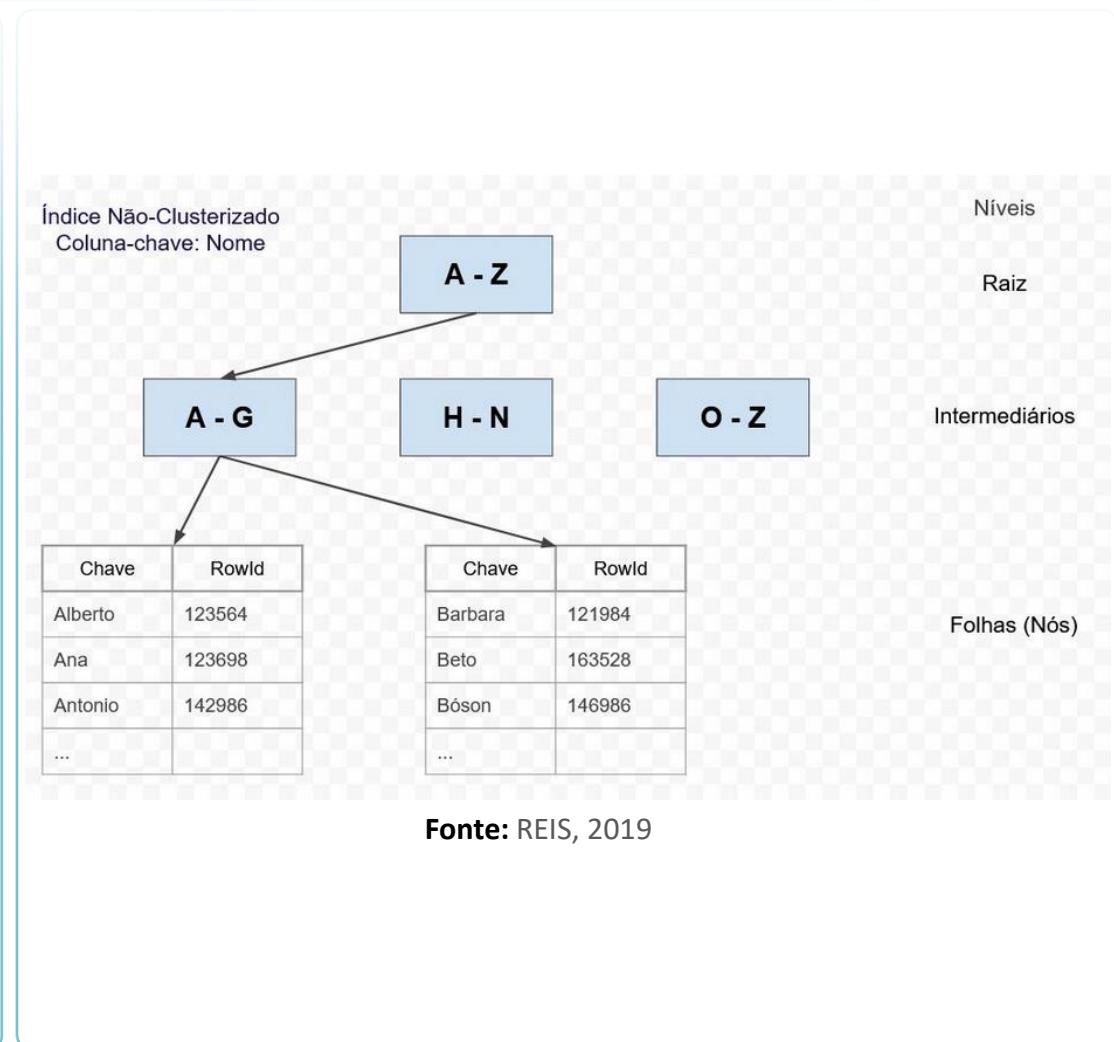
Fonte: HARRISON, 2017



Fonte: ORACLE, 2022

TIPOS DE ÍNDICES

- Índices clusterizados, não clusterizados e de *hashing*
- Índice de árvore B*
- Índice de *bitmap*
- Índice TTL (*Time To Live*)
- Índice Geoespacial
- Índice de várias tabelas
- Índice booleano
- Índice funcional



VOLTANDO AO PROBLEMA...

- Então vamos supor que eu seja chamado para diagnosticar um problema, observe o cenário:
- Olho para as métricas críticas de desempenho do sistema operacional, duas coisas se destacam: a CPU e o I/O no disco primário. Tanto a média de carga da CPU quanto as latências de E/S do disco sugerem que o sistema precisa de mais capacidade de CPU e I/O.



Fonte: <https://livecoins.com.br/permissoa-para-bancos-guardarem-bitcoin-pode-ser-sinal-de-desespero-do-sistema-financeiro/>.

UTILIZAÇÃO DE ÍNDICES

- Lembrem-se novamente das dicas de DESIGN, conheçam o problema de negócio.
- Uma pergunta simples seria a quantidade de transações de gravação e quantas de recuperação teremos nas funcionalidades principais.
- Qual a prioridade do negócio?
- Índice é uma das ferramentas de otimização mais conhecidas e utilizadas pelos desenvolvedores de bancos de dados.



Fonte: <https://livecoins.com.br/permissoa-para-bancos-guardarem-bitcoin-pode-ser-sinal-de-desespero-do-sistema-financeiro/>.

UTILIZAÇÃO DE ÍNDICES

- A indexação em tabelas pode aumentar significativamente a performance em consultas ao banco de dados.
- Só que pode diminuir a velocidade de transações como *inserts* e *updates*.
- **Porque é necessário balancear a árvore!**
- Senão não faz sentido manter o índice, concorda?
- **Reorganização de índices devem ser feitos periodicamente por este motivo!**



Fonte: <https://livecoins.com.br/permissoa-para-bancos-guardarem-bitcoin-pode-ser-sinal-de-desespero-do-sistema-financeiro/>.

OTIMIZAÇÃO, MONITORAMENTO E OPERAÇÃO EM BANCOS DE DADOS

■ REFERENCIAS BIBLIOGRÁFICAS

HARRISON, Guy & Michael. MongoDB Performance Tuning: Optimizing MongoDB Databases and their Applications. Apress. 2021.

DEVMEDIA, 2007. Entendendo e usando índices. Disponível em: <<https://www.devmedia.com.br/entendendo-e-usando-indices/6567#1>>. Acesso em: 28 Fev. 2023.

PEDROZO, Wendel Góes & **VAZ** Maria Salete Marcom Gomes, 2007. Arquitetura para Seleção de Índice no SGBD PostgreSQL, utilizando abordagem baseada em custos do Otimizador . Disponível em: <<https://www.inf.unioeste.br/~clodis/BDI/A11.pdf>>. Acesso em: 19 Mar. 2023.

CANTARELLI, Elisa Maria Pivetta, 2022. Árvores Binárias Balanceadas. Disponível em: <<https://docplayer.com.br/10942308-Arvores-binarias-balanceadas.html>>. Acesso em: 19 Mar. 2023.

HARRISON, Guy, 2017. Getting started with MongoDB explain(). Disponível em: <<https://www.dbkoda.com/blog/2017/11/12/MongoDBExplain>>. Acesso em: 19 Mar. 2023.

ORACLE, 2022. Oracle Enterprise Manager Database Tuning with the Oracle Tuning Pack. Disponível em: <https://docs.oracle.com/cd/B10501_01/em.920/a86647/sqlexpla.htm>. Acesso em: 19 Mar. 2023.

REIS, Fábio 2019. O que são Índices em Bancos de Dados – Indexação em Tabelas. Disponível em: <<http://www.bosontreinamentos.com.br/bancos-de-dados/o-que-sao-indices-em-bancos-de-dados-indexacao-em-tabelas/>>. Acesso em: 19 Mar. 2023.

