

Lista de Exercícios

MC536 – Bancos de Dados: Teoria e Prática Instituto de Computação Universidade Estadual de Campinas	Armazenamento, Indexação e Consulta 2013 André Santanchè
---	--

A Asdrúbal&Asdrúbal Engenharia executa projetos em todo o Brasil. Em seu sistema de controle de projetos estão duas tabelas:

Projeto(<u>IdProjeto</u> , Orçamento, Cidade) Despesa(<u>TipoDespesa</u> , <u>NrDocumento</u> , IdFornecedor, IdProjeto, Valor)
--

A tabela Projeto registra uma entrada para cada projeto:

Coluna	Descrição
IdProjeto	Identificador do projeto.
Orçamento	Orçamento total alocado para o projeto.
Cidade	Cidade em que está sendo executado o projeto.

A tabela Despesa registra cada despesa associada a cada projeto:

Coluna	Descrição
TipoDespesa	Tipo da despesa que pode ser: 'S' → serviço, 'M' → material de consumo, 'P' → peça, 'A' → aluguel de equipamento.
NrDocumento	Número do documento associado à despesa (e.g., nota fiscal).
IdFornecedor	Fornecedor do produto ou serviço. Chave estrangeira para a tabela de Fornecedor.
IdProjeto	Identifica o projeto ao qual está associada a despesa. Chave estrangeira para a tabela de Projeto.
Valor	Valor da despesa.

Questão 1

Considere três soluções de indexação ou ordenação da tabela de Despesa:

- índice na forma de árvore B+ organizado pelo IdProjeto;
- índice bitmap associado a IdProjeto;
- não há índice para IdProjeto, mas há uma rotina de ordenação externa para este campo.

Para os seguintes cenários alternativos apresentados a seguir recomende uma ou mais das soluções (a), (b) ou (c) e comente a diferença da sua aplicação em relação às demais soluções. Comente aspectos como consumo de espaço em disco (permanente ou temporário) e desempenho. Considere a existência daquele cenário como exclusivo, ou seja, a solução é apresentada para aquele cenário e ignora qualquer outro cenário.

I) Uma rotina controle de projeto precisa do total de despesas de um dado projeto a partir do seu IdProjeto. Esta rotina é disparada uma grande quantidade de vezes diariamente.

II) Ao final de cada mês é gerado um relatório apresentando o total de despesas por projeto.

III) Os dados da tabela Despesa são usados diariamente por rotinas analíticas que fazem cruzamentos de dados e respondem perguntas como: qual o fornecedor X que atende ao

projeto Y cujo valor da despesa está entre R\$ 5.000 e R\$ 10.000; qual o tipo da despesa é mais caro em cada projeto.

Questão 2

Dados os seguintes parâmetros:

- Os valores das despesas são organizados nas seguintes faixas:
 - baixa → despesa ≤ 1.000 ;
 - regular → $1.000 < \text{despesa} \leq 5.000$;
 - alta → $5.000 < \text{despesa} \leq 20.000$;
 - muito alta → despesa > 20.000 .
- A base foi projetada para ter no máximo 1.000 fornecedores e 250 projetos.
- O `idProjeto` é numérico e seu valor crescente - i.e., projetos novos tem `idProjeto` maior que projetos antigos.

Projete um índice bitmap para a tabela Despesa, cujo objetivo é responder de forma eficiente às seguintes consultas:

- a) Despesas do tipo X apresentam valores na faixa Y (baixa, regular, alta ou muito alta)?
- b) Há alguma despesa do tipo X, cujo valor seja muito alto e que tenha sido fornecida por Y?
- b) Projetos com `idProjeto` menor que X têm equipamentos alugados?

Defina quantos bits serão aplicados em cada parte do índice e para que finalidade. Justifique porque seu índice atenderá às consultas (a), (b) e (c) de forma eficiente.

Questão 3

Um sistema de monitoramento da bolsa de valores mantém um cadastro local que replica alguns dados das ações. Sabe-se que em geral aproximadamente 20% das ações têm valor acima da média. Como os dados destas ações é acessado com muita frequência pelo sistema, a empresa decidiu implantar um índice que otimize o acesso às ações acima da média. Indique que índices seriam adequados para este cenário e justifique, considerando os seguintes cenários:

- a) Há grande atualização dos dados das ações durante um mesmo dia.
- b) Os dados são atualizados uma vez por semana.