

# *Hashing*

## *Organização Direta de Arquivos*

### Estrutura de Dados II

## *Aula 08*

# Arquivos Diretos

Os principais métodos de organização e acesso de registros em arquivos diretos são:

1. Dicionário de Dados
2. Endereçamento Indireto, ou *Hashing*



## *Como já foi discutido...*

Diferente da organização sequencial, onde os registros são pesquisados um a um pela ordem de ocorrência, na **organização direta** há uma **relação** definida **entre o valor da chave de um registro e seu endereço no dispositivo de armazenamento** (de acesso direto).



# *Também já foi discutido...*

Nos arquivos com organização direta, **os dados são armazenados e recuperados através do uso da relação entre os registros e seus respectivos endereços.**



# *Busca em Arquivos Diretos*

A **busca** por um determinado registro R **não implica** na necessidade de **varredura sequencial da base de dados**. Esta é efetuada através do acesso ao endereço de R, diretamente, obtido a partir de sua chave de pesquisa.



# Hashing



No hashing, adota-se uma função que, **a partir do valor de chave** de um registro R de dados, **define-se o provável endereço** de R no arquivo:

$$f(\text{Chave}) = \text{endereço}$$

# Hashing



No hashing, a relação chave e endereço dos registros de uma arquivo (direto) é estabelecida por uma função matemática.

$$f(\text{Chave}) = \text{endereço}$$

Esta função é denominada **função hash** ou **função de mapeamento**.

# Função Hashing



$$f(\text{Chave}) = \text{endereço}$$

O endereço gerado com a aplicação da função hash, a partir da chave de busca, é denominado **endereço base**.

# Função Hashing



$$f(\text{Chave}) = \text{endereço}$$

Com a aplicação do hashing, para se efetuar uma operação sobre um registro com **chave de busca C**, **aplica-se a função hash** e obtém-se o **endereço base E**. De posse deste endereço E, a base de dados (arquivo) é acessada de forma direta.

# Função Hashing



Que função usar?

Que características esta função deve apresentar?



# Função Hashing



Considerando como exemplo um sistema de consulta a preços de um estabelecimento comercial:

|    | CÓDIGO (BARRA) | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|----------------|-------------|-------|------------|----------|
| 01 | 284            | lápiz       | 1.20  | voyage     | 08/2049  |
| 02 | 935            | apontador   | 1.50  | dantena    | 08/2030  |
| 03 | 128            | régua       | 2.20  | xavier     | 06/2038  |
| 04 | 443            | caneta      | 2.50  | vitória    | 04/2041  |
| 05 | 339            | caderno     | 6.50  | voyage     | 04/2031  |
| 06 | 845            | borracha    | 1.90  | alba       | 08/2050  |
| 07 | 223            | agenda      | 5.50  | suzano     | 10/2060  |
| 08 | 695            | resma A4    | 15.00 | chamex     | 04/2042  |
| 09 | 241            | calculadora | 25.50 | calc       | 05/2032  |
| 10 | 803            | esquadro    | 2.30  | vermelho   | 06/2080  |

# Função Hashing



Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , quais os endereços dos produtos constantes na base abaixo?

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 01 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 02 | 935    | apontador   | 1.50  | dantena    | 08/2030  |
| 03 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 04 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 05 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 06 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 07 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 08 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |

| CHAVE | ENDEREÇO |
|-------|----------|
| 284   | ?        |
| 935   | ?        |
| 443   | ?        |
| 339   | ?        |
| 223   | ?        |
| 695   | ?        |
| 241   | ?        |
| 803   | ?        |

# Função Hashing



Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , obtém-se os endereços que seguem, para os produtos constantes na base abaixo:

| CHAVE | ENDEREÇO |
|-------|----------|
| 284   | 11       |
| 935   | 12       |
| 443   | 01       |
| 339   | 00       |
| 223   | 02       |
| 695   | 06       |
| 241   | 07       |
| 803   | 10       |

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 01 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 02 | 935    | apontador   | 1.50  | dantena    | 08/2030  |
| 03 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 04 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 05 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 06 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 07 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 08 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |

# Hashing

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 01 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 02 | 935    | apontador   | 1.50  | dantena    | 08/2030  |
| 03 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 04 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 05 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 06 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 07 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 08 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |

Considerando os endereços obtidos a partir da aplicação da função hash  $f(\text{Chave}) = \text{Chave} \bmod 13$  a base de dados deve ser reorganizada?

| CHAVE | ENDEREÇO |
|-------|----------|
| 284   | 11       |
| 935   | 12       |
| 443   | 01       |
| 339   | 00       |
| 223   | 02       |
| 695   | 06       |
| 241   | 07       |
| 803   | 10       |

|     | CÓDIGO | PRODUTO   | PREÇO | FORNECEDOR | VALIDADE |
|-----|--------|-----------|-------|------------|----------|
| 00  | 339    | caderno   | 6.50  | voyage     | 04/2031  |
| 01  | 443    | caneta    | 2.50  | vitória    | 04/2041  |
| 02  | 223    | agenda    | 5.50  | suzano     | 10/2060  |
| 03  |        |           |       |            |          |
| 04  |        |           |       |            |          |
| 05  |        |           |       |            |          |
| ... |        |           |       |            |          |
| 12  | 935    | apontador | 1.50  | dantena    | 08/2030  |

# Hashing



Como efetuar a consulta ao produto com código (chave de busca) igual a **223**?

E a consulta da chave com código 16?

|            | CÓDIGO | PRODUTO   | PREÇO | FORNECEDOR | VALIDADE |
|------------|--------|-----------|-------|------------|----------|
| <b>00</b>  | 339    | caderno   | 6.50  | voyage     | 04/2031  |
| <b>01</b>  | 443    | caneta    | 2.50  | vitória    | 04/2041  |
| <b>02</b>  | 223    | agenda    | 5.50  | suzano     | 10/2060  |
| <b>03</b>  |        |           |       |            |          |
| <b>04</b>  |        |           |       |            |          |
| <b>05</b>  |        |           |       |            |          |
| <b>...</b> |        |           |       |            |          |
| <b>12</b>  | 935    | apontador | 1.50  | dantena    | 08/2030  |

# Hashing



A base de dados original segue abaixo. No processo inicial de cálculo dos endereços, propositadamente, excluimos os produtos **128** e **845**.

|    | CÓDIGO (BARRA) | PRODUTO         | PREÇO       | FORNECEDOR    | VALIDADE       |
|----|----------------|-----------------|-------------|---------------|----------------|
| 01 | 284            | lápiz           | 1.20        | voyage        | 08/2049        |
| 02 | 935            | apontador       | 1.50        | dantena       | 08/2030        |
| 03 | <b>128</b>     | <b>régua</b>    | <b>2.20</b> | <b>xavier</b> | <b>06/2038</b> |
| 04 | 443            | caneta          | 2.50        | vitória       | 04/2041        |
| 05 | 339            | caderno         | 6.50        | voyage        | 04/2031        |
| 06 | <b>845</b>     | <b>borracha</b> | <b>1.90</b> | <b>alba</b>   | <b>08/2050</b> |
| 07 | 223            | agenda          | 5.50        | suzano        | 10/2060        |
| 08 | 695            | resma A4        | 15.00       | chamex        | 04/2042        |
| 09 | 241            | calculadora     | 25.50       | calc          | 05/2032        |
| 10 | 803            | esquadro        | 2.30        | vermelho      | 06/2080        |

# Hashing

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , obtém-se os endereços dos produtos constantes na base abaixo.

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE | CHAVE      | ENDEREÇO  |
|----|--------|-------------|-------|------------|----------|------------|-----------|
| 01 | 284    | lápiz       | 1.20  | voyage     | 08/2049  | 284        | 11        |
| 02 | 935    | apontador   | 1.50  | dantena    | 08/2030  | 935        | 12        |
| 03 | 128    | régua       | 2.20  | xavier     | 06/2038  | <b>128</b> | <b>11</b> |
| 04 | 443    | caneta      | 2.50  | vitória    | 04/2041  | 443        | 01        |
| 05 | 339    | caderno     | 6.50  | voyage     | 04/2031  | 339        | 00        |
| 06 | 845    | borracha    | 1.90  | alba       | 08/2050  | <b>845</b> | <b>00</b> |
| 07 | 223    | agenda      | 5.50  | suzano     | 10/2060  | 223        | 02        |
| 08 | 695    | resma A4    | 15.00 | chamex     | 04/2042  | 695        | 06        |
| 09 | 241    | calculadora | 25.50 | calc       | 05/2032  | 241        | 07        |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  | 803        | 10        |

O que há em comum entre os endereços gerados pelas chaves 128 e 845?

# Hashing

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 01 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 02 | 935    | apontador   | 1.50  | dantena    | 08/2030  |
| 03 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 04 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 05 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 06 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 07 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 08 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |

Considerando os endereços obtidos a partir da aplicação da função hash  $f(\text{chave}) = \text{chave} \bmod 13$

como reorganizar a base de dados, em especial para acesso aos dados de chave **845**?

| CHAVE      | ENDEREÇO  |
|------------|-----------|
| 284        | 11        |
| 935        | 12        |
| <b>128</b> | <b>11</b> |
| 443        | 01        |
| 339        | 00        |
| <b>845</b> | <b>00</b> |
| 223        | 02        |
| ...        |           |

|           | CÓDIGO | PRODUTO   | PREÇO | FORNECEDOR | VALIDADE |
|-----------|--------|-----------|-------|------------|----------|
| <b>00</b> | 339    | caderno   | 6.50  | voyage     | 04/2031  |
| <b>01</b> | 443    | caneta    | 2.50  | vitória    | 04/2041  |
| <b>02</b> | 223    | agenda    | 5.50  | suzano     | 10/2060  |
| <b>03</b> |        |           |       |            |          |
| <b>04</b> |        |           |       |            |          |
| <b>05</b> |        |           |       |            |          |
| ...       |        |           |       |            |          |
| <b>12</b> | 935    | apontador | 1.50  | dantena    | 08/2030  |

# Hashing



Como efetuar a consulta ao produto com código (chave de busca) igual a **845**?

|            | CÓDIGO | PRODUTO   | PREÇO | FORNECEDOR | VALIDADE |
|------------|--------|-----------|-------|------------|----------|
| <b>00</b>  | 339    | caderno   | 6.50  | voyage     | 04/2031  |
| <b>01</b>  | 443    | caneta    | 2.50  | vitória    | 04/2041  |
| <b>02</b>  | 223    | agenda    | 5.50  | suzano     | 10/2060  |
| <b>03</b>  |        |           |       |            |          |
| <b>04</b>  |        |           |       |            |          |
| <b>05</b>  |        |           |       |            |          |
| <b>...</b> |        |           |       |            |          |
| <b>12</b>  | 935    | apontador | 1.50  | dantena    | 08/2030  |

# Estouro ou Colisão



Dada uma chave  $C$ , a partir da aplicação da função hash, obtém-se o endereço  $E$  do registro que contém  $C$ .

Mas duas chaves distintas podem gerar um mesmo endereço  $E$ .

A tentativa de armazenar um registro num endereço já ocupado é denominado **estouro** ou **colisão**.

# Hashing

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , obtém-se os endereços dos produtos constantes na base abaixo.

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 01 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 02 | 935    | apontador   | 1.50  | dantena    | 08/2030  |
| 03 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 04 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 05 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 06 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 07 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 08 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 09 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |

| CHAVE      | ENDEREÇO  |
|------------|-----------|
| 284        | 11        |
| 935        | 12        |
| <b>128</b> | <b>11</b> |
| 443        | 01        |
| 339        | 00        |
| <b>845</b> | <b>00</b> |
| 223        | 02        |
| 695        | 06        |
| 241        | 07        |
| 803        | 10        |

A chave 128 colidiu com a 284. E a chave 845 colidiu com a 339.



# Hashing

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , obtém-se os endereços dos produtos constantes na base abaixo.

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 01 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 02 | 935    | apontador   | 1.50  | dantena    | 08/2030  |
| 03 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 04 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 05 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 06 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 07 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 08 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 09 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |

| CHAVE      | ENDEREÇO  |
|------------|-----------|
| 284        | 11        |
| 935        | 12        |
| <b>128</b> | <b>11</b> |
| 443        | 01        |
| 339        | 00        |
| <b>845</b> | <b>00</b> |
| 223        | 02        |
| 695        | 06        |
| 241        | 07        |
| 803        | 10        |

Acontece a colisão de no máximo duas chaves num mesmo endereço? Justifique:



# Hashing



Sendo assim, para uso efetivo do Hashing, é preciso decidir como tratar as colisões/estouros.

As soluções clássicas mantêm uma estrutura auxiliar, denominada **tabela hash**, que relaciona chaves e endereços; para identificação, em especial, dos **endereços efetivos** das chaves que sofrem colisão.

# Hashing Perfeito

Caso seja identificada uma função de mapeamento que garanta a inexistência de colisões, não é preciso manter a tabela hash; o acesso a qualquer chave pode ser efetuado estritamente pela aplicação da função. Neste caso teríamos o que se denomina **hashing perfeito**.

# • Hashing



As funções de mapeamento clássicas ainda não evitam estouros; assim, o hashing é composto por dois processos:

1. aplicação da função de mapeamento
2. tratamento de colisão

# Função de Mapeamento

Uma função hash deve mapear chaves em inteiros dentro do intervalo  $[0..M-1]$  sendo  $M$  o número máximo de registros a serem manipulados pela aplicação, e/ou o espaço de memória disponível para a aplicação.



# Funções de Mapeamento



São funções de mapeamento clássicas:

**1. Resto da Divisão** onde:

$$f(\text{Chave}) = \text{Chave} \bmod M$$

sendo  $M$  o tamanho de memória disponível para armazenamento e/ou número máximo de dados manipulados pela aplicação computacional.

# Funções de Mapeamento



## 2. Resto da Divisão Variante onde:

$$f(\text{Chave}) = \text{Chave} \bmod P$$

sendo  $P$  o menor número primo maior que  $M$  constante no Resto da Divisão.

# Tratamento de Colisão



Definem o destino dos registros que sofrem colisão.

Os métodos clássicos de tratamento de colisões são:

1. Coalescido
2. Amostragem Linear
3. Duplo

# Hashing Coalescido



Havendo colisão, a chave associada ao seu respectivo endereço base, é relacionada a este, em tabela hash, por meio de um link. Desta forma, na tabela hash são mantidos: chave, endereço e link. E a chave colidida é armazenada na primeira posição livre localizada após o seu endereço base.

# Hashing Coalescido

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , e hashing coalescido, montar tabela hash:

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 01 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 02 | 935    | apontador   | 1.50  | dantena    | 08/2030  |
| 03 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 04 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 05 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 06 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 07 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 08 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 09 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |

| CHAVE      | ENDEREÇO  |
|------------|-----------|
| 284        | 11        |
| 935        | 12        |
| <b>128</b> | <b>11</b> |
| 443        | 01        |
| 339        | 00        |
| <b>845</b> | <b>00</b> |
| 223        | 02        |
| 695        | 06        |
| 241        | 07        |
| 803        | 10        |

## Tabela Hash

|    | CHAVE | LINK |
|----|-------|------|
| 00 |       | -1   |
| 01 |       | -1   |
| 02 |       | -1   |
| 03 |       | -1   |
| 04 |       | -1   |
| 05 |       | -1   |
| 06 |       | -1   |
| 07 |       | -1   |
| 08 |       | -1   |
| 09 |       | -1   |
| 10 |       | -1   |
| 11 |       | -1   |
| 12 |       | -1   |

**Hashing Coalescido:** Havendo colisão, a chave associada ao seu respectivo endereço base, é relacionada a este, em tabela hash, por meio de um link. Desta forma, na tabela hash são mantidos: chave, endereço e link. E a chave colidida é armazenada na primeira posição livre localizada após o seu endereço base.

# Hashing Coalescido

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , e hashing coalescido, obtém-se a tabela hash:

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 01 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 02 | 935    | apontador   | 1.50  | dantena    | 08/2030  |
| 03 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 04 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 05 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 06 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 07 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 08 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 09 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |

## Tabela Hash

|    | CHAVE | LINK |
|----|-------|------|
| 00 | 128   | 02   |
| 01 | 443   | -1   |
| 02 | 339   | 03   |
| 03 | 845   | 04   |
| 04 | 223   | -1   |
| 05 |       | -1   |
| 06 | 695   | -1   |
| 07 | 241   | -1   |
| 08 |       | -1   |
| 09 |       | -1   |
| 10 | 803   | -1   |
| 11 | 284   | 00   |
| 12 | 935   | -1   |



O arquivo principal deve ser remodelado?

# Hashing Coalescido

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , e hashing coalescido, obtém-se a tabela hash e o arquivo de dados abaixo.

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 00 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 01 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 02 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 03 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 04 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 05 |        |             |       |            |          |
| 06 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 07 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 08 |        |             |       |            |          |
| 09 |        |             |       |            |          |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |
| 11 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 12 | 935    | apontador   | 1.50  | dantena    | 08/2030  |

## Tabela Hash

|    | CHAVE | LINK |
|----|-------|------|
| 00 | 128   | 02   |
| 01 | 443   | -1   |
| 02 | 339   | 03   |
| 03 | 845   | 04   |
| 04 | 223   | -1   |
| 05 |       | -1   |
| 06 | 695   | -1   |
| 07 | 241   | -1   |
| 08 |       | -1   |
| 09 |       | -1   |
| 10 | 803   | -1   |
| 11 | 284   | 00   |
| 12 | 935   | -1   |



Como localizar (consultar) o registro 803?

# Hashing Coalescido

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , e hashing coalescido, obtém-se a tabela hash e o arquivo de dados abaixo.

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 00 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 01 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 02 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 03 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 04 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 05 |        |             |       |            |          |
| 06 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 07 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 08 |        |             |       |            |          |
| 09 |        |             |       |            |          |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |
| 11 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 12 | 935    | apontador   | 1.50  | dantena    | 08/2030  |

## Tabela Hash

|    | CHAVE | LINK |
|----|-------|------|
| 00 | 128   | 02   |
| 01 | 443   | -1   |
| 02 | 339   | 03   |
| 03 | 845   | 04   |
| 04 | 223   | -1   |
| 05 |       | -1   |
| 06 | 695   | -1   |
| 07 | 241   | -1   |
| 08 |       | -1   |
| 09 |       | -1   |
| 10 | 803   | -1   |
| 11 | 284   | 00   |
| 12 | 935   | -1   |



Como localizar (consultar) o registro 223?

# Hashing Coalescido

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , e hashing coalescido, obtém-se a tabela hash e o arquivo de dados abaixo.

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 00 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 01 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 02 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 03 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 04 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 05 |        |             |       |            |          |
| 06 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 07 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 08 |        |             |       |            |          |
| 09 |        |             |       |            |          |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |
| 11 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 12 | 935    | apontador   | 1.50  | dantena    | 08/2030  |

## Tabela Hash

|    | CHAVE | LINK |
|----|-------|------|
| 00 | 128   | 02   |
| 01 | 443   | -1   |
| 02 | 339   | 03   |
| 03 | 845   | 04   |
| 04 | 223   | -1   |
| 05 |       | -1   |
| 06 | 695   | -1   |
| 07 | 241   | -1   |
| 08 |       | -1   |
| 09 |       | -1   |
| 10 | 803   | -1   |
| 11 | 284   | 00   |
| 12 | 935   | -1   |

Diz-se que o registro com chave 223 tem **endereço base 02**, e **endereço efetivo 04**.

# Hashing Coalescido

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , e hashing coalescido, obtém-se a tabela hash e o arquivo de dados abaixo.

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 00 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 01 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 02 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 03 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 04 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 05 |        |             |       |            |          |
| 06 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 07 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 08 |        |             |       |            |          |
| 09 |        |             |       |            |          |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |
| 11 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 12 | 935    | apontador   | 1.50  | dantena    | 08/2030  |

## Tabela Hash

|    | CHAVE | LINK |
|----|-------|------|
| 00 | 128   | 02   |
| 01 | 443   | -1   |
| 02 | 339   | 03   |
| 03 | 845   | 04   |
| 04 | 223   | -1   |
| 05 |       | -1   |
| 06 | 695   | -1   |
| 07 | 241   | -1   |
| 08 |       | -1   |
| 09 |       | -1   |
| 10 | 803   | -1   |
| 11 | 284   | 00   |
| 12 | 935   | -1   |

O método de tratamento de colisão em estudo é denominado coalescido, por produzir uma **coalha** entre os dados, em meio aos endereços base e efetivo. Observar a chave 845, de endereço 00, onde consta o 128, de endereço base 11.

# Hashing Coalescido

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , e hashing coalescido, obtém-se a tabela hash e o arquivo de dados abaixo.

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 00 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 01 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 02 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 03 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 04 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 05 |        |             |       |            |          |
| 06 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 07 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 08 |        |             |       |            |          |
| 09 |        |             |       |            |          |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |
| 11 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 12 | 935    | apontador   | 1.50  | dantena    | 08/2030  |

## Tabela Hash

|    | CHAVE | LINK |
|----|-------|------|
| 00 | 128   | 02   |
| 01 | 443   | -1   |
| 02 | 339   | 03   |
| 03 | 845   | 04   |
| 04 | 223   | -1   |
| 05 |       | -1   |
| 06 | 695   | -1   |
| 07 | 241   | -1   |
| 08 |       | -1   |
| 09 |       | -1   |
| 10 | 803   | -1   |
| 11 | 284   | 00   |
| 12 | 935   | -1   |



Onde reside a tabela hash?

# Hashing Coalescido

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , e hashing coalescido, obtém-se a tabela hash e o arquivo de dados abaixo.

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 00 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 01 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 02 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 03 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 04 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 05 |        |             |       |            |          |
| 06 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 07 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 08 |        |             |       |            |          |
| 09 |        |             |       |            |          |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |
| 11 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 12 | 935    | apontador   | 1.50  | dantena    | 08/2030  |

## Tabela Hash

|    | CHAVE | LINK |
|----|-------|------|
| 00 | 128   | 02   |
| 01 | 443   | -1   |
| 02 | 339   | 03   |
| 03 | 845   | 04   |
| 04 | 223   | -1   |
| 05 |       | -1   |
| 06 | 695   | -1   |
| 07 | 241   | -1   |
| 08 |       | -1   |
| 09 |       | -1   |
| 10 | 803   | -1   |
| 11 | 284   | 00   |
| 12 | 935   | -1   |



Como é definido o tamanho da tabela hash?

# Hashing Coalescido

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , e hashing coalescido, obtém-se a tabela hash e o arquivo de dados abaixo.

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 00 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 01 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 02 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 03 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 04 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 05 |        |             |       |            |          |
| 06 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 07 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 08 |        |             |       |            |          |
| 09 |        |             |       |            |          |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |
| 11 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 12 | 935    | apontador   | 1.50  | dantena    | 08/2030  |

## Tabela Hash

|    | CHAVE | LINK |
|----|-------|------|
| 00 | 128   | 02   |
| 01 | 443   | -1   |
| 02 | 339   | 03   |
| 03 | 845   | 04   |
| 04 | 223   | -1   |
| 05 |       | -1   |
| 06 | 695   | -1   |
| 07 | 241   | -1   |
| 08 |       | -1   |
| 09 |       | -1   |
| 10 | 803   | -1   |
| 11 | 284   | 00   |
| 12 | 935   | -1   |



Como é definida a função hash?

Além da consulta, que outras operações são necessárias para manutenção de dados mantidos em arquivos diretos organizados por hashing?



# Operações sobre Arquivos Diretos

1. Consulta
2. Inserção
3. Remoção
4. Alteração



# Hashing Coalescido

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , e hashing coalescido, obtém-se a tabela hash e o arquivo de dados abaixo.

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 00 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 01 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 02 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 03 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 04 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 05 |        |             |       |            |          |
| 06 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 07 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 08 |        |             |       |            |          |
| 09 |        |             |       |            |          |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |
| 11 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 12 | 935    | apontador   | 1.50  | dantena    | 08/2030  |

## Tabela Hash

|    | CHAVE | LINK |
|----|-------|------|
| 00 | 128   | 02   |
| 01 | 443   | -1   |
| 02 | 339   | 03   |
| 03 | 845   | 04   |
| 04 | 223   | -1   |
| 05 |       | -1   |
| 06 | 695   | -1   |
| 07 | 241   | -1   |
| 08 |       | -1   |
| 09 |       | -1   |
| 10 | 803   | -1   |
| 11 | 284   | 00   |
| 12 | 935   | -1   |



Como incluir o registro de chave 09?

# Hashing Coalescido

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , e hashing coalescido, obtém-se a tabela hash e o arquivo de dados abaixo.

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 00 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 01 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 02 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 03 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 04 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 05 |        |             |       |            |          |
| 06 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 07 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 08 |        |             |       |            |          |
| 09 |        |             |       |            |          |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |
| 11 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 12 | 935    | apontador   | 1.50  | dantena    | 08/2030  |

## Tabela Hash

|    | CHAVE | LINK |
|----|-------|------|
| 00 | 128   | 02   |
| 01 | 443   | -1   |
| 02 | 339   | 03   |
| 03 | 845   | 04   |
| 04 | 223   | -1   |
| 05 |       | -1   |
| 06 | 695   | -1   |
| 07 | 241   | -1   |
| 08 |       | -1   |
| 09 |       | -1   |
| 10 | 803   | -1   |
| 11 | 284   | 00   |
| 12 | 935   | -1   |



Como incluir o registro de chave 11?

Descrever a operação de **inserção** em arquivo com organização direta mantido por hashing.



# Hashing Coalescido

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , e hashing coalescido, obtém-se a tabela hash e o arquivo de dados abaixo.

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 00 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 01 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 02 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 03 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 04 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 05 |        |             |       |            |          |
| 06 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 07 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 08 |        |             |       |            |          |
| 09 |        |             |       |            |          |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |
| 11 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 12 | 935    | apontador   | 1.50  | dantena    | 08/2030  |

## Tabela Hash

|    | CHAVE | LINK |
|----|-------|------|
| 00 | 128   | 02   |
| 01 | 443   | -1   |
| 02 | 339   | 03   |
| 03 | 845   | 04   |
| 04 | 223   | -1   |
| 05 |       | -1   |
| 06 | 695   | -1   |
| 07 | 241   | -1   |
| 08 |       | -1   |
| 09 |       | -1   |
| 10 | 803   | -1   |
| 11 | 284   | 00   |
| 12 | 935   | -1   |



Como excluir o registro de chave 241?

# Hashing Coalescido

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , e hashing coalescido, obtém-se a tabela hash e o arquivo de dados abaixo.

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 00 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 01 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 02 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 03 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 04 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 05 |        |             |       |            |          |
| 06 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 07 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 08 |        |             |       |            |          |
| 09 |        |             |       |            |          |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |
| 11 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 12 | 935    | apontador   | 1.50  | dantena    | 08/2030  |

## Tabela Hash

|    | CHAVE | LINK |
|----|-------|------|
| 00 | 128   | 02   |
| 01 | 443   | -1   |
| 02 | 339   | 03   |
| 03 | 845   | 04   |
| 04 | 223   | -1   |
| 05 |       | -1   |
| 06 | 695   | -1   |
| 07 | 241   | -1   |
| 08 |       | -1   |
| 09 |       | -1   |
| 10 | 803   | -1   |
| 11 | 284   | 00   |
| 12 | 935   | -1   |



Como excluir o registro de chave 128?

Descrever a operação de **exclusão** em arquivo com organização direta mantido por hashing.



# Hashing Coalescido

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , e hashing coalescido, obtém-se a tabela hash e o arquivo de dados abaixo.

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 00 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 01 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 02 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 03 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 04 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 05 |        |             |       |            |          |
| 06 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 07 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 08 |        |             |       |            |          |
| 09 |        |             |       |            |          |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |
| 11 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 12 | 935    | apontador   | 1.50  | dantena    | 08/2030  |

## Tabela Hash

|    | CHAVE | LINK |
|----|-------|------|
| 00 | 128   | 02   |
| 01 | 443   | -1   |
| 02 | 339   | 03   |
| 03 | 845   | 04   |
| 04 | 223   | -1   |
| 05 |       | -1   |
| 06 | 695   | -1   |
| 07 | 241   | -1   |
| 08 |       | -1   |
| 09 |       | -1   |
| 10 | 803   | -1   |
| 11 | 284   | 00   |
| 12 | 935   | -1   |



Como alterar o registro de chave 223?

◦ O hashing apresenta vantagens e/ou desvantagens em relação ao dicionário de dados? Justifique sua resposta:





Considerando:

(1) as chaves 34, 12, 25, 49, 45, 13 e 92 e

(2) a função de mapeamento  $f(\text{chave}) = \text{chave} \bmod 11$ ;

Esquematizar a tabela hash e o arquivo de dados (com situação problema hipotética adequada) obtidos a partir do tratamento de colisão coalescido.

- Partindo da resposta final dada à questão anterior, esquematizar as exclusões das chaves 34 e 49.



# Hashing por Amostragem Linear



Neste método o registro colidido é armazenado na primeira posição livre encontrada a partir de ponto de colisão. Vale destacar que a busca por posição livre considera a estrutura circular.

# Hashing por Amostragem Linear

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , e hashing por amostragem linear, montar tabela hash:

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 01 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 02 | 935    | apontador   | 1.50  | dantena    | 08/2030  |
| 03 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 04 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 05 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 06 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 07 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 08 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 09 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |

| CHAVE      | ENDEREÇO  |
|------------|-----------|
| 284        | 11        |
| 935        | 12        |
| <b>128</b> | <b>11</b> |
| 443        | 01        |
| 339        | 00        |
| <b>845</b> | <b>00</b> |
| 223        | 02        |
| 695        | 06        |
| 241        | 07        |
| 803        | 10        |

## Tabela Hash

|    | CHAVE |
|----|-------|
| 00 |       |
| 01 |       |
| 02 |       |
| 03 |       |
| 04 |       |
| 05 |       |
| 06 |       |
| 07 |       |
| 08 |       |
| 09 |       |
| 10 |       |
| 11 |       |
| 12 |       |

O registro colidido é armazenado na primeira posição livre encontrada a partir de ponto de colisão. Vale destacar que a busca por posição livre considera a estrutura circular.

# Hashing por Amostragem Linear

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , e hashing por amostragem linear, montar tabela hash:

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 01 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 02 | 935    | apontador   | 1.50  | dantena    | 08/2030  |
| 03 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 04 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 05 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 06 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 07 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 08 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 09 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |

## Tabela Hash

|    | CHAVE |
|----|-------|
| 00 | 128   |
| 01 | 443   |
| 02 | 339   |
| 03 | 845   |
| 04 | 223   |
| 05 |       |
| 06 | 695   |
| 07 | 241   |
| 08 |       |
| 09 |       |
| 10 | 803   |
| 11 | 284   |
| 12 | 935   |



Como fica o arquivo de dados?

# Hashing por Amostragem Linear

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , e hashing por amostragem linear, montar tabela hash:

## Tabela Hash

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 00 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 01 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 02 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 03 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 04 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 05 |        |             |       |            |          |
| 06 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 07 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 08 |        |             |       |            |          |
| 09 |        |             |       |            |          |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |
| 11 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 12 | 935    | apontador   | 1.50  | dantena    | 08/2030  |

|    | CHAVE |
|----|-------|
| 00 | 128   |
| 01 | 443   |
| 02 | 339   |
| 03 | 845   |
| 04 | 223   |
| 05 |       |
| 06 | 695   |
| 07 | 241   |
| 08 |       |
| 09 |       |
| 10 | 803   |
| 11 | 284   |
| 12 | 935   |

# Hashing por Amostragem Linear

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , e hashing por amostragem linear, montar tabela hash:

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 00 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 01 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 02 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 03 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 04 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 05 |        |             |       |            |          |
| 06 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 07 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 08 |        |             |       |            |          |
| 09 |        |             |       |            |          |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |
| 11 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 12 | 935    | apontador   | 1.50  | dantena    | 08/2030  |

## Tabela Hash

|    | CHAVE |
|----|-------|
| 00 | 128   |
| 01 | 443   |
| 02 | 339   |
| 03 | 845   |
| 04 | 223   |
| 05 |       |
| 06 | 695   |
| 07 | 241   |
| 08 |       |
| 09 |       |
| 10 | 803   |
| 11 | 284   |
| 12 | 935   |



Como consultar o registro de chave 339?

# Hashing por Amostragem Linear

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , e hashing por amostragem linear, montar tabela hash:

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 00 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 01 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 02 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 03 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 04 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 05 |        |             |       |            |          |
| 06 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 07 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 08 |        |             |       |            |          |
| 09 |        |             |       |            |          |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |
| 11 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 12 | 935    | apontador   | 1.50  | dantena    | 08/2030  |

## Tabela Hash

|    | CHAVE |
|----|-------|
| 00 | 128   |
| 01 | 443   |
| 02 | 339   |
| 03 | 845   |
| 04 | 223   |
| 05 |       |
| 06 | 695   |
| 07 | 241   |
| 08 |       |
| 09 |       |
| 10 | 803   |
| 11 | 284   |
| 12 | 935   |



Como consultar o registro de chave 10?

# Hashing por Amostragem Linear

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , e hashing por amostragem linear, montar tabela hash:

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 00 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 01 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 02 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 03 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 04 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 05 |        |             |       |            |          |
| 06 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 07 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 08 |        |             |       |            |          |
| 09 |        |             |       |            |          |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |
| 11 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 12 | 935    | apontador   | 1.50  | dantena    | 08/2030  |

## Tabela Hash

|    | CHAVE |
|----|-------|
| 00 | 128   |
| 01 | 443   |
| 02 | 339   |
| 03 | 845   |
| 04 | 223   |
| 05 |       |
| 06 | 695   |
| 07 | 241   |
| 08 |       |
| 09 |       |
| 10 | 803   |
| 11 | 284   |
| 12 | 935   |



Como incluir o registro de chave 09?

# Hashing por Amostragem Linear

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , e hashing por amostragem linear, montar tabela hash:

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 00 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 01 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 02 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 03 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 04 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 05 |        |             |       |            |          |
| 06 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 07 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 08 |        |             |       |            |          |
| 09 |        |             |       |            |          |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |
| 11 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 12 | 935    | apontador   | 1.50  | dantena    | 08/2030  |

## Tabela Hash

|    | CHAVE |
|----|-------|
| 00 | 128   |
| 01 | 443   |
| 02 | 339   |
| 03 | 845   |
| 04 | 223   |
| 05 |       |
| 06 | 695   |
| 07 | 241   |
| 08 |       |
| 09 |       |
| 10 | 803   |
| 11 | 284   |
| 12 | 935   |



Como incluir o registro de chave 23?

# Hashing por Amostragem Linear

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , e hashing por amostragem linear, montar tabela hash:

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 00 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 01 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 02 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 03 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 04 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 05 |        |             |       |            |          |
| 06 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 07 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 08 |        |             |       |            |          |
| 09 |        |             |       |            |          |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |
| 11 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 12 | 935    | apontador   | 1.50  | dantena    | 08/2030  |

## Tabela Hash

|    | CHAVE |
|----|-------|
| 00 | 128   |
| 01 | 443   |
| 02 | 339   |
| 03 | 845   |
| 04 | 223   |
| 05 |       |
| 06 | 695   |
| 07 | 241   |
| 08 |       |
| 09 |       |
| 10 | 803   |
| 11 | 284   |
| 12 | 935   |



Como excluir o registro de chave 241?

# Hashing por Amostragem Linear

Aplicando a função hash:  $f(\text{Chave}) = \text{Chave} \bmod 13$ , e hashing por amostragem linear, montar tabela hash:

|    | CÓDIGO | PRODUTO     | PREÇO | FORNECEDOR | VALIDADE |
|----|--------|-------------|-------|------------|----------|
| 00 | 128    | régua       | 2.20  | xavier     | 06/2038  |
| 01 | 443    | caneta      | 2.50  | vitória    | 04/2041  |
| 02 | 339    | caderno     | 6.50  | voyage     | 04/2031  |
| 03 | 845    | borracha    | 1.90  | alba       | 08/2050  |
| 04 | 223    | agenda      | 5.50  | suzano     | 10/2060  |
| 05 |        |             |       |            |          |
| 06 | 695    | resma A4    | 15.00 | chamex     | 04/2042  |
| 07 | 241    | calculadora | 25.50 | calc       | 05/2032  |
| 08 |        |             |       |            |          |
| 09 |        |             |       |            |          |
| 10 | 803    | esquadro    | 2.30  | vermelho   | 06/2080  |
| 11 | 284    | lápiz       | 1.20  | voyage     | 08/2049  |
| 12 | 935    | apontador   | 1.50  | dantena    | 08/2030  |

## Tabela Hash

|    | CHAVE |
|----|-------|
| 00 | 128   |
| 01 | 443   |
| 02 | 339   |
| 03 | 845   |
| 04 | 223   |
| 05 |       |
| 06 | 695   |
| 07 | 241   |
| 08 |       |
| 09 |       |
| 10 | 803   |
| 11 | 284   |
| 12 | 935   |



Como excluir o registro de chave 128?

Hã vantagens e/ou desvantagens do método de tratamento de colisão por amostragem linear, em relação ao coalescido? Justifique:





Considerando:

(1) as chaves: 34 – 12 – 49 – 31 – 45 – 13 e 92,

(2) a função de mapeamento:  $f(\text{chave}) = \text{chave} \bmod 11$

(3) o método de tratamento de colisão amostragem linear; esquematizar a tabela hash e o arquivo de dados (relativo à situação hipotética) correspondente.

- ° Partindo da resposta dada à questão anterior, esquematizar as exclusões das chaves 34 e 49.



# Hashing Duplo



Método variante do Amostragem Linear que apresenta duas funções: uma para mapeamento e outra para calcular o deslocamento dos registro colididos na tabela hash para localização de posição livre, de forma a melhor distribuir os registros na ocupação da estrutura.

Na Amostragem Linear, pode-se dizer que a o deslocamento é constante e de uma unidade.



Considerando:

(1) as chaves: 34 – 12 – 49 – 31 – 45 – 13 e 92

(2) a função de mapeamento:

$$f(\text{chave}) = \text{chave} \bmod 11$$

(3) o método de tratamento de colisão duplo

(4) a função de deslocamento:

$$f(\text{chave}) = (\text{chave} \bmod 3) + 1$$

esquematizar a tabela hash e o arquivo de dados correspondente.

Efetuar a consulta ao registro  
com chave 49.



Efetuar a consulta ao registro  
com chave 45.



Efetuar a inserção do registro  
com chave 14.



Efetuar a exclusão do registro  
com chave 34.



# Rehashing



O rehashing, como o próprio nome sugere, consiste em refazer a estrutura hash.

# Rehashing



Dada uma estrutura hash inicial, de tamanho  $N$  a ser transformada numa estrutura de tamanho  $T$ :

- ❑ Se  $T > N$  o objetivo é diminuir a frequência de colisões e estouros.
- ❑ Se  $T = N$  o objetivo é efetuar uma limpeza, reorganizar a estrutura hash.
- ❑ Se  $T < N$  o objetivo é poupar espaço.

# Rehashing



A tabela hash e o arquivo de dados devem ser ajustados e, no processo, deve ser criada uma estrutura auxiliar para controle das alterações efetuadas.



Considerando:

1. Chaves(endereços): 27(5) – 18(7) – 29(7) – 28(6) – 39(6) – 13(2) – 16(5) – 42(9) – 46(2)
2. Função hash inicial:  $f(\text{chave}) = \text{chave} \bmod 11$
3. Tratamento de colisão por amostragem linear
4. Função hash de expansão:  $g(\text{chave}) = \text{chave} \bmod 19$
5. Chaves(novos endereços): 27(8) – 18(18) – 29(10) – 28(9) – 39(1) – 13(13) – 16(16) – 42(4) – 46(8)

Efetuar processo de rehashing.

Descreva os passos gerais de execução do processo de rehashing.





Esquematizar o rehashing de:

(1) chaves: 11 – 5 – 23 – 22 – 38 – 55 – 39;

(2) função inicial:  $f(\text{chave}) = \text{chave} \bmod 11$ ;

(3) função de expansão:  $f(\text{chave}) = \text{chave} \bmod 17$ ;

(4) tratamento de colisão: amostragem linear.

# Funções de Mapeamento

## Exercício

Pesquisar e apresentar outras funções de mapeamento.

Descrever funcionamento da função e apresentar o intervalo de valores gerados por estas (endereços).

Pesquisar, inclusive, funções aplicadas sobre chaves do tipo string.

# Método de Tratamento de Colisão

## Exercício

Pesquisar e apresentar o método ENCADEAMENTO DE EXCEDENTES.

Apresentar exemplo (esquema).

Descrever operações.

# Complementar Estudos...



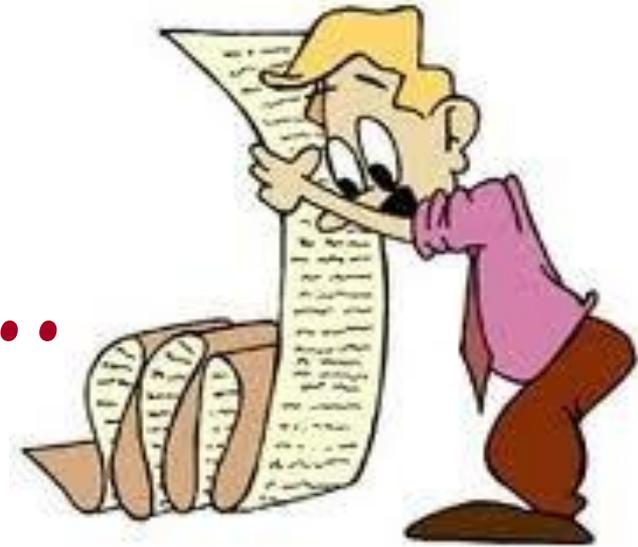
**File Organization and Processing**

*Allan L Tharp*

## **Capítulo 3**

**Direct File Organization**

*Próximo passo...*



# Arquivos Sequenciais Indexados