



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

RELATÓRIO 3

Tópicos Emergentes em Bancos de Dados

Santo André – SP

2019

QUESTÃO 1

Uma relação de 600000 tuplas foi criada contendo três atributos, att1, att2 e att3. Os dois primeiros atributos contendo dados tipo *int* sequenciais e não nulos, e o att3 contendo valores também do tipo *int* aleatórios. Seis diferentes Queries foram analisadas neste experimento, cujos códigos SQL seguem na tabela 1:

```
P1) SELECT * FROM rel WHERE att1=100003 and att2 = 100003
P2) SELECT * FROM Rel where att1 in (5, 1000, 2058, 3285) and att2 between 1 and 5000
P3) SELECT * FROM Rel where att1 = 253879;
P4) SELECT * FROM Rel where att1 = 63 or att1 = 405282;
P5) SELECT * FROM Rel where att2 = 2019;
P6) SELECT * FROM Rel where att2 = 314165 or att2 = 2;
```

Tabela 1 - Consultas utilizadas no experimento

O experimento consistiu em realizar as consultas acima em diferentes cenários se utilizando de diferentes casos de indexação:

- 1º caso sem indexação;
- 2º caso com indexação composta, B+ Tree nos atributos att1 e att2;
- 3º caso com indexação simples dupla nos atributos att1 e att2;
- 4º caso com indexação composta, B+ Tree nos atributos att1 e att2 e indexação BRIN no atributo att2;

Para este experimento, o processamento paralelo foi desativado, bem como Bitmap Hash. O script utilizado nos diversos casos se encontra no arquivo *Q1.txt*.

Resultados:

Os tempos de execução foram registrados na tabela 2, que também contém informações de qual modo de pesquisa o otimizador se utilizou:

	Tempos de execução (ms)				
	Sem Indexação	Idx. Composta	Idx. Simples dupla	Idx Composta/idx BRIN	
P1	43	0.1	0.1	0.1	<div>Seq Scan utilizada</div> <div>indexação composta utilizada</div> <div>indexação B tree no atributo 2 utilizada</div> <div>indexação B tree no atributo 1 utilizada</div>
P2	95	0.1	0.1	0.1	
P3	37	0.1	0.1	0.1	
P4	45	55	0.1	49	
P5	69	67	0.1	72	
P6	92	92	93	94	

Tabela 2 - Tempos de execução de cada query nos 4 casos realizados

Os tempos de execução dos dados da tabela 2 por diferentes Queries e casos foram plotados em um gráfico - Gráfico 1- para melhor visualização dos resultados obtidos:

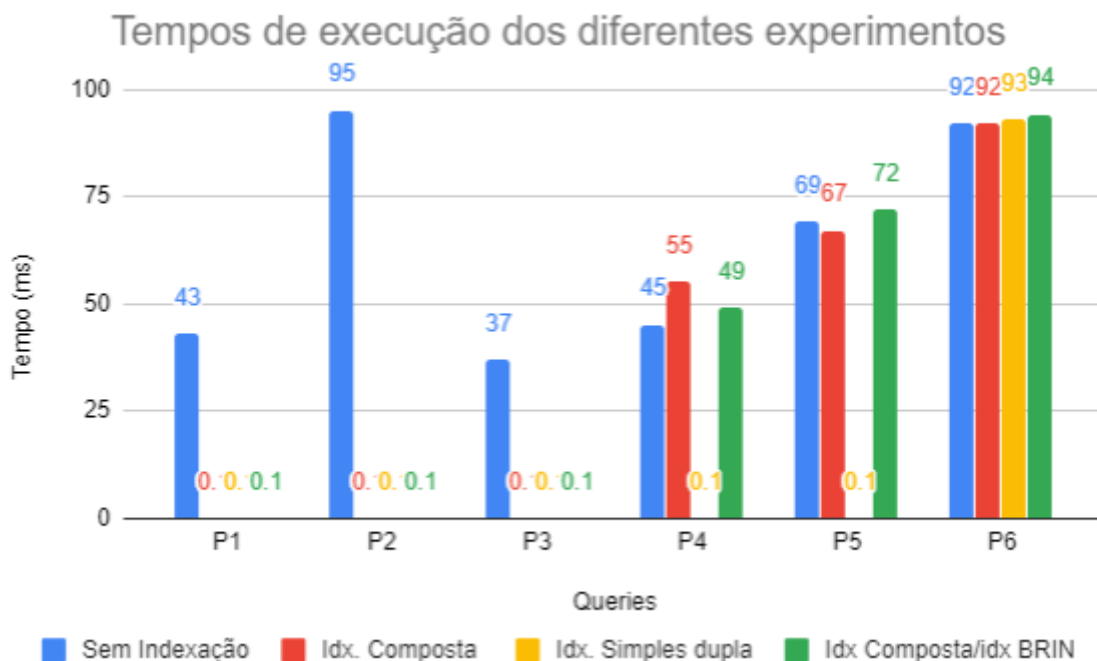


Gráfico 1 - Tempos de execução por experimento

Conclusões:

- I. A indexação dupla simples se mostrou a mais favorável para a maior parte das queries;
- II. O Otimizador não utilizou indexação BRIN em nenhuma das consultas para o cenário 4 em que foi configurado.

QUESTÃO 2

Para todos os resultados o processamento paralelo foi desligado. Todos os comandos SQL utilizados constam no arquivo *Q2.txt*.

Foi identificado que além das PK's o único índice existente era o *idx_depto* sobre o campo *coddepto* da tabela *depto*.

T1: Utilizou os índices *pk_emp* e *idx_depto*, retorno em tempo inferior a 1ms, por isso motivo não foi alterada;

T2: Foram criadas 3 versões para a query com faixas de valores distintas para o *between*, dentro dos vales *mix* e *max* presentes na tabela. A execução inicial não utilizou nenhum índice, para nenhuma das três versões. Por se tratar de uma faixa de valores de quantidade de horas cogitou-se a utilização do índice BRIN. O índice foi

criado, porém, não foi utilizado em nenhuma das 3 queries utilizadas. Decidiu-se então por não utilizá-lo no projeto;

T3: Assim como em T2 foram criadas 3 faixas, as quais utilizaram o índice *pk_emp* na execução inicial. Foi feita análise do índice BRIN para o campo *vlsalario*, porém, também não foi utilizado e não será adotado para o projeto;

T4: Como T4 utiliza o campo *nmlocalização* que é um campo de texto não cogitou-se a utilização do índice BRIN., visto que esse tem melhores resultados em faixas de valores. A execução inicial desta transação não utilizou nenhum dos índices pré-existentes.

T5: Das transações avaliadas, T5 possui o maior tempo de execução, não utilizando nenhum índice na execução inicial. Foi testado o uso do índice BRIN para faixa de salário do campo *vlsalario*, exatamente igual ao índice criado para T2. O índice criado não foi utilizado na execução. Por esse motivo decidimos não utilizá-lo no projeto.

Conclusões

- I. Apesar de ser utilizado especialmente para faixas de valores, a criação do índice BRIN não teve nenhum efeito na execução de nenhuma das transações, por esse motivo não foi utilizado no projeto;
- II. Em contrapartida, foi feita a tentativa de criação de índice btree para algumas das transações, o qual apresentou resultados positivos, porém, como não é o foco da análise não será relatado com profundidade.