

# Análise de desempenho para o banco de dados MariaDB

---

Cristiano Daitx Ribeiro  
Igor Antunes Fogaça  
Marcos Vinícius Treviso

17 de julho de 2015

Banco de Dados II - Universidade Federal do Pampa

- Introdução
  - MariaDB
  - Índices
- Ambiente
- Resultados
- Avaliação
- Conclusão

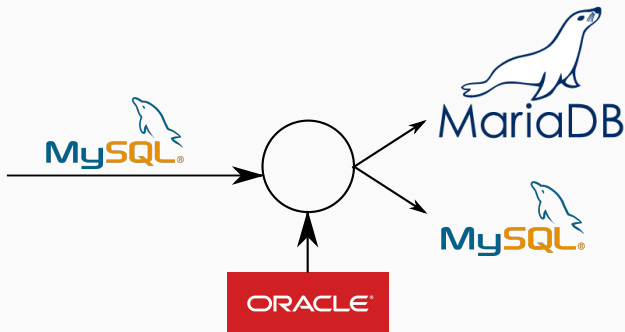
- Introdução

MariaDB

Índices

- Ambiente
- Resultados
- Avaliação
- Conclusão

- Sistema de Gerenciamento de Banco de Dados Relacional
  - Funcionalidade e substituição para o MySQL
  - Autores originais do MySQL
  - Aprimoramentos de recursos, otimizações
  - Criado pelo próprio criador do MySQL: Michael Widenius



- Árvore B
  - $>$ ,  $>=$ ,  $=$ ,  $<=$ ,  $<$ ,  $\neq$ , BETWEEN
  - LIKE (apenas com prefixo)
- Hash
  - $>$ ,  $>=$ ,  $=$ ,  $<=$ ,  $<$ ,  $\neq$
  - LIKE sem substrings
- Árvore R
  - $>$ ,  $>=$ ,  $=$ ,  $<=$ ,  $<$ ,  $\neq$ , BETWEEN
  - LIKE
  - Apenas campos NOT NULL

- Introdução

MariaDB

Índices

- Ambiente
- Resultados
- Avaliação
- Conclusão

- Notebook HP Pavilion dv6

---

<b>Processador</b>	Intel i7-2670QM 2.20GHz
<b>Memória Principal</b>	8 GB RAM DDR3 1333 MHz
<b>Disco</b>	TOSHIBA MK1059GS 1 TB
<b>Cache</b>	L3 Cache 6 MB

---

- SGBD: MariaDB versão 10.0.20
- Engine de armazenamento: InnoDB

# Ambiente

- Banco de dados
  - 5 tabelas com tamanho:  $10^2, 10^3, 10^4, 10^5, 10^6$
  - Índice primário sobre campo *id*
  - Índice secundário sobre campo *país*
  - Índice secundário composto sobre campos *nome, sobrenome*
  - Índices utilizando árvore B

<u>idrobos</u> (int)	nome (int)	sobrenome (int)	país (int)	senha (int)	idade (int)
$[0,  T ]$	$[0, 10^3]$	$[0, 10^3]$	$[0, 250]$	$[0, 10^6]$	$[0, 10^2]$

- Média de 15 tempos de execução
- Eliminação da cache:  
`SET GLOBAL query_cache_size, query_cache_type = 0, 0;`
- Atribuição aleatória para população das tabelas



- Introdução
  - MariaDB
  - Índices
- Ambiente
- Resultados
- Avaliação
- Conclusão

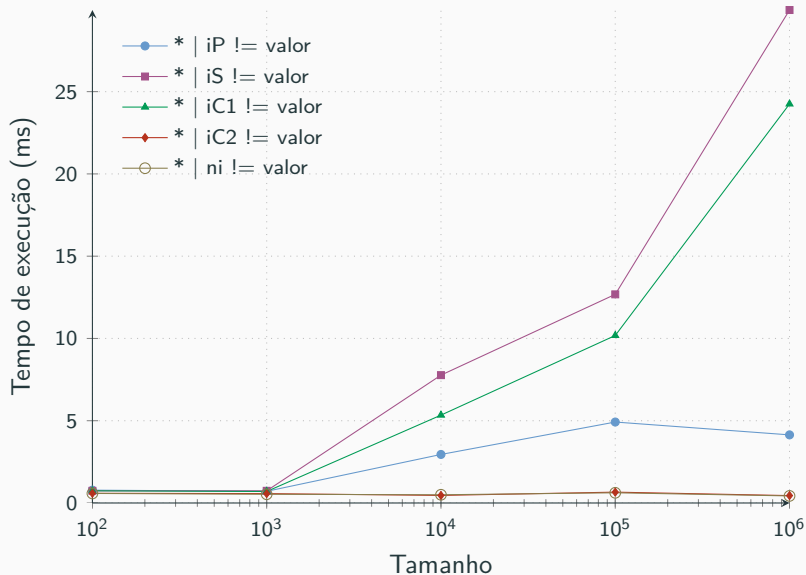
# Resultados

- Conjunto f) coleção 1:
- **Consulta 1:** `SELECT * FROM tabela WHERE iP != valor`
- **Consulta 2:** `SELECT * FROM tabela WHERE iS != valor`
- **Consulta 3:** `SELECT * FROM tabela WHERE iC1 != valor`
- **Consulta 4:** `SELECT * FROM tabela WHERE iC2 != valor`
- **Consulta 5:** `SELECT * FROM tabela WHERE ni != valor`

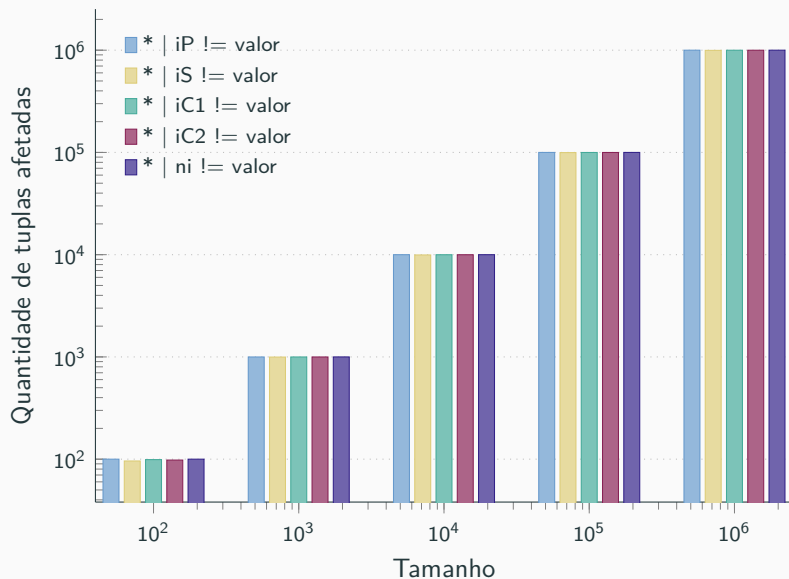
Consulta	100	1000	10000	100000	1000000
*   iP != valor	0.78 (100)	0.71 (999)	2.95 (9998)	4.92 (99997)	4.14 (999998)
*   iS != valor	0.74 (96)	0.74 (997)	7.77 (9940)	12.68 (99631)	29.96 (995929)
*   iC1 != valor	0.74 (99)	0.70 (999)	5.34 (9989)	10.18 (99894)	24.25 (999017)
*   iC2 != valor	0.59 (98)	0.57 (999)	0.46 (9989)	0.66 (99897)	0.45 (999012)
*   ni != valor	0.58 (100)	0.53 (1000)	0.50 (9999)	0.62 (99999)	0.43 (999997)

Tempo de execução (ms) - Quantidade de tuplas afetadas

# Resultados



# Resultados



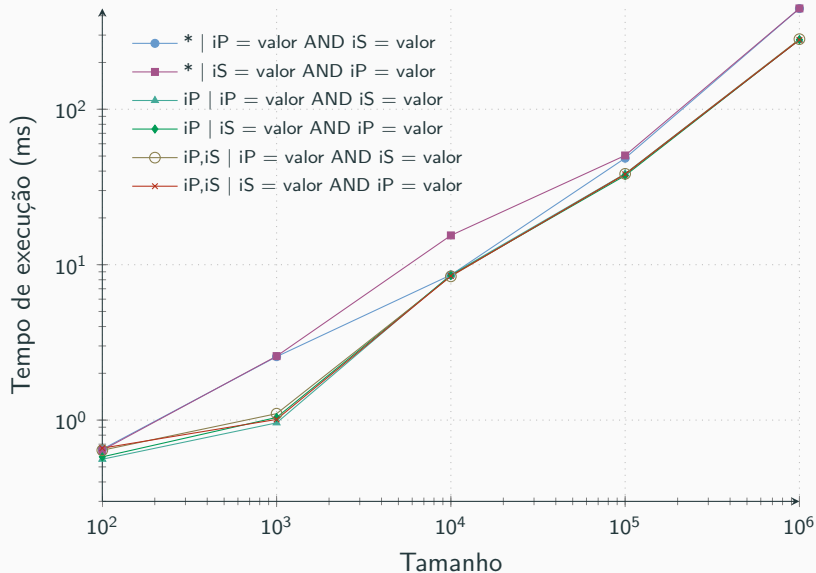
# Resultados

- Conjunto f) coleção 2:
- **Consulta 1:** `SELECT * FROM tabela WHERE iP = valor AND iS = valor`
- **Consulta 2:** `SELECT * FROM tabela WHERE iS = valor AND iP = valor`
- **Consulta 3:** `SELECT iP FROM tabela WHERE iP = valor AND iS = valor`
- **Consulta 4:** `SELECT iP FROM tabela WHERE iS = valor AND iP = valor`
- **Consulta 5:** `SELECT iP,iS FROM tabela WHERE iP = valor AND iS = valor`
- **Consulta 6:** `SELECT iP,iS FROM tabela WHERE iS = valor AND iP = valor`

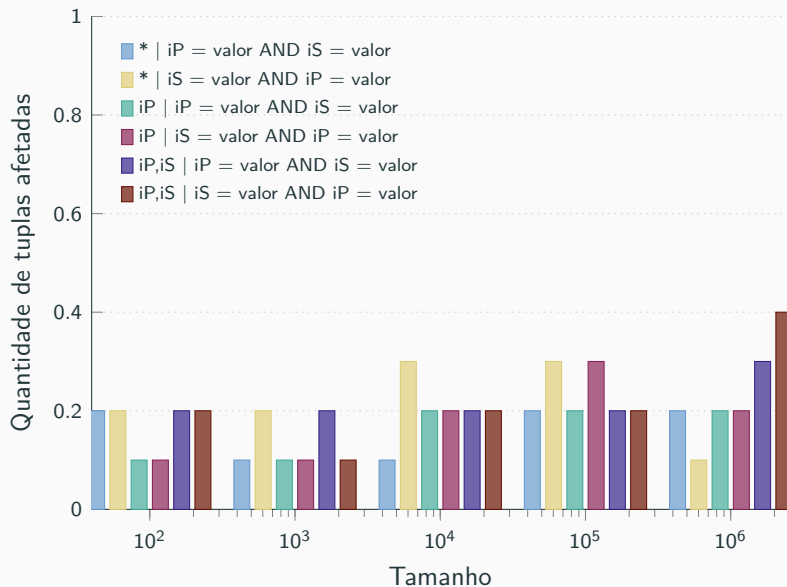
Consulta	100	1000	10000	100000	1000000
Consulta 1	0.65 (0.2)	2.56 (0.1)	8.59 (0.1)	48.46 (0.2)	444.57 (0.2)
Consulta 2	0.64 (0.2)	2.58 (0.2)	15.43 (0.3)	50.44 (0.3)	445.45 (0.1)
Consulta 3	0.56 (0.1)	0.96 (0.1)	8.62 (0.2)	38.47 (0.2)	279.38 (0.2)
Consulta 4	0.58 (0.1)	1.04 (0.1)	8.52 (0.2)	47.44 (0.3)	280.41 (0.2)
Consulta 5	0.64 (0.2)	1.10 (0.2)	8.41 (0.2)	37.43 (0.2)	282.51 (0.3)
Consulta 6	0.66 (0.2)	1.01 (0.1)	8.46 (0.2)	38.16 (0.2)	279.42 (0.4)

Tempo de execução (ms) - Quantidade de tuplas afetadas

# Resultados



# Resultados



- Introdução
  - MariaDB
  - Índices
- Ambiente
- Resultados
- Avaliação
- Conclusão



- **Consulta 1:** `SELECT * FROM tabela WHERE iP != valor`
  - Índice primário
  - Operação de desigualdade
  - Todos os dados da tabela

Tamanho	Tipo de Acesso	Chaves Possíveis	Chave Usada	Linhas	Extra
$10^2$	variável	primária	primária	99	where
$10^3$	variável	primária	primária	997	where
$10^4$	variável	primária	primária	5053	where
$10^5$	variável	primária	primária	47096	where
$10^6$	variável	primária	primária	479274	where

- **Consulta 2:** `SELECT * FROM tabela WHERE is != valor`
  - Índice secundário
  - Operação de desigualdade
  - Todos os dados da tabela

Tamanho	Tipo de Acesso	Chaves Possíveis	Chave Usada	Linhas	Extra
$10^2$	linear	secundária	nenhuma	100	where
$10^3$	linear	secundária	nenhuma	1000	where
$10^4$	linear	secundária	nenhuma	10104	where
$10^5$	linear	secundária	nenhuma	93912	where
$10^6$	linear	secundária	nenhuma	958268	where

- **Consulta 3:** `SELECT * FROM tabela WHERE iC1 != valor`
  - Primeiro elemento do índice composto
  - Operação de desigualdade
  - Todos os dados da tabela

Tamanho	Tipo de Acesso	Chaves Possíveis	Chave Usada	Linhas	Extra
$10^2$	linear	composta <sup>1</sup>	nenhuma	100	where
$10^3$	linear	composta <sup>1</sup>	nenhuma	1000	where
$10^4$	linear	composta <sup>1</sup>	nenhuma	10104	where
$10^5$	linear	composta <sup>1</sup>	nenhuma	93912	where
$10^6$	linear	composta <sup>1</sup>	nenhuma	958268	where

- **Consulta 4:** `SELECT * FROM tabela WHERE ic2 != valor`
  - Segundo elemento do índice composto
  - Operação de desigualdade
  - Todos os dados da tabela

Tamanho	Tipo de Acesso	Chaves Possíveis	Chave Usada	Linhas	Extra
$10^2$	linear	nenhuma	nenhuma	100	where
$10^3$	linear	nenhuma	nenhuma	1000	where
$10^4$	linear	nenhuma	nenhuma	10104	where
$10^5$	linear	nenhuma	nenhuma	93912	where
$10^6$	linear	nenhuma	nenhuma	958268	where

- **Consulta 5:** `SELECT * FROM tabela WHERE ni != valor`
  - Sem índice
  - Operação de desigualdade
  - Todos os dados da tabela

Tamanho	Tipo de Acesso	Chaves Possíveis	Chave Usada	Linhas	Extra
$10^2$	linear	nenhuma	nenhuma	100	where
$10^3$	linear	nenhuma	nenhuma	1000	where
$10^4$	linear	nenhuma	nenhuma	10104	where
$10^5$	linear	nenhuma	nenhuma	93912	where
$10^6$	linear	nenhuma	nenhuma	958268	where

- Foi usado índice apenas na consulta sobre a chave primária
- Por que linear?
  - Operação de desigualdade
  - Distribuição uniforme e aleatória
- Por que nenhuma chave possível para a Consulta 4 (segundo elemento do índice composto)?
  - MariaDB ordena pelo primeiro elemento, e depois pelo segundo
  - `return (a.first < b.first) ? true : (a.second < b.second)`
- Por que as linhas previstas não ficaram parecidas com o encontrado?
  - Estimativa de custos
  - Atribuição pseudoaleatória

$$n_r - \frac{n_r}{V(A, r)}$$

- **Tempo de execução:**
- Consulta 1 (chave primária)
  - Utilizou o índice primário → Tempo de execução baixo
- Consulta 2 (chave secundária)
  - Não utilizou o índice secundário → Tempo de execução alto conforme o tamanho da tabela cresce
- Consulta 3 (primeiro elemento da chave composta)
  - Não utilizou o índice composto
  - Crescimento semelhante a Consulta 2
- Consulta 4 (segundo elemento da chave composta)
  - Não utilizou o índice composto
  - Busca linear
- Consulta 5 (primeiro elemento da chave composta)
  - Busca linear
  - Comportamento muito semelhante a Consulta 4

- **Consulta 1:** `SELECT * FROM tabela WHERE iP = valor AND iS = valor`
  - Índice primário e secundário
  - Operação de AND sobre igualdade
  - Todos os dados da tabela

Tamanho	Tipo de Acesso	Chaves Possíveis	Chave Usada	Linhas	Extra
$10^2$	NULL	nenhuma	nenhuma	NULL	impossível
$10^3$	NULL	nenhuma	nenhuma	NULL	impossível
$10^4$	NULL	nenhuma	nenhuma	NULL	impossível
$10^5$	NULL	nenhuma	nenhuma	NULL	impossível
$10^6$	NULL	nenhuma	nenhuma	NULL	impossível

- Impossível → Só há uma chance de ocorrer
- Caso ocorra, é utilizado o índice primário



- **Consulta 2:** `SELECT * FROM tabela WHERE iS = valor AND iP = valor`
  - Índice primário e secundário
  - Operação de AND sobre igualdade
  - Todos os dados da tabela

Tamanho	Tipo de Acesso	Chaves Possíveis	Chave Usada	Linhas	Extra
$10^2$	NULL	nenhuma	nenhuma	NULL	impossível
$10^3$	NULL	nenhuma	nenhuma	NULL	impossível
$10^4$	NULL	nenhuma	nenhuma	NULL	impossível
$10^5$	NULL	nenhuma	nenhuma	NULL	impossível
$10^6$	NULL	nenhuma	nenhuma	NULL	impossível

- Impossível → Só há uma chance de ocorrer
- Caso ocorra, é utilizado o índice primário

- **Consulta 3:** `SELECT iP FROM tabela WHERE iP = valor AND iS = valor`
  - Índice primário e secundário
  - Operação de AND sobre igualdade
  - Apenas campo do índice primário

Tamanho	Tipo de Acesso	Chaves Possíveis	Chave Usada	Linhas	Extra
$10^2$	NULL	nenhuma	nenhuma	NULL	impossível
$10^3$	NULL	nenhuma	nenhuma	NULL	impossível
$10^4$	NULL	nenhuma	nenhuma	NULL	impossível
$10^5$	NULL	nenhuma	nenhuma	NULL	impossível
$10^6$	NULL	nenhuma	nenhuma	NULL	impossível

- Impossível → Só há uma chance de ocorrer
- Caso ocorra, é utilizado o índice primário

- **Consulta 4:** `SELECT iP FROM tabela WHERE iS = valor AND iP = valor`
  - Índice primário e secundário
  - Operação de AND sobre igualdade
  - Apenas campo do índice primário

Tamanho	Tipo de Acesso	Chaves Possíveis	Chave Usada	Linhas	Extra
$10^2$	NULL	nenhuma	nenhuma	NULL	impossível
$10^3$	NULL	nenhuma	nenhuma	NULL	impossível
$10^4$	NULL	nenhuma	nenhuma	NULL	impossível
$10^5$	NULL	nenhuma	nenhuma	NULL	impossível
$10^6$	NULL	nenhuma	nenhuma	NULL	impossível

- Impossível → Só há uma chance de ocorrer
- Caso ocorra, é utilizado o índice primário

- **Consulta 5:** `SELECT iP,iS FROM tabela WHERE iP = valor AND iS = valor`
  - Índice primário e secundário
  - Operação de AND sobre igualdade
  - Apenas campo do índice primário e do secundário

Tamanho	Tipo de Acesso	Chaves Possíveis	Chave Usada	Linhas	Extra
$10^2$	NULL	nenhuma	nenhuma	NULL	impossível
$10^3$	NULL	nenhuma	nenhuma	NULL	impossível
$10^4$	NULL	nenhuma	nenhuma	NULL	impossível
$10^5$	NULL	nenhuma	nenhuma	NULL	impossível
$10^6$	NULL	nenhuma	nenhuma	NULL	impossível

- Impossível → Só há uma chance de ocorrer
- Caso ocorra, é utilizado o índice primário

- **Consulta 6:** `SELECT iP,iS FROM tabela WHERE iS = valor AND iP = valor`
  - Índice primário e secundário
  - Operação de AND sobre igualdade
  - Apenas campo do índice primário e do secundário

Tamanho	Tipo de Acesso	Chaves Possíveis	Chave Usada	Linhas	Extra
$10^2$	NULL	nenhuma	nenhuma	NULL	impossível
$10^3$	NULL	nenhuma	nenhuma	NULL	impossível
$10^4$	NULL	nenhuma	nenhuma	NULL	impossível
$10^5$	NULL	nenhuma	nenhuma	NULL	impossível
$10^6$	NULL	nenhuma	nenhuma	NULL	impossível

- Impossível → Só há uma chance de ocorrer
- Caso ocorra, é utilizado o índice primário

- Por que a ação foi “impossível” para todos?

$$P(iP, iS|T) = P(iP|T) \times P(iS|T) = \frac{1}{|T|} \times \frac{C(v_{iS}, T)}{|T|}$$

- Caso dê a sorte de ocorrer
  - Primeiro é analisado o índice primário
  - Como ele é chave primária, é devolvido apenas um valor
- Busca no índice primário

- Introdução
  - MariaDB
  - Índices
- Ambiente
- Resultados
- Avaliação
- Conclusão

- A distribuição uniforme e aleatória afetou os resultados diretamente
- Nem sempre a criação de índices ajuda a melhorar o desempenho
- O comando EXPLAIN ajuda bastante a entender o plano de execução da consulta
- Alguns tempos de execução iniciais afetaram o resultado final para a primeira coleção



- Eficiência do MariaDB
  - Comparar com outros SGBDs
  - Analisar com dados de aplicações reais
- Outros comandos
  - OR
  - ORDER BY
  - JOINS
- Outras instruções
  - INSERT
  - UPDATE
  - DELETE

Página de documentação do MariaDB. Disponível em:  
<<https://mariadb.com/kb/en/mariadb/>>.

CHEIRAN, J. F. Transparências de aulas de Banco de Dados II.  
Universidade Federal do Pampa, 2015.

# Análise de desempenho para o banco de dados MariaDB

Cristiano Daitx Ribeiro  
Igor Antunes Fogaça  
Marcos Vinícius Treviso

Banco de Dados II

17 de julho de 2015

Universidade Federal do Pampa