

Análise de Otimização de Consultas

Adriel Schmitz, Leonardo José Gonçalves

17 de Maio de 2018, Chapecó-SC

- SGBD



- Banco Usado

→ STATS

Tipos de dados: Numérico,
String e de Tempo.

Tamanho: 658.4 MB

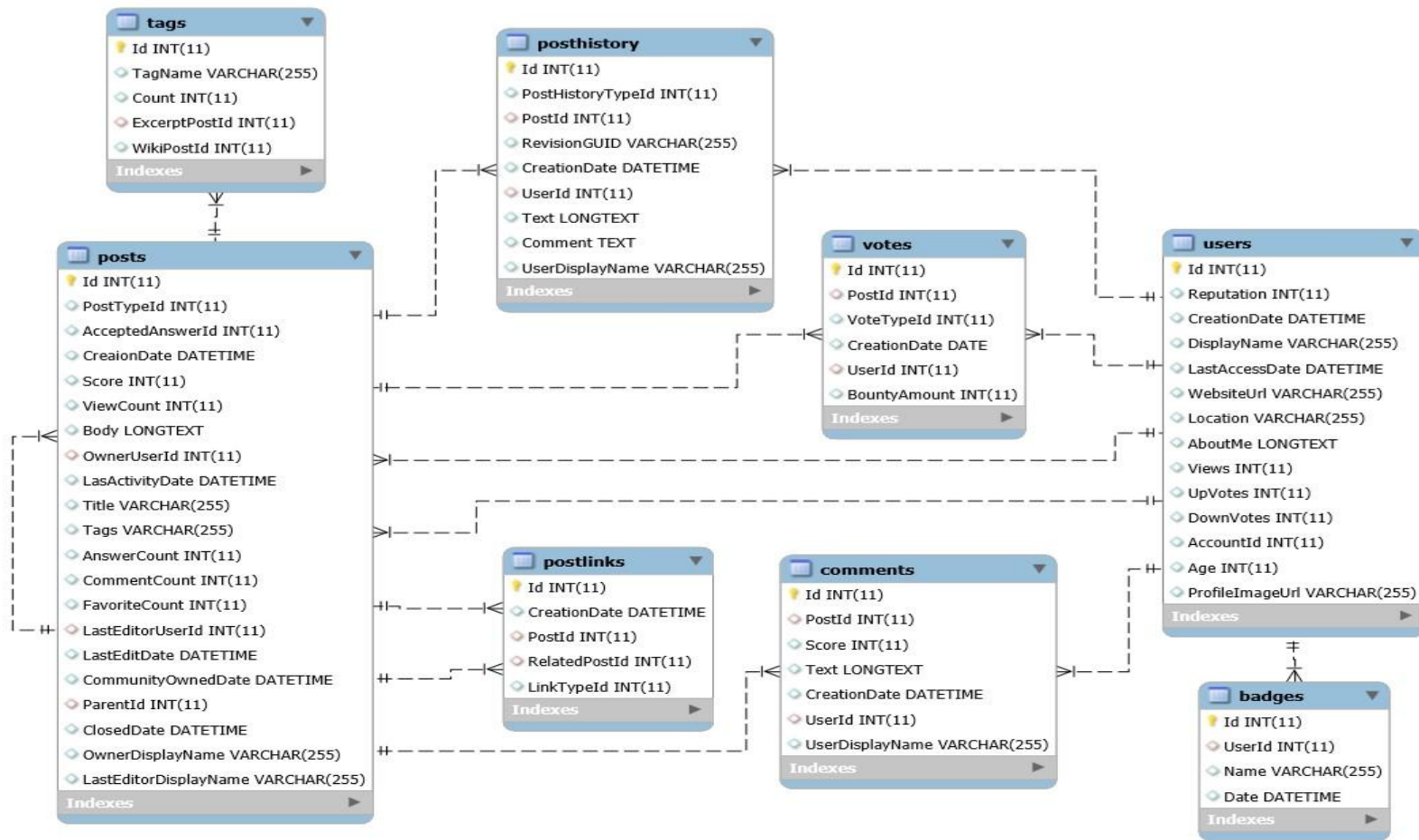
Linhas: 1,027,838

Tabelas: 8

Colunas: 71

Disponível: <<https://relational.fit.cvut.cz/dataset/Stats>>

● Banco Usado



● Consulta 1



```
SELECT u.displayname, posts.title from posts JOIN votes ON  
posts.id = votes.postid JOIN (SELECT users.id, users.DisplayName  
from users WHERE users.DisplayName = 'Preet') as U ON U.id =  
votes.userid;
```

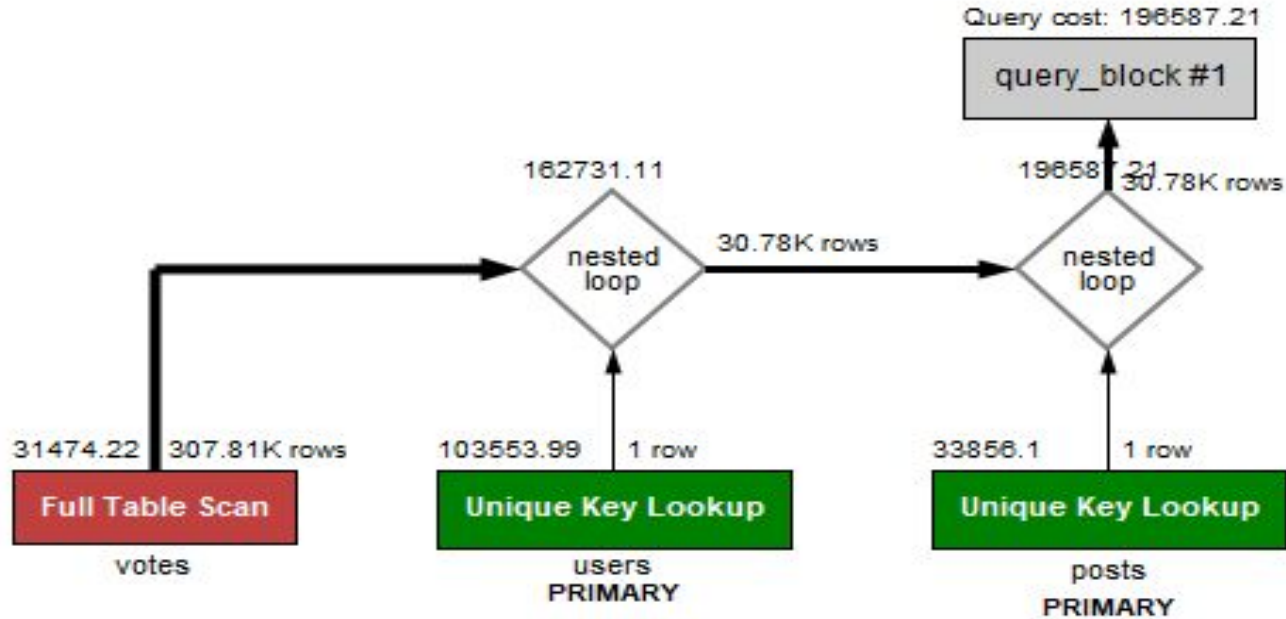
Basicamente, seleciona todos os *posts* que o *usuário* 'Preet' *votou*.

Resultado:

DisplayName	title
Preet	The Two Cultures: statistics vs. machine learning?
Preet	Sites for predictive modeling competitions
Preet	Python as a statistics workbench

- Consulta 1

Plano:





● Consulta 1

Tempo:

Timing (as measured at client side):

Execution time: 0:00:2.51600000

Timing (as measured by the server):

Execution time: 0:00:2.52588917

Table lock wait time: 0:00:0.00049200

Errors:

Had Errors: NO

Warnings: 0

Rows Processed:

Rows affected: 0

Rows sent to client: 3

Rows examined: 362856

Temporary Tables:

Temporary disk tables created: 0

Temporary tables created: 0

Joins per Type:

Full table scans (Select_scan): 1

Joins using table scans (Select_full_join): 0

Joins using range search (Select_full_range_join): 0

Joins with range checks (Select_range_check): 0

Joins using range (Select_range): 0

Sorting:

Sorted rows (Sort_rows): 0

Sort merge passes (Sort_merge_passes): 0

Sorts with ranges (Sort_range): 0

Sorts with table scans (Sort_scan): 0

Index Usage:

No Index used

Other Info:

Event Id: 133

Thread Id: 70

- **Consulta 1**

```
EXPLAIN ANALYZE SELECT u.displayname, posts.title from posts JOIN votes  
ON posts.id = votes.postid JOIN (SELECT users.id, users.DisplayName from  
users WHERE users.DisplayName = 'Preet') as U ON U.id = votes.userid;
```

Resultado:

```
displayname | title  
-----+-----  
Preet       | The Two Cultures: statistics vs. machine learning?  
Preet       | Sites for predictive modeling competitions  
Preet       | Python as a statistics workbench  
(3 rows)
```


● Consulta 1

Tempos, Planos:

```
QUERY PLAN
-----
Nested Loop (cost=0.71..1583.57 rows=1 width=67) (actual time=0.093..25.823 rows=3 loops=1)
-> Nested Loop (cost=0.42..1583.11 rows=1 width=17) (actual time=0.080..25.780 rows=3 loops=1)
    -> Seq Scan on users (cost=0.00..1526.06 rows=1 width=17) (actual time=0.056..25.747 rows=1 loops=1)
        Filter: ((displayname)::text = 'Preet's::text)
        Rows Removed by Filter: 40324
    -> Index Scan using idx_24935_votes_stripped_userid_fkey on votes (cost=0.42..56.80 rows=24 width=16) (actual time=0.018..0.023 rows=3 loops=1)
        Index Cond: (userid = users.id)
-> Index Scan using idx_24911_posts_id_key on posts (cost=0.29..0.45 rows=1 width=66) (actual time=0.009..0.010 rows=1 loops=3)
    Index Cond: (id = votes.postid)
Planning time: 1.835 ms
Execution time: 26.022 ms
(11 rows)
```

- **Comparação**

- MySQL

Execução: 2.516 ms

Planejamento: --

Total: 2.516 ms

- PostgreSQL

Execução: 26.022 ms

Planejamento: 1.185 ms

Total: 27.207 ms



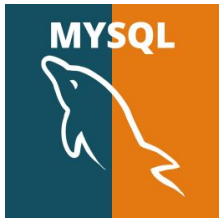
● Consulta 2

```
SELECT users.DisplayName, COUNT(comments.id) FROM users  
JOIN comments ON users.id = comments.UserId GROUP  
BY(comments.id);
```

Quantidade de *comentários* que cada *usuário* postou.

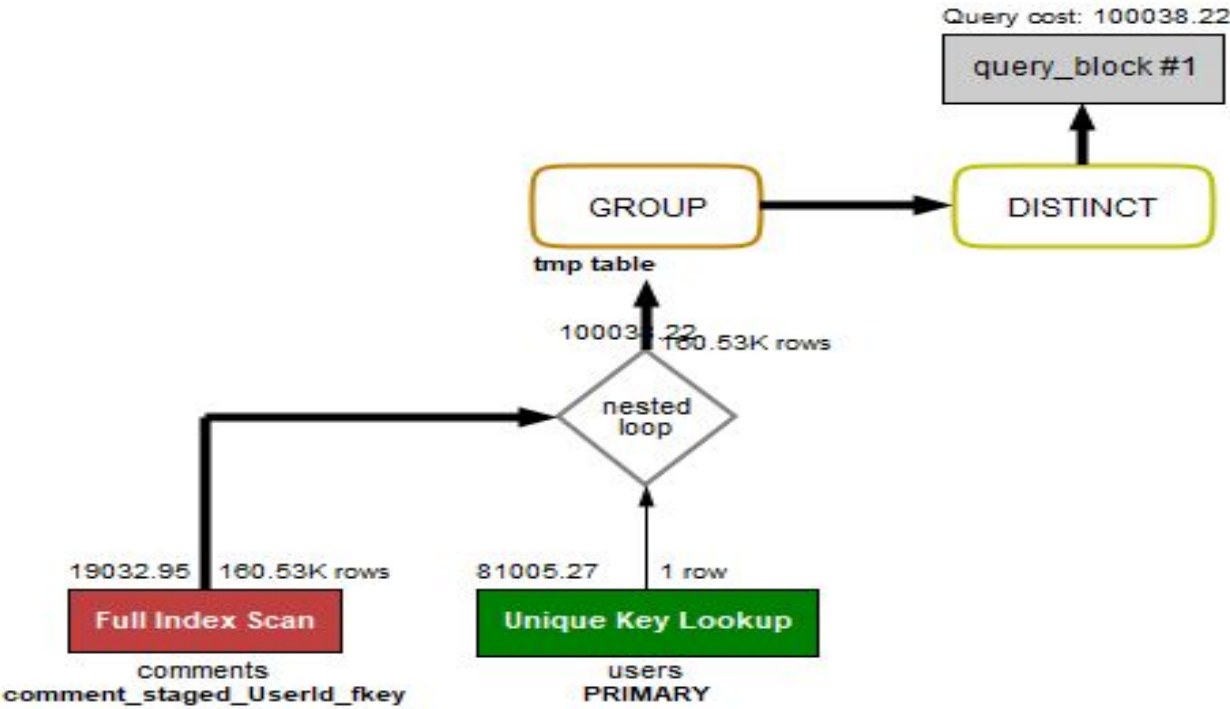
Resultado:

DisplayName	COUNT(comments.id)
Jarrood Dixon	1
Shane	1
Harlan	1
Vince	1
csaillespie	1
wahalulu	1
Sharpie	1
arokus	1
Noah Snyder	1
radek	1



- **Consulta 2**

Plano:





● Consulta 2

Tempo:

Timing (as measured at client side):

Execution time: 0:00:3.37500000

Timing (as measured by the server):

Execution time: 0:00:3.37025991

Table lock wait time: 0:00:0.00017100

Errors:

Had Errors: NO

Warnings: 0

Rows Processed:

Rows affected: 0

Rows sent to client: 10

Rows examined: 345785

Temporary Tables:

Temporary disk tables created: 0

Temporary tables created: 1

Joins per Type:

Full table scans (Select_scan): 1

Joins using table scans (Select_full_join): 0

Joins using range search (Select_full_range_join): 0

Joins with range checks (Select_range_check): 0

Joins using range (Select_range): 0

Sorting:

Sorted rows (Sort_rows): 0

Sort merge passes (Sort_merge_passes): 0

Sorts with ranges (Sort_range): 0

Sorts with table scans (Sort_scan): 0

Index Usage:

At least one Index was used

Other Info:

Event Id: 155

Thread Id: 70

- **Consulta 2**

```
EXPLAIN ANALYZE SELECT users.DisplayName, COUNT(comments.id)
FROM users JOIN comments ON users.id = comments.UserId GROUP
BY(comments.id);
```

- **Consulta 2**

R.I.P :((

```
bd2=# users.DisplayName, COUNT(comments.id) FROM users JOIN comments ON users.id
    = comments.UserId GROUP BY(comments.id);
ERROR:  syntax error at or near "users"
LINE 1: users.DisplayName, COUNT(comments.id) FROM users JOIN commen...
          ^
bd2=# |
```

- **Comparação**

➤ MySQL

Execução: 3.375 ms

Planejamento: --

Total: 3.375 ms

➤ PostgreSQL

Execução: { }

Planejamento:{ }

Total: { }



- **Consulta 3 sem índice**

```
SELECT users.displayname, COUNT(comments.id) FROM users  
JOIN comments ON users.id = comments.userid WHERE  
users.displayname = 'Preetz' GROUP BY(users.id);
```

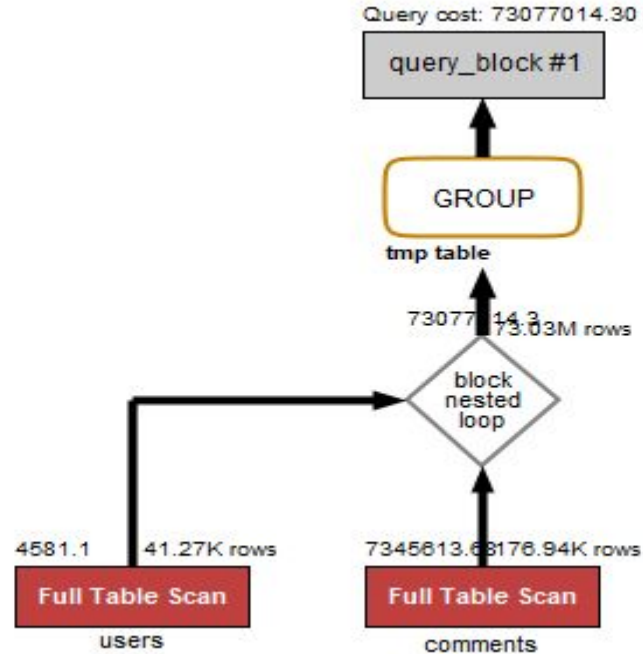
Quantidade de *comentários* que o *usuário* 'Preetz' postou.

Resultado:

	displayname	COUNT(comments.id)
	Preetz	2

- **Consulta 3 sem índice**

Plano:





- **Consulta 3 sem índice**

Tempo:

Timing (as measured at client side):

Execution time: 0:00:1.84400000

Timing (as measured by the server):

Execution time: 0:00:1.83921509

Table lock wait time: 0:00:0.00020000

Errors:

Had Errors: NO

Warnings: 0

Rows Processed:

Rows affected: 0

Rows sent to client: 1

Rows examined: 214631

Temporary Tables:

Temporary disk tables created: 0

Temporary tables created: 1

Joins per Type:

Full table scans (Select_scan): 1

Joins using table scans (Select_full_join): 1

Joins using range search (Select_full_range_join): 0

Joins with range checks (Select_range_check): 0

Joins using range (Select_range): 0

Sorting:

Sorted rows (Sort_rows): 0

Sort merge passes (Sort_merge_passes): 0

Sorts with ranges (Sort_range): 0

Sorts with table scans (Sort_scan): 0

Index Usage:

No Index used

Other Info:

Event Id: 43

Thread Id: 67

- **Consulta 3 sem índice**

```
EXPLAIN ANALYZE SELECT users.displayname, COUNT(comments.id)
FROM users JOIN comments ON users.id = comments.userid WHERE
users.displayname = 'Preet's' GROUP BY(users.id);
```

Resultado:

```
displayname | count
-----+-----
Preet's     |      2
(1 row)
```

- **Consulta 3 sem índice**

Tempos e Planos:

```
QUERY PLAN
-----
HashAggregate (cost=10464.83..10464.84 rows=1 width=25) (actual time=53.917..53.917 rows=1 loops=1)
  Group Key: users.id
    -> Hash Join (cost=1526.08..10464.81 rows=4 width=25) (actual time=7.837..53.907 rows=2 loops=1)
      Hash Cond: (comments.userid = users.id)
        -> Seq Scan on comments (cost=0.00..8285.05 rows=174305 width=16) (actual time=0.071..26.204 rows=174305 loops=1)
        -> Hash (cost=1526.06..1526.06 rows=1 width=17) (actual time=7.710..7.710 rows=1 loops=1)
          Buckets: 1024 Batches: 1 Memory Usage: 9kB
          -> Seq Scan on users (cost=0.00..1526.06 rows=1 width=17) (actual time=0.083..7.700 rows=1 loops=1)
            Filter: ((displayname)::text = 'Preet's '::text)
            Rows Removed by Filter: 40324
Planning time: 1.047 ms
Execution time: 54.073 ms
(12 rows)
```

- **Comparação**

- MySQL

Execução: 1.844 ms

Planejamento: --

Total: 1.844 ms

- PostgreSQL

Execução: 54.073 ms

Planejamento: 1.047 ms

Total: 55,012 ms