

Para os exercícios abaixo, você pode fazer considerações necessárias sobre índices, tamanho de bloco, quantidade de tuplas ou quais outros dados que não foram fornecidos.

- 1) Defina plano de execução de consulta. Dê um exemplo.
- 2) Suponha a consulta $\sigma_{\text{nomeAluno} = \text{"Luis Fonsi"} \wedge \text{cidadeAluno} = \text{"San Juan"}}$ (Aluno). Considere que a tabela Aluno está organizada pelo número do aluno e que existe um índice primário esparsa e outro índice secundário por intérprete. Apresente três diferentes planos de execução consulta para a expressão fornecida.
- 3) Suponha a consulta $\sigma_{\text{numeroAluno} > 10000}$ (Aluno) e que a Relação **Aluno** está ordenada fisicamente por número do aluno qual seria o custo dessa consulta para busca linear? É possível fazer uma busca binária neste caso? Justifique sua resposta.
- 4) Suponha a consulta $\sigma_{\text{numeroAluno} = 123456}$ (Aluno) e que a Relação **Aluno** está ordenada fisicamente por número do aluno. Qual seria o custo dessa consulta para busca linear e busca binária, sabendo que número do aluno é chave primária? Se houver um índice primário esparsa (ordenado) qual seria o custo da consulta usando este índice, considerando que cabem 20 entradas de índice em cada bloco?
- 5) Suponha que para o cálculo da consulta $\sigma_{\text{nomeAluno} = \text{"John Lennon"}}$ (Aluno) haja um índice árvore-B+ primário para o atributo *nomeAluno*, com 10 entradas de índice por nó da árvore. Considerando os dados fornecidos no final da lista, calcule o custo para a busca linear, busca binária e busca com índice, considerando que *nomeAluno* não é chave (primária ou candidata).
- 6) Suponha que para o cálculo da consulta $\sigma_{\text{nomeAluno} = \text{"John Lennon"}}$ (Aluno) haja um índice árvore-B+ primário para o atributo *nomeAluno*, com 10 entradas de índice por nó da árvore. Considerando os dados fornecidos no final da lista, calcule o custo para a busca linear, busca binária e busca com índice, considerando que *nomeAluno* é chave (primária ou candidata).
- 7) Suponha que para o cálculo da consulta $\sigma_{\text{cidadeAluno} = \text{"Fortaleza"}}$ (Aluno) haja um índice árvore-B secundário para o atributo *cidade*, com 30 entradas de índice por nó da árvore. Considerando os dados fornecidos no final da lista, calcule o custo para a busca linear e busca com índice, considerando que *cidadeAluno* não é chave primária.
- 8) Suponha que para o cálculo da consulta $\sigma_{\text{nomeAluno} \geq \text{"Valeska Popozuda"}}$ (Aluno) haja um índice árvore-B+ primário para o atributo *nomeAluno*, com 20 entradas de índice por nó da árvore. Considerando os dados fornecidos no final da lista, calcule o custo para a busca linear e busca com índice, considerando que *nomeAluno* não é chave primária.
- 9) Suponha que para o cálculo da consulta $\sigma_{\text{nomeAluno} \geq \text{"Valeska Popozuda"}}$ (Aluno) haja um índice árvore-B+ secundário para o atributo *nomeAluno*, com 20 entradas de índice por nó da árvore. Considerando os dados fornecidos no final da lista, calcule o custo para a busca linear e busca com índice, considerando que *nomeAluno* não é chave primária.
- 10) Suponha que para o cálculo da consulta $\sigma_{\text{nomeAluno} < \text{"John Lennon"}}$ (Aluno) haja um índice árvore-B secundário para o atributo *nomeAluno*, com 20 entradas de índice por nó da árvore. Considerando os dados fornecidos no final da lista, calcule o custo para a busca linear e busca com índice, considerando que *nomeAluno* não é chave primária.
- 11) Suponha que para o cálculo da consulta $\sigma_{\text{nomeAluno} < \text{"John Lennon"} \wedge \text{media} > 7}$ (Aluno) haja um índice árvore-B+ primário para o atributo *nomeAluno*, com 10 entradas de índice por nó da árvore e um índice secundário denso ordenado para média:
 - a. Apresente quatro diferentes planos de execução de consulta para a expressão dada
 - b. Calcule o custo total de cada plano de consulta
 - c. Análise os resultados e explique qual é o melhor plano, justificando sua resposta.
- 12) Suponha a consulta $\sigma_{\text{nomeAluno} = \text{"Luis Fonsi"} \wedge \text{cidadeAluno} = \text{"San Juan"}}$ (Aluno). Considere que a tabela Musica está organizada por código da música e que existe um índice primário esparsa ordenado (30 entradas de índice por bloco) e outro índice secundário esparsa ordenado por intérprete (40 entradas de índice por bloco):
 - a. Apresente três diferentes planos de execução de consulta para a expressão dada
 - b. Calcule o custo total de cada plano de consulta, considerando $N_{\text{musica}} = 10.000$, $V(\text{nomeAluno}, \text{Aluno}) = 3000$; $V(\text{cidadeAluno}, \text{Aluno}) = 1000$.
 - c. Análise os resultados e explique qual é o melhor plano, justificando sua resposta.
- 13) Considerando os dados disponibilizados no final da lista e sabendo que cabem 21 blocos na memória da máquina, qual é o custo para classificar a Relação **Aluno** usando o algoritmo sort-merge-externo?

$F_{\text{aluno}} = 40$

$V(\text{nomeAluno}, \text{Aluno}) = 100.000$ $V(\text{cidadeAluno}, \text{Aluno}) = 500$

$N_{\text{aluno}} = 200.000$ (200 mil registros)

f_i – fan-out médio dos nós internos do índice i para índices estruturados em árvore

HT_i – número de níveis no índice i (altura do índice). Se árvore balanceada, $HT_i = \lceil \log_{f_i}(V(A, r)) \rceil$

Quando não for especificado, utilize $n=20$ para quantidade de ponteiros em um índice.