

Instituto de Ciências Matemáticas e de Computação

Universidade de São Paulo

Compactação e Reuso de espaço (parte 2)

SCC0607 – Estrutura de Dados III Anderson Canale Garcia

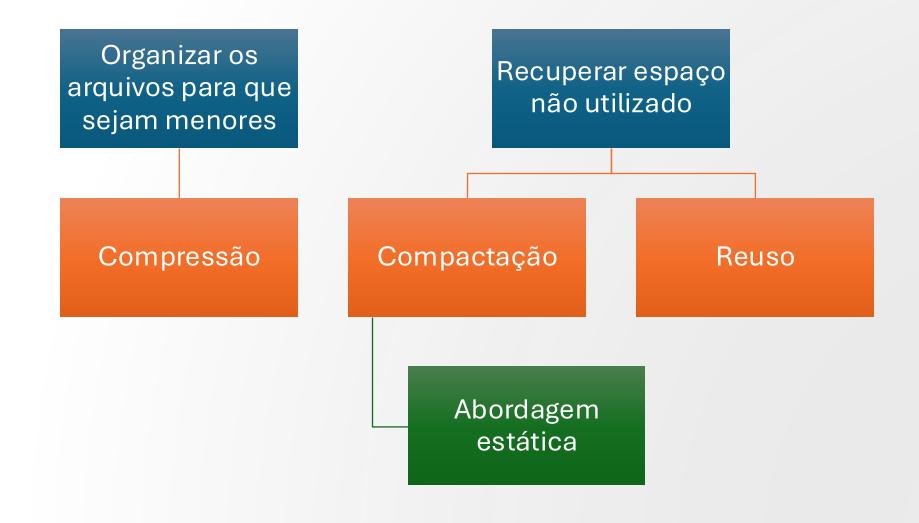
Baseado nos materiais dos professores:
Cristina D. Aguiar
Leandro C. Cintra
M.C.F de Oliveira
Rosane Minghim

Sumário

- Contexto
- Compactação e reuso de espaço
 - Abordagem estática
 - Abordagem dinâmica
 - Registros de tamanho fixo
 - Registros de tamanho variável

Na última aula...

Organizar arquivos visando performance



Compactação e reuso de espaço

Lidando com a deterioração da organização dos arquivos à medida que um arquivo é modificado

Alterações de registros

Adição

Relativamente simples

Remoção

Atualização

Remoção seguida de adição

Quando um registro é removido, deve-se posteriormente reutilizar o espaço do registro

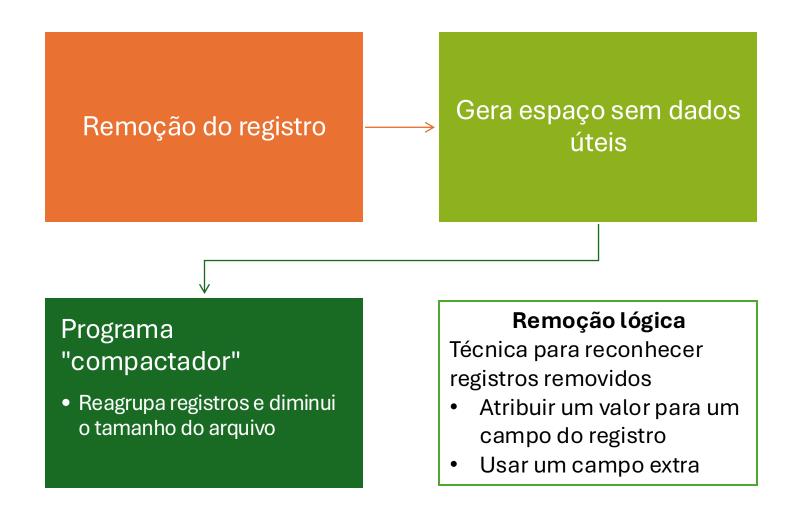
Compactação

• Busca por regiões do arquivo que não contêm dados

Recupera os espaços perdidos

E como fazer?

Abordagem estática



Abordagem estática



Não faz nada em um intervalo de tempo Δt



Durante ∆t (Remoção lógica)

Registros removidos são marcados, porém não são reaproveitados Novas inserções são realizadas no final do arquivo Buscas desconsideram os registros marcados como removidos



Após Δt (Remoção física)

Programa é executado para reconstruir o arquivo Todos os registros removidos são descartados

Exemplo

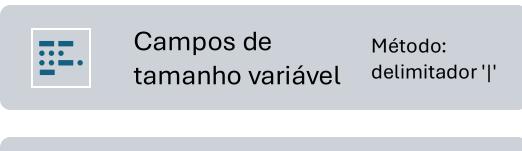
- Arquivo de dados
 - o Campos de tamanho variável
 - Método: delimitador '|'
 - o Registros de tamanho fixo
 - Tamanho do registro: 30 bytes

Exemplo

Dados originais

Nome	Endereço	Cidade
Maria	Rua 1	São Carlos
João	Rua A	Rio Claro
Pedro	Rua XV	São Carlos
Ana	Rua Augusto Paiva	Ibaté

Características do arquivo





Tamanho do registro: 30 bytes

Exemplo – Arquivo original

0	1	2	3	4	5	6	7	8	9	0	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
М	А	R	Ι	А		R	U	А	b	1		S	А	0	b	С	А	R	L	0	S		X	X	X	X	X	X	X
30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
J	0	Α	0		R	U	Α	b	Α		R	I	0	b	C	L	Α	R	0		X	X	X	X	X	X	X	X	X
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	7 9	80	81	82	83	84	85	86	87	88	89
Р	Е	D	R	0		R	U	Α	b	X	V		S	Α	0	b	С	Α	R	L	0	S		X	X	X	X	X	X
90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119
Α	N	А		R	U	Α	b	А	כ	G	U	S	Τ	0	b	Р	Α	I	٧	Α		Ι	В	Α	Τ	ш		X	X
120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149

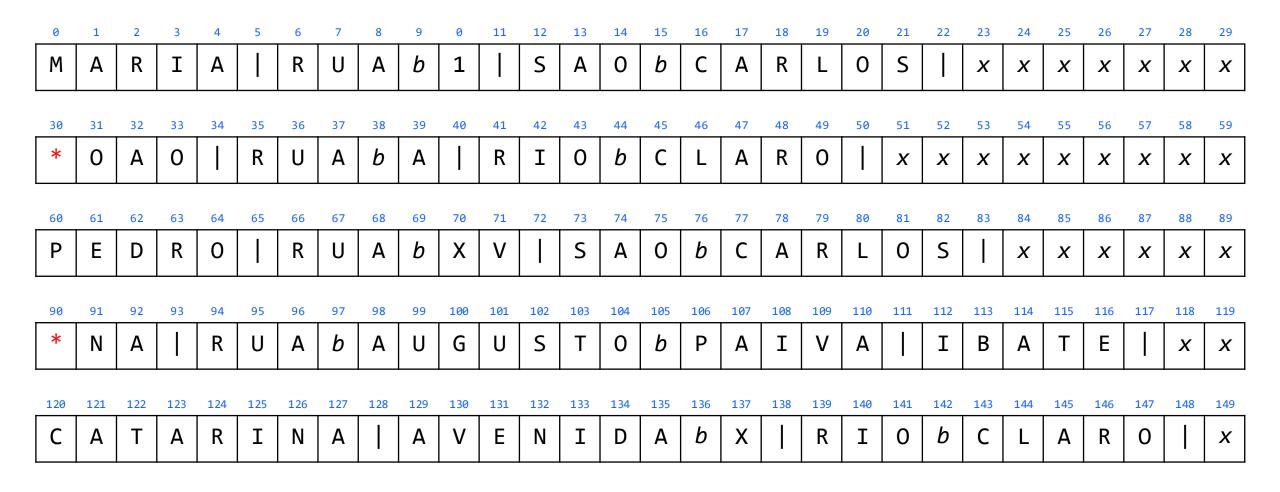
3. Remova os registros com RRN = 1 e RRN = 3

Exemplo – Arquivo após remoção lógica

0	1	2	3	4	5	6	7	8	9	0	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
М	Α	R	I	А		R	U	Α	b	1		S	Α	0	b	C	Α	R	L	0	S		X	X	X	X	X	X	X
30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
*	0	А	0		R	U	Α	b	Α		R	Ι	0	b	C	Ш	Α	R	0		X	X	X	X	X	X	X	X	X
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89
Р	Е	D	R	0		R	U	Α	b	X	V		S	Α	0	b	С	Α	R	L	0	S		X	X	X	X	X	X
90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119
*	N	Α		R	U	А	b	Α	U	G	U	S	Т	0	b	Р	Α	Ι	V	Α		I	В	Α	Т	Е		X	X
120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149

5. Insira um novo registro: CATARINA, AVENIDA X, RIO CLARO

Exemplo – Arquivo após inserção



6. Compacte o arquivo

Exemplo – Arquivo compactado

0	1	2	3	4	5	6	7	8	9	0	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
M	Α	R	Ι	Α		R	U	Α	b	1		S	А	0	b	С	Α	R	L	0	S		X	X	X	X	X	X	X
30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
Р	Е	D	R	0		R	U	А	b	X	٧		S	А	0	b	С	Α	R	L	0	S		X	X	X	X	X	X
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89
С	Α	Т	Α	R	Ι	N	Α		А	V	Е	Ν	Ι	D	А	b	X		R	I	0	b	С	L	А	R	0		X
90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119
120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149

Abordagem estática



Ideal para arquivos estáticos ou com baixa frequência de alterações/deleções



Processo em lote, geralmente realizado em períodos específicos (fim do ano fiscal, por exemplo)



Garante alta utilização do espaço, mas pode ser ineficiente para sistemas altamente dinâmicos.

E se...

... o sistema exigir alta disponibilidade?

 A compactação em lote pode gerar indisponibilidade durante o processo.

... a taxa de inserções e deleções for alta?

 A fragmentação do arquivo pode aumentar significativamente, impactando o desempenho.

Organizar arquivos visando performance



Reuso de espaço

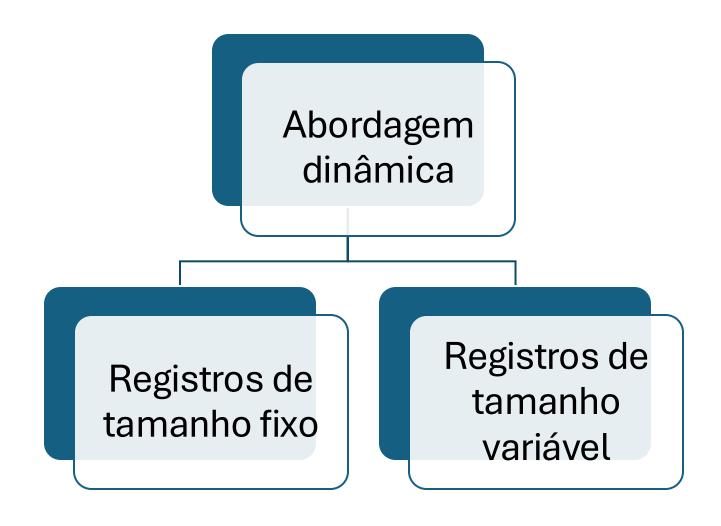
Abordagem estática

- Ideal para arquivos estáticos ou com baixa frequência de alterações/deleções
- Precisamos garantir
 - 1. Que os registros removidos sejam marcados/identificados

Abordagem dinâmica

- Indicada para aplicações interativas que acessam arquivos altamente voláteis
- Precisamos garantir
 - 1. Que os registros removidos sejam marcados/identificados
 - 2. Que o espaço do registro removido possa ser localizado e reutilizado

Como localizar espaços vazios?



Abordagem dinâmica: Registros de **tamanho fixo**

Abordagem dinâmica: registros de tamanho fixo

- Solução
 - Lista encadeada de registros eliminados
- Características
 - Lista: constitui-se dos RRNs dos registros marcados como logicamente removidos
 - Cabeça da lista: armazenada no registro de cabeçalho do arquivo
 - Inserção e reuso de espaço: ocorrem sempre no início da lista

Implementação: PILHA

Operações

Remoção

- Marca o registro como logicamente removido
- Insere o registro na lista de registros logicamente removidos (empilha)

Inserção

- Remove o registro da lista de registros logicamente removidos (desempilha)
- Insere os dados no espaço do registro desempilhado

Atualização

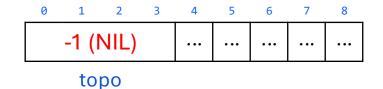
- Ocorre no próprio registro
 - Por que?

Pergunta

• Onde e como mantemos nossa pilha?

Exemplo – Arquivo original

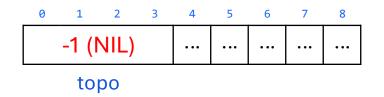
Registro de cabeçalho



	6.01.	o ac	0.0.	J. O O																									
0	1	2	3	4	5	6	7	8	9	0	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
М	Α	R	I	Α		R	U	А	b	1		S	А	0	b	С	Α	R	L	0	S		X	X	X	X	X	X	X
30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
J	0	Α	0		R	U	Α	b	Α		R	Ι	0	b	С	L	Α	R	0		X	X	X	X	X	X	X	X	X
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89
Р	Е	D	R	0		R	U	Α	b	Χ	V		S	Α	0	b	С	Α	R	L	0	S		X	X	X	X	X	X
90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119
Α	N	Α		R	U	Α	b	Α	U	G	U	S	Τ	0	b	Р	Α	Ι	V	Α		Ι	В	Α	Т	Е		X	X
120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149

Exemplo – Arquivo original

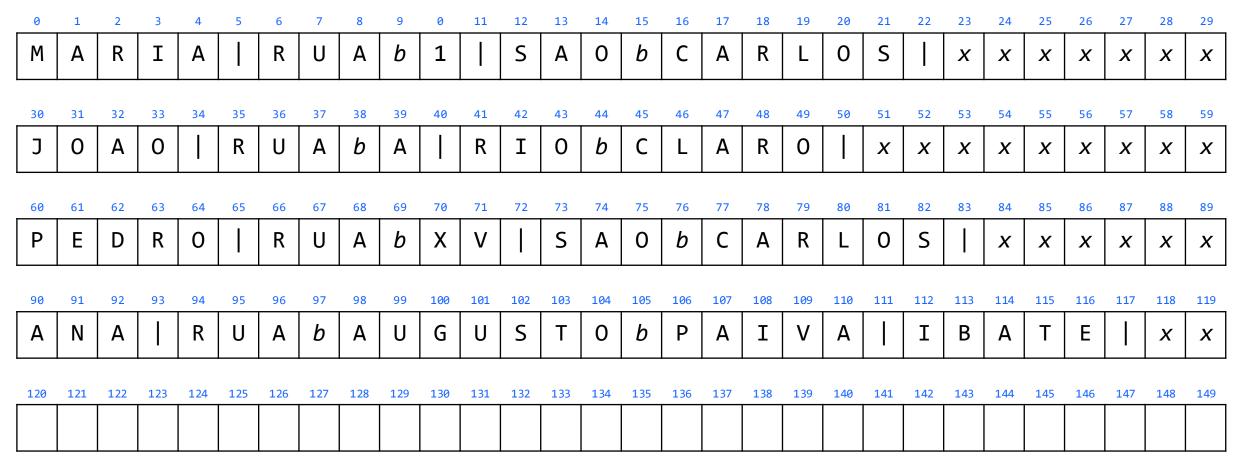
Registro de cabeçalho



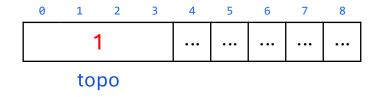
Remova o registro com

RRN = 1

Registro	de	da	dos
----------	----	----	-----



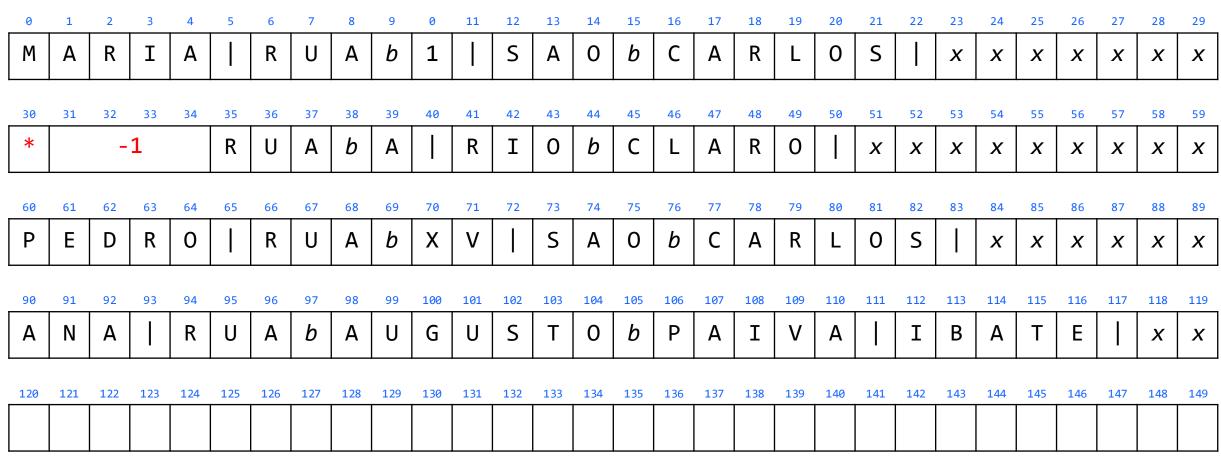
Registro de cabeçalho



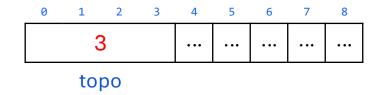
Remova o registro com

RRN = 3

Registro de	e dados
-------------	---------



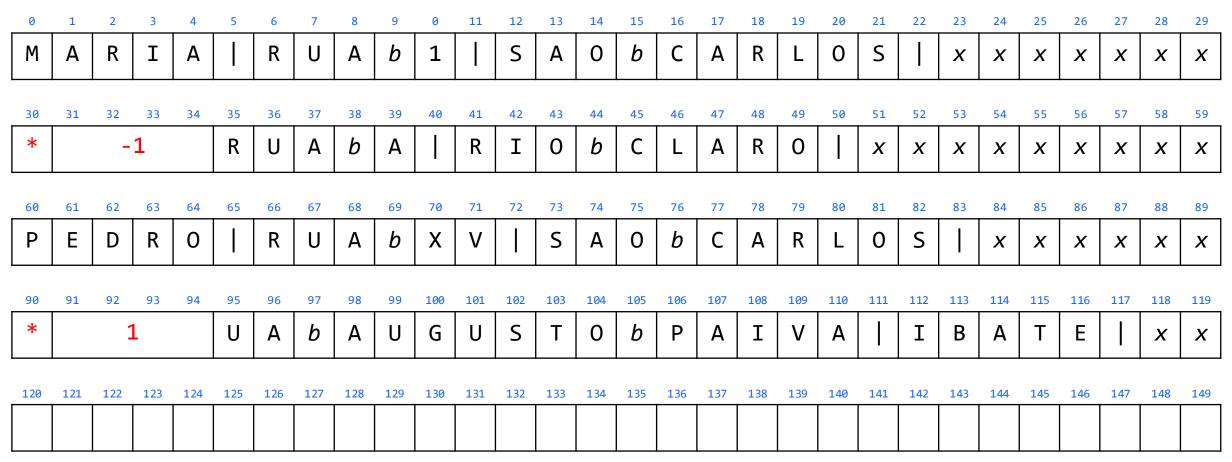
Registro de cabeçalho



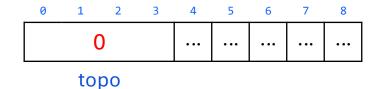
Remova o registro com

RRN = 0

Registro	de	da	dos
----------	----	----	-----

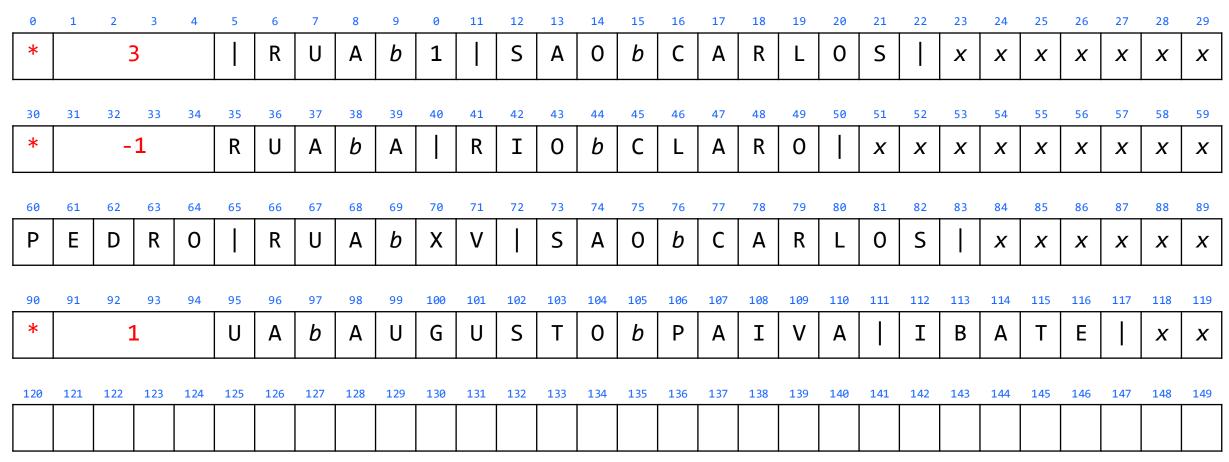


Registro de cabeçalho



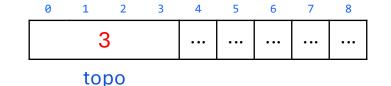
Insira novo registro:

CATARINA, AVENIDA X, RIO CLARO



Exemplo – Arquivo com registro inserido

Registro de cabeçalho



	5 .01.	o ac																											
0	1	2	3	4	5	6	7	8	9	0	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
С	Α	Т	Α	R	I	N	А		А	V	E	N	I	D	Α	b	X		R	I	0	b	С	L	А	R	0		X
30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
*		-	1		R	U	Α	b	Α		R	Ι	0	b	С	L	Α	R	0		X	X	X	X	X	X	X	X	X
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89
Р	Е	D	R	0		R	U	Α	b	X	V		S	Α	0	b	С	Α	R	L	0	S		X	X	X	X	X	X
90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119
*		1	L		J	А	b	Α	U	G	כ	S	Τ	0	b	Р	Α	Ι	>	А		Ι	В	А	Т	Ε		X	X
120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149

Como localizar espaços vazios?



Abordagem dinâmica: Registros de **tamanho variável**

Abordagem dinâmica: registros de tamanho variável

- Solução
 - Lista encadeada de registros eliminados
- Características
 - Lista: constitui-se dos byte offsetts dos registros marcados como logicamente removidos
 - Cabeça da lista: armazenada no registro de cabeçalho do arquivo
 - Dado adicional: necessário guardar também o tamanho do registro
 - Inserção e reuso de espaço: ??

Implementação: LISTA

Operações

Remoção

- Marca o registro como logicamente removido
- Insere o registro na lista de registros logicamente removidos (empilha ou insere ordenado na lista)

Inserção

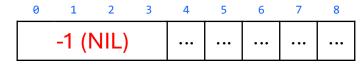
- Remove o registro da lista de registros logicamente removidos (de acordo com o tamanho solicitado)
- Insere os dados no espaço do registro desempilhado

Atualização

- Pode requerer remoção e depois inserção
 - Por que?

Exemplo – Arquivo original

Registro de cabeçalho

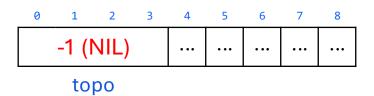


topo

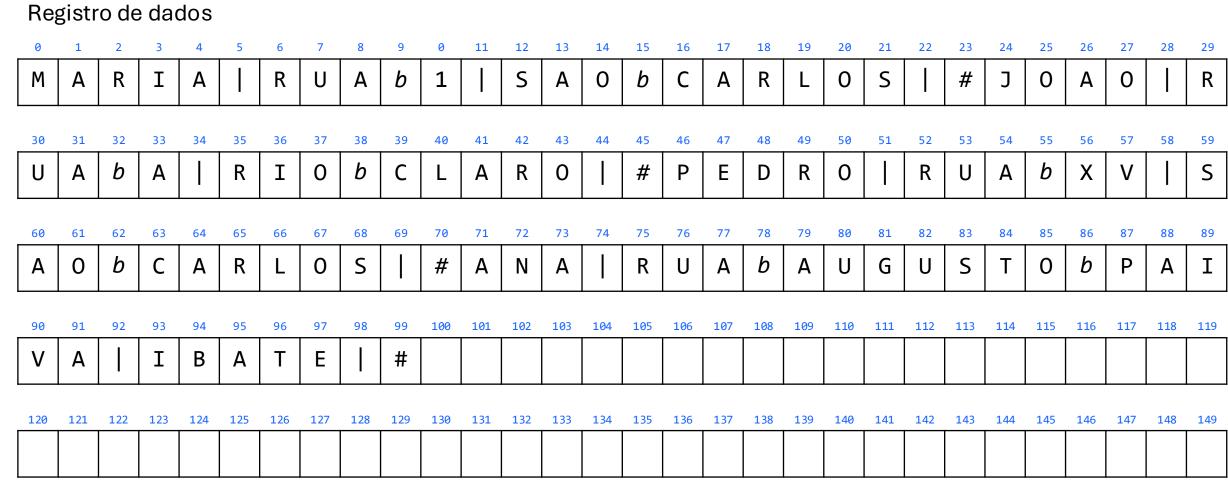
		0.0																											
0	1	2	3	4	5	6	7	8	9	0	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
М	Α	R	I	Α		R	U	А	b	1		S	Α	0	b	С	Α	R	L	0	S		#	J	0	Α	0		R
30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
U	Α	b	Α		R	I	0	b	С	L	А	R	0		#	Р	E	D	R	0		R	U	А	b	X	V		S
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	7 9	80	81	82	83	84	85	86	87	88	89
Α	0	b	С	Α	R	L	0	S		#	Α	N	Α		R	U	Α	b	Α	U	G	U	S	Т	0	b	Р	Α	I
90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119
V	Α		Ι	В	Α	Т	E		#																				
120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149

Exemplo – Arquivo original

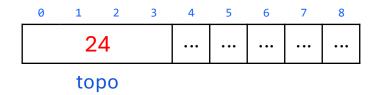
Registro de cabeçalho



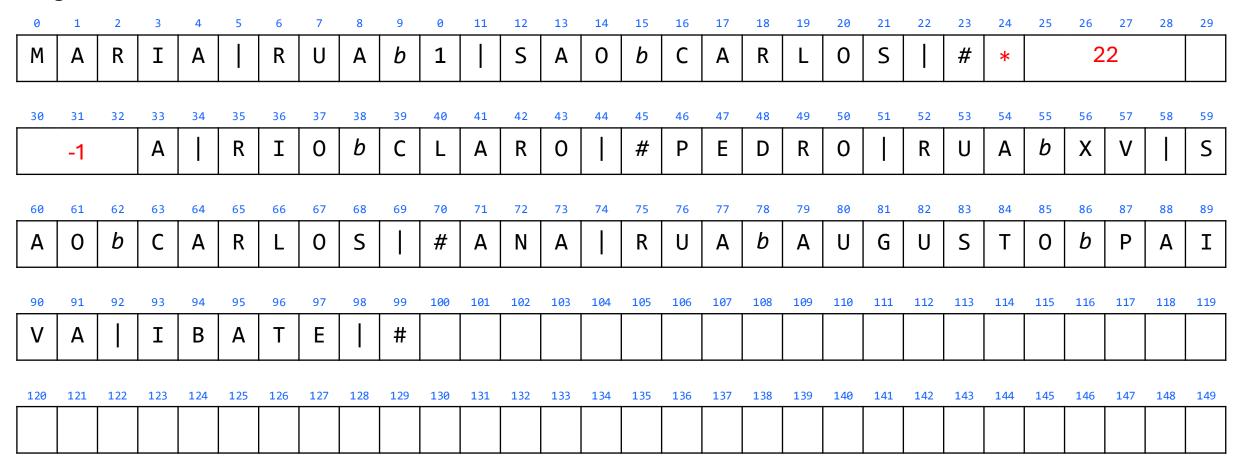
Remova os dados do segundo registro



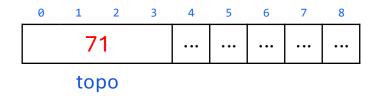
Registro de cabeçalho



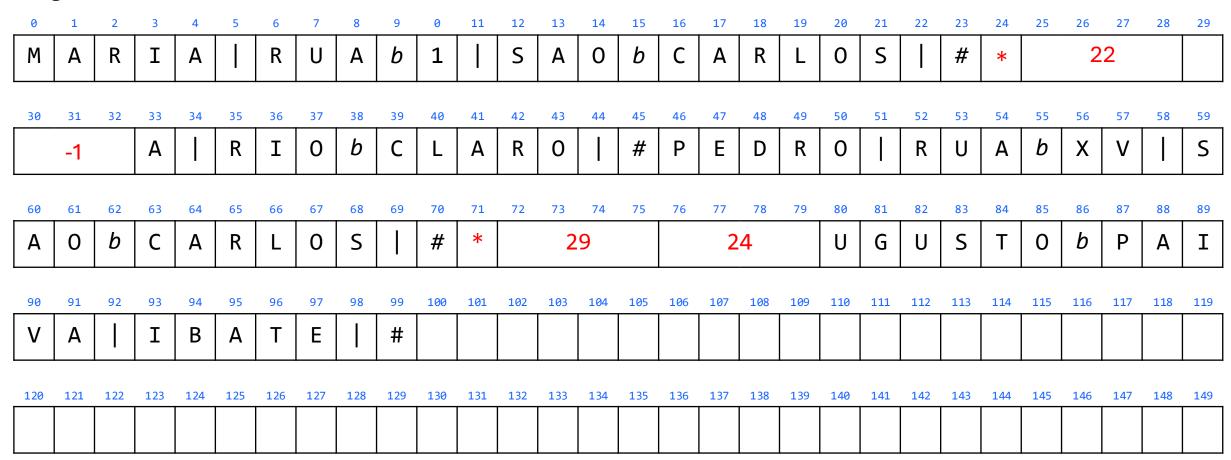
Remova os dados do quarto registro



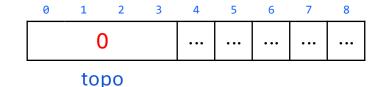
Registro de cabeçalho



Remova os dados do primeiro registro



Registro de cabeçalho



	gioti	0 0.0																											
0	1	2	3	4	5	6	7	8	9	0	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
*		2	4			7	1		b	1		S	А	0	b	С	Α	R	L	0	S		#	*		2	2		
30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
	-1		Α		R	I	0	b	С	L	Α	R	0		#	Р	E	D	R	0		R	U	Α	b	Х	V	1	S
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89
Α	0	b	С	Α	R	L	0	S		#	*		2	9			2	4		U	G	U	S	Τ	0	b	Р	Α	I
90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119
V	Α		I	В	Α	Т	E		#																				
120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149

Registros de tamanho variável: Reuso do espaço

Realiza uma busca sequencial na lista

Se encontrou espaço disponível no tamanho adequado

Então reaproveita o espaço para armazenar o novo registro, usando uma estratégia de alocação

Senão insere o novo registro no final do arquivo

O tamanho do registro que foi removido deve ser do tamanho adequado, ou seja, "grande o suficiente" para que os dados do novo registro usem aquele espaço

Estratégias de alocação

First-Fit

• Utiliza o primeiro espaço que servir

Best-Fit

• Escolhe o espaço mais justo possível

Worst-Fit

• Escolhe o maior espaço possível

Estratégias de alocação: Analogia com estacionamento de carros



First-Fit

O carro estaciona na primeira vaga que encontra, sem se preocupar se ela é grande ou pequena, desde que caiba.



Best-Fit

O carro estaciona na menor vaga disponível que ainda seja grande o suficiente para ele, ocupando o espaço mais justo.



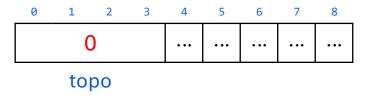
Worst-Fit

O carro estaciona na maior vaga disponível, mesmo que sobre muito espaço, para liberar as vagas menores para outros veículos.

Reuso de Espaço: Estratégia First-Fit

Registro de cabeçalho

Dariatus de de de s



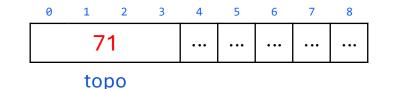
Insira novo registro:

RITA, RUA X, IBATE

		.	dat	dos																	0 S # * 22 50 51 52 53 54 55 56 57 58 59									
0	1	2	3	4	5	6	7	8	9	0	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	
*		2	4			7	1		b	1		S	Α	0	b	С	Α	R	L	0	S		#	*		2	2			
30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	
	-1		Α		R	Ι	0	b	С	L	Α	R	0		#	Р	E	D	R	0		R	U	Α	b	X	V		S	
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	7 9	80	81	82	83	84	85	86	87	88	89	
Α	0	b	С	Α	R	Ш	0	S		#	*		2	9			2	4		U	G	U	S	Т	0	b	Р	Α	I	
90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	
V	Α		Ι	В	Α	Τ	Ε		#																					
120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	

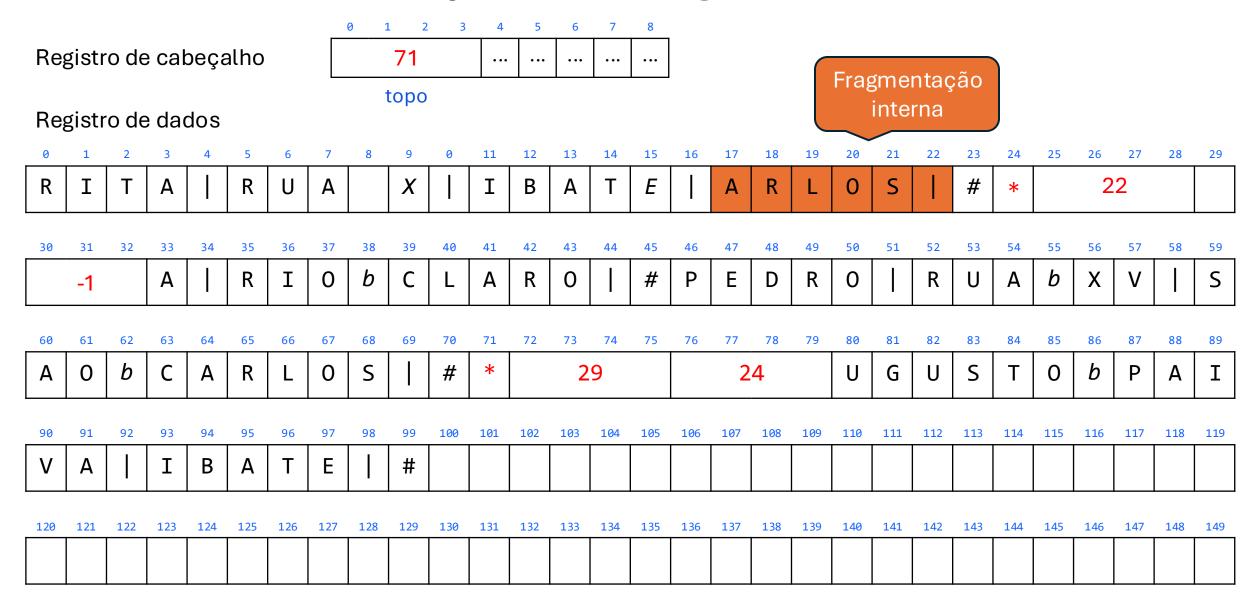
Reuso de Espaço: Estratégia First-Fit

Registro de cabeçalho



riogistro de dados																													
0	1	2	3	4	5	6	7	8	9	0	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
R	Ι	Т	Α		R	U	A				Ε		Α	R	L	0	S		# *			22							
30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
	-1		Α		R	Ι	0	b	С	L	Α	R	0		#	Р	E	D	R	0		R	U	Α	b	X	V		S
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89
Α	0	b	С	Α	R	L	0	S		#	*		29			24			U	G	U	S	Τ	0	b	Р	Α	I	
90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119
V	Α		I	В	Α	Т	E		#																				
120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149

Reuso de Espaço: Estratégia First-Fit



Reuso de Espaço: Estratégia Best-Fit

Insira novo registro: Registro de cabeçalho RITA, RUA X, IBATE topo Registro de dados (tamanho = 18)26 24 b S R b 24 0

Ε

129

130

131

132

135

136

137

139

28

Α

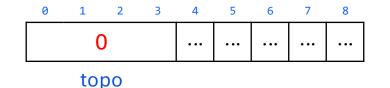
118

22

146

Reuso de Espaço: Estratégia Best-Fit

Registro de cabeçalho



riogioti o de dados																													
0	1	2	3	4	5	6	7	8	9	0	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
*	24					71				1		S	А	0	b	С	А	R	L	0	S		#	R	I	Т	А		R
30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
U	Α	b	X		I	В	Α	Т	Е		A	R	0	—	#	Р	E	D	R	0		R	J	Α	b	X	٧		S
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89
Α	0	b	С	Α	R	L	0	S		#	*		29			-1			U	G	U	S	Т	0	b	Р	Α	I	
90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119
V	А		I	В	Α	Т	Е		#																				
120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149