

Compactação e Reuso de espaço

SCC0607 – Estrutura de Dados III

Anderson Canale Garcia

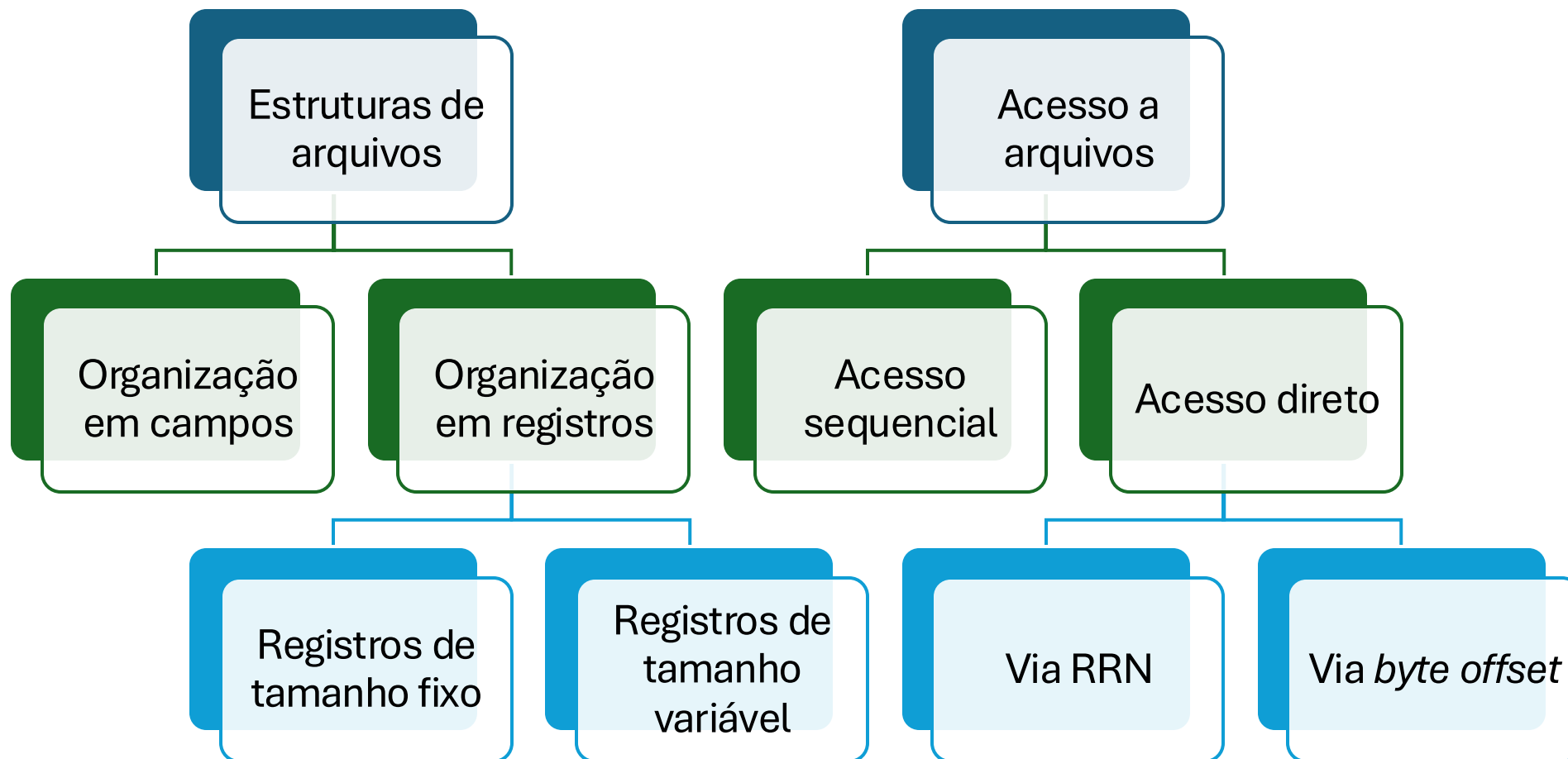
Baseado no material de:
Cristina D. Aguiar

Sumário

- Contexto
- Compressão (revisão)
- Compactação e reuso de espaço
 - Abordagem estática

O que estamos estudando?

Vimos...



Agora...

- Continuaremos falando sobre organização de arquivos
- Nova motivação: melhorar a performance!
 - Como organizar os arquivos para que sejam menores?
 - Compressão
 - Como recuperar espaço não utilizado?
 - Compactação
 - Reuso

Compressão de dados

- Por que comprimir?
 - Reduzir espaço utilizado => custos de armazenamento
 - Acelerar tempo de transmissão
 - Processar dados sequenciais mais rapidamente
- Como funciona?
 - Codificando a informação de forma mais eficiente
 - A variedade de técnicas é enorme

Compressão de dados – algumas técnicas

- Redução de redundância
- Omissão de sequências repetidas
- Códigos de tamanho variável

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

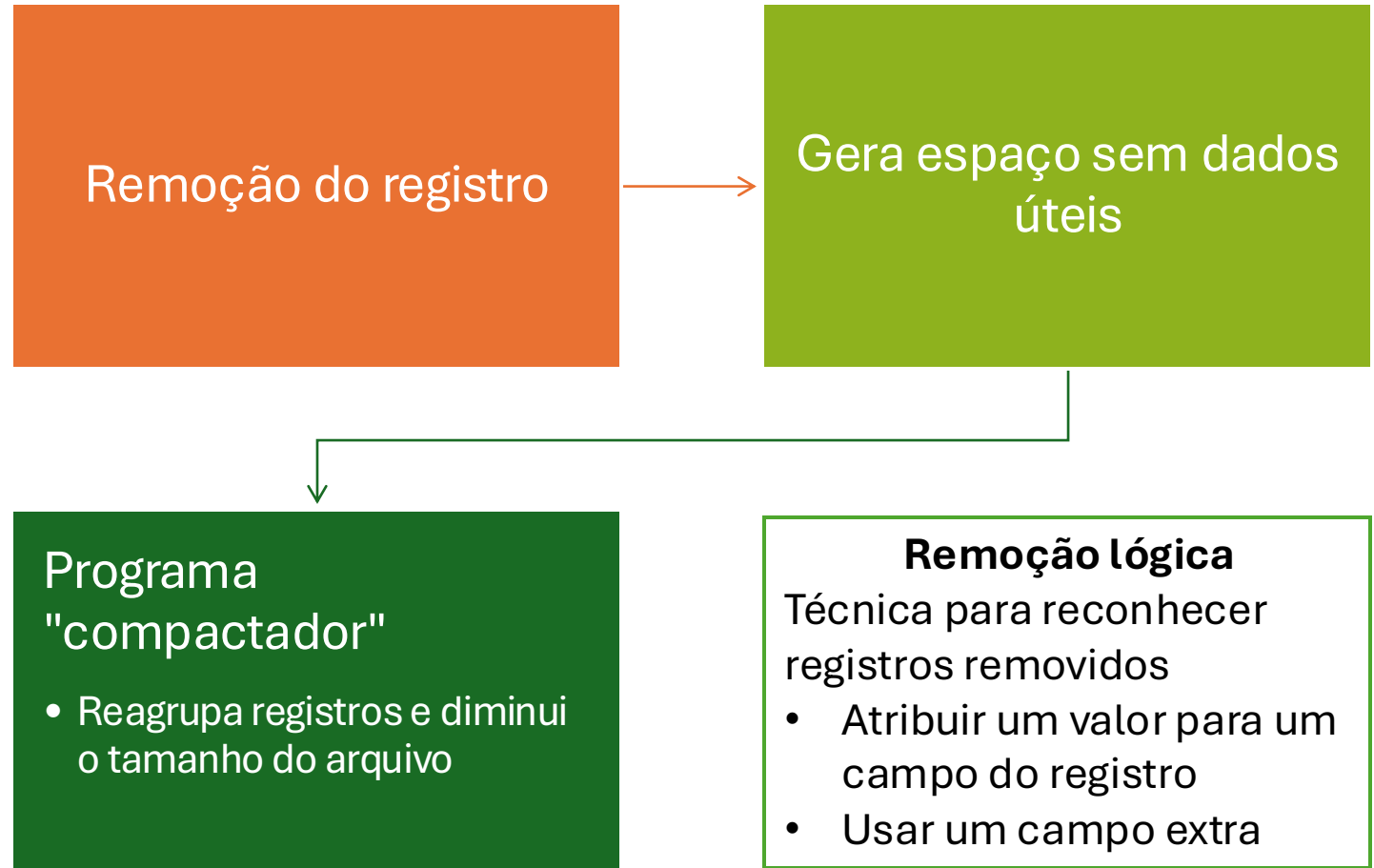
Quando um registro é removido, deve-se posteriormente reutilizar o espaço do registro

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 Δt (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
 - Campos de tamanho variável
 - Método: delimitador '|'
 - 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



Campos de
tamanho variável

Método:
delimitador '|'



Registros de
tamanho fixo

Tamanho do
registro: 30
bytes

Exemplo

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	A	R	I	A		R	U	A	b	1		S	A	O	b	C	A	R	L	O	S		x	x	x	x	x	x	x

J	O	A	O		R	U	A	b	A		R	I	O	b	C	L	A	R	O		x	x	x	x	x	x	x	x
---	---	---	---	--	---	---	---	---	---	--	---	---	---	---	---	---	---	---	---	--	---	---	---	---	---	---	---	---

P	E	D	R	O		R	U	A	<i>b</i>	X	V		S	A	O	<i>b</i>	C	A	R	L	O	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
A	N	A		R	U	A	<i>b</i>	A	U	G	U	S	T	O	<i>b</i>	P	A	I	V	A		I	B	A	T	E		x	x

[illegible]

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
M	A	R	I	A		R	U	A	b	1		S	A	O	b	C	A	R	L	O	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	O	A	O		R	U	A	b	A		R	I	O	b	C	L	A	R	O		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
P	E	D	R	O		R	U	A	b	X	V		S	A	O	b	C	A	R	L	O	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
A	N	A		R	U	A	b	A	U	G	U	S	T	O	b	P	A	I	V	A		I	B	A	T	E		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

1. Recupere os dados do registro com **RRN = 1**

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
M	A	R	I	A		R	U	A	b	1		S	A	O	b	C	A	R	L	O	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	O	A	O		R	U	A	b	A		R	I	O	b	C	L	A	R	O		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
P	E	D	R	O		R	U	A	b	X	V		S	A	O	b	C	A	R	L	O	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
A	N	A		R	U	A	b	A	U	G	U	S	T	O	b	P	A	I	V	A		I	B	A	T	E		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

2. Recupere os dados do registro com **RRN = 4**

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
M	A	R	I	A		R	U	A	b	1		S	A	O	b	C	A	R	L	O	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	O	A	O		R	U	A	b	A		R	I	O	b	C	L	A	R	O		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
P	E	D	R	O		R	U	A	b	X	V		S	A	O	b	C	A	R	L	O	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
A	N	A		R	U	A	b	A	U	G	U	S	T	O	b	P	A	I	V	A		I	B	A	T	E		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
M	A	R	I	A		R	U	A	b	1		S	A	O	b	C	A	R	L	O	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
*	O	A	O		R	U	A	b	A		R	I	O	b	C	L	A	R	O		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
P	E	D	R	O		R	U	A	b	X	V		S	A	O	b	C	A	R	L	O	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	A		R	U	A	b	A	U	G	U	S	T	O	b	P	A	I	V	A		I	B	A	T	E		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

4. Recupere os dados do registro com **RRN = 1**

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
M	A	R	I	A		R	U	A	b	1		S	A	O	b	C	A	R	L	O	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
*	O	A	O		R	U	A	b	A		R	I	O	b	C	L	A	R	O		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
P	E	D	R	O		R	U	A	b	X	V		S	A	O	b	C	A	R	L	O	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	A		R	U	A	b	A	U	G	U	S	T	O	b	P	A	I	V	A		I	B	A	T	E		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

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	A	R	I	A		R	U	A	b	1		S	A	O	b	C	A	R	L	O	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
*	O	A	O		R	U	A	b	A		R	I	O	b	C	L	A	R	O		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
P	E	D	R	O		R	U	A	b	X	V		S	A	O	b	C	A	R	L	O	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	A		R	U	A	b	A	U	G	U	S	T	O	b	P	A	I	V	A		I	B	A	T	E		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
C	A	T	A	R	I	N	A		A	V	E	N	I	D	A	b	X		R	I	O	b	C	L	A	R	O		x

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	A	R	I	A		R	U	A	b	1		S	A	O	b	C	A	R	L	O	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
P	E	D	R	O		R	U	A	b	X	V		S	A	O	b	C	A	R	L	O	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
C	A	T	A	R	I	N	A		A	V	E	N	I	D	A	b	X		R	I	O	b	C	L	A	R	O		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

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	A	R	I	A		R	U	A	b	1		S	A	O	b	C	A	R	L	O	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
P	E	D	R	O		R	U	A	b	X	V		S	A	O	b	C	A	R	L	O	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
C	A	T	A	R	I	N	A		A	V	E	N	I	D	A	b	X		R	I	O	b	C	L	A	R	O		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

6. Compacte o arquivo

Abordagem estática - características

- Técnica pode ser aplicada a
 - Registros de tamanho fixo
 - Registros de tamanho variável
- Frequência para se aplicar a técnica
 - Depende da aplicação
 - Depende da porcentagem de registros marcados como removidos