



RENAN HENRIQUE GOMES DAMAZIO ASSUNÇÃO RA 21038114
ALEX ARANTES GONÇALVES RA 21011214

RELATÓRIO 4

Tópicos Emergentes em Bancos de Dados

QUESTÃO 1

De acordo com as instruções contidas no script provido, uma relação foi criada com 60000 tuplas possuindo 4 atributos, o primeiro e o segundo do tipo inteiro de preenchimento obrigatório - *not null* - o terceiro também do tipo inteiro de preenchimento não obrigatório e quarto do tipo varchar de tamanho máximo de 300 caracteres. O fillfactor configurado para 95%

Foi informado que certas operações diárias críticas de modificação são lentas, que a expectativa de preenchimento do 3º atributo era de 20% e que não haviam informações sobre o atributo 4.

Ao considerar as informações apresentadas, foi verificada a quantidade de linhas preenchidas do att3, o que resultou em uma quantidade de mais de 27000 tuplas preenchidas com este atributo, 46% de preenchimento ou seja o dobro do que se esperava.

O comprimento dos dados registrados no atributo 4 foi analisada. O comprimento máximo é de 64 bytes, que corresponde a maioria das tuplas conforme pode ser verificado no Gráfico 1

Proporção de tuplas por Comprimento do att4

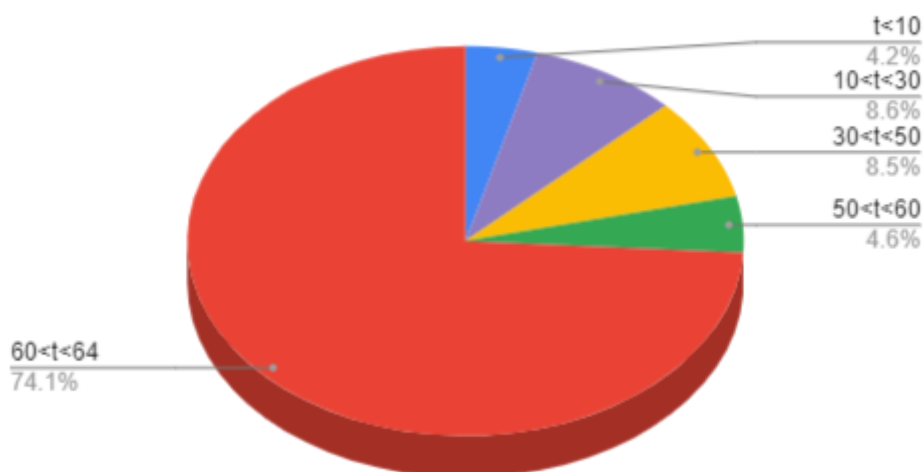


Gráfico 1 - Proporção de tuplas por comprimento do atributo 4

Considerando o att4 como 64 bytes, e porcentagem de preenchimento do atributo 3 como 46%, é possível estimar a média do fator de blocagem atual:

$$FB_1 = (8192 \times 0.95) / (4(\text{att1}) + 4(\text{att2}) + 4(\text{att3}) + 64(\text{att4})) = 102 \text{ tuplas/ bloco}$$

$$FB_2 = (8192 \times 0.95) / (4(\text{att1}) + 4(\text{att2}) + 64(\text{att4}) + 8(\text{bitmap})) = 97 \text{ tuplas/ bloco}$$

$$FB_{\text{média}} = (25.9 \times 0,46) + (24.6 \times 0,54) = \mathbf{99 \text{ Tuplas/ Bloco}}$$

Como as operações críticas diárias são de modificações, propomos a redução do fillfactor para 80%, uma vez que não há informações sobre expectativas de preenchimento do atributo 4. Assim sendo a nova estimativa do fator de blocagem seria:

$$FB_1 = (8192 \times 0.80) / (4 \text{ (att1)} + 4 \text{ (att2)} + 4 \text{ (att3)} + 64 \text{ (att4)}) = 86 \text{ tuplas/ bloco}$$

$$FB_2 = (8192 \times 0.80) / (4 \text{ (att1)} + 4 \text{ (att2)} + 64 \text{ (att4)} + 8 \text{ (bitmap)}) = 82 \text{ tuplas/ bloco}$$

$$FB_{\text{média}} = (25.9 \times 0.46) + (24.6 \times 0.54) = \mathbf{84 \text{ Tuplas/ Bloco}}$$

o que diminuiria o fator de blocagem em 16% , mas disponibilizaria mais espaço para expansão dos dados no atributo 4 sem que o sistema crie novas páginas permitindo aumento da eficiência na alteração de dados.

QUESTÃO 2

Para análise desta questão todos os comandos SQL utilizados constam no arquivo Q2.txt.

Baseado nos parâmetros propostos nesta questão, primeiramente foi determinado o tamanho da soma dos atributos de tamanho fixo discriminados na tabela 1 Por conta da estabilidade e pouca variação (95%) dos atributos snemp e nemp, os mesmos foram contabilizados de maneira fixa também.

Nome	Tipo	Tamanho
coddepto	int	4
codempresponsavel	int	4
codemp	int	4
snemp	varchar(30) (95%)	29
nemp	varchar(40) (95%)	38
codempgerente	int	4
dtnast	date	4
codesex	char	1
vlsalario	real	4
Total		92

Tabela 1 -

fixos

tamanho atributos

Dos atributos variáveis, como o emp.ender possui maior tamanho e uma variação de expectativas de 30 a 80% de preenchimento, foi utilizado no cálculo do fator de blocagem.

➤ Relação (92 atributos fixos + 255 variáveis)

- Expectativa #1 30% de preenchimento
- Expectativa #1 80% de preenchimento
- Pressuposto ($92 + 255 \times 30\%$) = 169
- Expectativa ($92 + 255 \times 80\%$) = 296

Cálculos:

- $(8192 / 169 + 3) = 47.6$ tuplas por página
- $(8192 / 296 + 3) = 27.3$ tuplas por página
- $27.3 \times (169 + 3) = 4712.45$ espaço
- $8192 - 4712 = 3480$
- $3480 / 8192 = 42\%$

Como resultado o fillfactor foi configurado para 58%.

Embora como resultado desta configuração o fator de blocagem seja menor, esta forma se garante que operações de modificação do banco não comprometam o desempenho.

Nenhum index foi utilizado para melhorar o desempenho da consulta proposta na questão pois o fullscan é necessário para contabilização e agrupamento de número de empregados por departamento. A indexação neste caso apenas contribuiria para que o desempenho nas operações de escrita fossem comprometidas.