



Sdružení uživatelů Informix & DB2

Adaptive Linear Index Cleaning (ALICE) režim čištění indexových stránek

Ing. Jan Musil, IBM ČR
Community of Practice for
CEEMEA

Agenda

- Co je to B-Tree Scanner ?
- Jak B-Tree Scannery pracují ?
- Vlastnosti B-Tree Scannerů
- Metody prohledávání indexových stránek
- Konfigurace B-Tree Scannerů
- Výpočet ALICE módu
- Porovnání ALICE módů
- Dynamické nastavení ALICE módu
- Monitorování
- Dynamická konfigurace B-Tree Scannerů

Co je to B-Tree Scanner ?

- Zajišťuje čištění indexových stránek
- Provádí kompresi listových stránek B-Tree
- Nahrazuje funkcionalitu *btcleaner* od verzí IDS 7.31.UD8, 9.21.UC7, 9.30.UC6 a vyšších
- Důraz na efektivnost, flexibilitu a jednoduchost administrace

Jak B-Tree Scannery pracují ?

- Po smazání záznamu v tabulce se neodstraňuje indexový záznam, ale označí se pomocí Delete Flagu (df)
- Asynchronní mazání záznamů v indexových stránkách (čištění indexových stránek) provádí B-Tree Scannery

0000 05f4	0002 14dd	0004 0090	0045 07a7
0000 026c	0000 0000	key value	rowid df
key value	rowid	df	key value rowid df
key value	rowid	df	rowid df
			0036 000f
002c 000a	0022 000a	0018 000s	3275 4a3f

- Indexový záznam je uvolněn pro mazání až po ukončení transakce (*committed deleted item*)

Vlastnosti B-Tree Scannerů

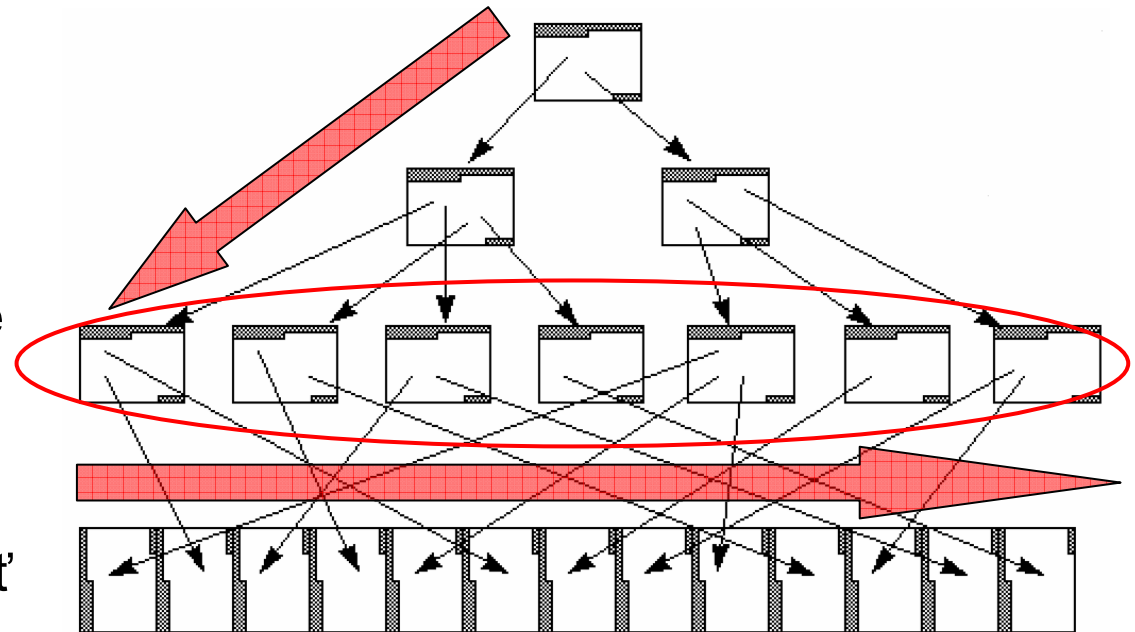
- Funkcionalita je implementovaná prostřednictvím *btscanner* vláken
- Btscanner vlákna mohou být vytvářena a konfigurována dynamicky
- Měřítko pro čištění indexu je počet přečtených *committed deleted items*, nebo-li míra nadbytečných (neužitečných) čtení provedených databázovým serverem
- Počet těchto nadbytečných čtení je vyjádřena hodnotou *hit-count (dirty hits)*
- Pokud hodnota *hit-count* překročí stanovenou mez, je index umístěn do tzv. hot-list seznamu pro čištění
- Btscanner vybere index z hot-listu a prohledává listové stránky za účelem jejich čištění
- Rozsah prohledávané oblasti (počet listových stránek) je určen metodou prohledávání
- Btscanner odstraní *committed deleted items* z prohledávané oblasti listových stránek a v případě potřeby tyto stránky sloučí

Metody prohledávání indexových stránek

- leaf
- range
- ALICE

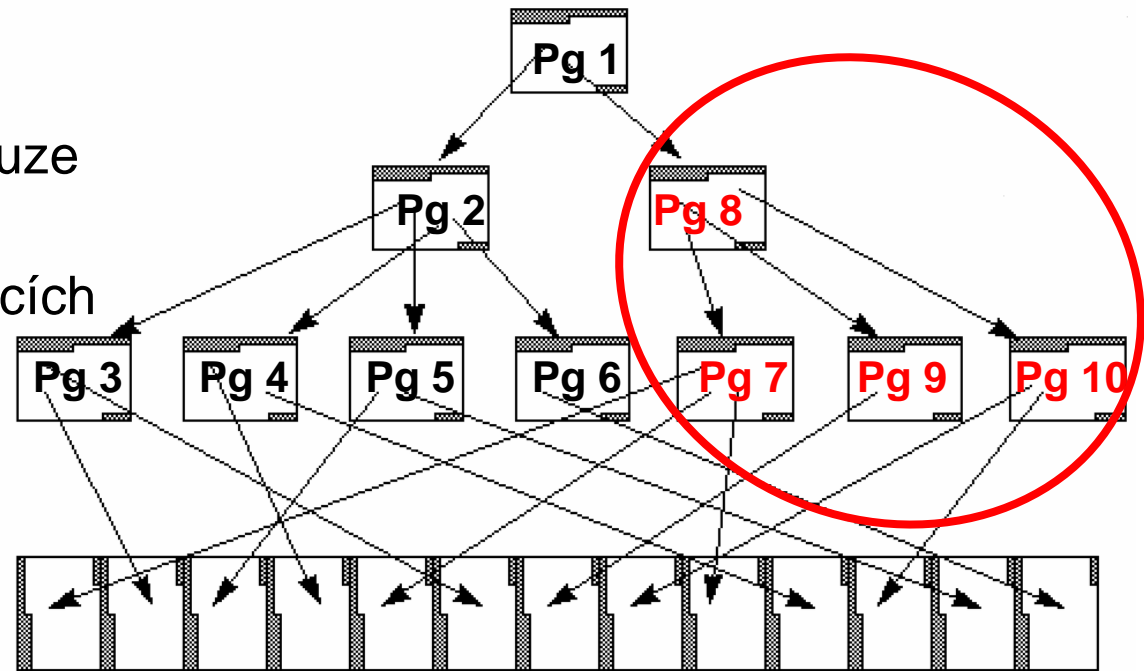
Metoda prohledávání: leaf

- Výhody
 - Prohledává všechny indexové uzly
 - Velmi efektivní pro malé indexy
 - Čistí jak attached, tak detached indexy
 - Používá sdílenou paměť
- Nevýhody
 - Každý uzel musí načíst do sdílené paměti
 - Pomalé vzhledem k velkému počtu I/O operací



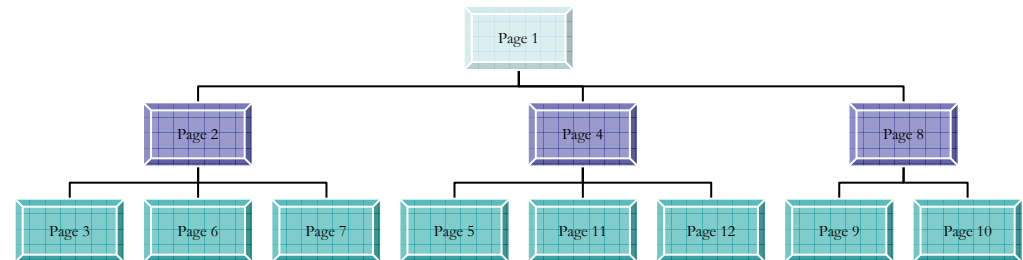
Metoda prohledávání: range

- Výhody
 - Prohledává stránky pouze ve stanovené oblasti
 - Používá Light I/O v blocích po 128 stránkách
- Nevýhody
 - Nečistí uzly, které nebyly zapsány ze sdílené paměti na disk
 - Nelze použít u tabulek s více attached indexy

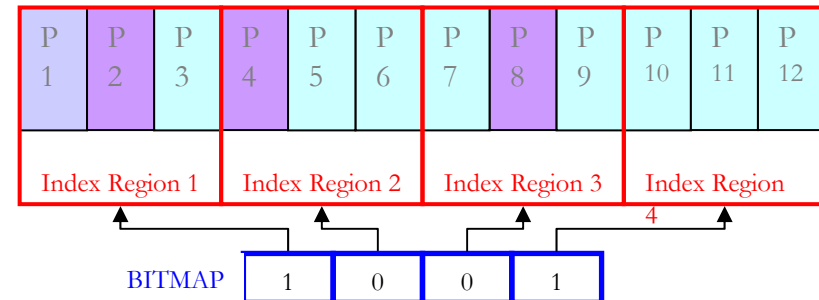


Metoda prohledávání: ALICE

- ALICE=Adaptive Linear Index Cleaning
- IDS 10.00.xC5 a vyšší
- Indexové stránky rozdělené do stejně velkých oblastí
- Každá oblast je označena bitovou hodnotou 0 (není třeba čistit) nebo 1 (je třeba čistit)
- Linearizace je provedena podle logických adres stránek na disku, každá oblast (region) je fyzicky spojitá
- Za účelem čištění se prohledávají pouze regiony s bitmap hodnotou 1



Lineární reprezentace indexu



- 64x účinnější než range scan
- Na rozdíl od range scanu, má ALICE samoladící schopnost

Konfigurace B-Tree Scannerů

```
BTSCANNER [num=scanner_threads] [,threshold=dirty_hits] [,rangesize=size]  
          [,alice=mode] [,compression=level]
```

- num
 - Počet bt scanner vláken (default=1)
- threshold
 - Hranice počtu hit-count, při kterém je index umístěn do hot listu (default=5000)
- rangesize
 - Velikost indexu (fragmentu) v kB, od které se použije range scan (default= -1 : OFF)
- alice
 - Určuje velikost počáteční velikost bitmapy pro sledování *committed delete items*
 - Hodnota od 1 – 12 (default=0 : OFF)
 - Optimální hodnota pro malé a střední systémy s velikostí indexů do 1 GB je 6 nebo 7
- compression
 - Úroveň, při které se sloučí dvě částečně zaplněné indexové stránky
 - Hodnoty: *low*, *med* (medium), *high*, *default*
 - Default: med

Výpočet ALICE módu

- 1 light I/O scan = 256 pg (velikost stránky 2kB)
nebo 128 pg (velikost stránky 4kB)
 - 512 kB
- Počet light I/O scanů potřebných pro přečtení jednoho regionu při $alice=N$ je $2^{(12-N)}$
 - Při hodnotě $alice=6$ reprezentuje 1 bit velikost regionu: $2^{(12-6)} * 512 \text{ kB} = 64 * 512 \text{ kB} = 32 \text{ MB}$
 - Při $alice=6$ reprezentuje 1 B paměti: $32 \text{ MB} * 8 = 256 \text{ MB}$, tedy 131 072 (128k) 2k-stránek indexu

Porovnání ALICE módů

ALICE mód	Počet 2k stránek mapovaných 1 bajtem paměti	Počet light I/O čtení na region	Velikost bitmapy pro index 1 GB	Počet regionů pro index 1 GB
6	131 072	512	4 B	32
10	8 192	32	64 B	512
12	2 048	8	256 B	2048

- Vyšší ALICE mód = větší oblast v paměti pro bitmapu = méně light I/O
- Vyšší ALICE mód = více regionů = méně indexových stránek v regionu = efektivnější čištění

Dynamické nastavení ALICE módu

- Účinnost I/O = počet I/O / počet I/O bez nalezení *committed deleted items*
- Pokud je účinnost I/O menší než 75 % (vyhodnocuje se vždy po každých 5 čištění), pak se ALICE mód zvyšuje o jednu
 - Dvojnásobek paměti
 - Poloviční počet light I/O
 - Poloviční počet indexových stránek v regionu
- Pro hodnoty ALICE=1 (8 B paměti) a ALICE=2 (16 B paměti) se dynamické nastavení neprovádí !

\$ onstat -C alice

```
IBM Informix Dynamic Server Version 11.50.FC1 -- On-Line -- Up 10 days 02:24:24 -- 48080
Kbytes
```

```
Btree Cleaner Info
```

```
ALICE Cleaning Statistics
```

```
=====
```

```
System ALICE Info: Mode = 6, Eff = 30 %, Adj = 5
```

Partnum	Mode	BM_Sz	Used_Pg	Examined	Dirty_Pg	# I/O	Found	Eff	Adj
0x00100013	6	64	97	0	0	0	0	0.0 %	0
0x001001cd	0	0	26	0	0	0	0	0.0 %	0
0x00a00011	6	64	91	0	0	0	0	0.0 %	0
0x00a00013	6	64	25	24	3	3	1	33.3 %	1
0x00a00019	6	64	470	470	3	3	2	66.7 %	1
0x00a00022	6	64	26	0	0	0	0	0.0 %	0
0x00a00022	6	64	26	0	0	0	0	0.0 %	0

\$ onstat -C map

```
IBM Informix Dynamic Server Version 11.50.FC1 -- On-Line -- Up 10 days 02:25:05 -- 48080
Kbytes
```

```
Btree Cleaner Info
```

```
ALICE Bitmap of Deleted Index Items
```

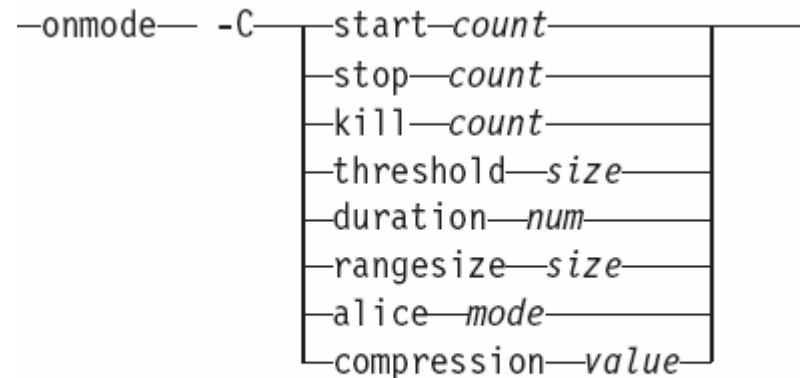
```
=====
```

```
Partnum Key Map
```

```
0x00100013 2 0000: 80000000 00000000
0x0010008b 3 0000: 80000000 00000000
0x001000c7 1 0000: 80000000 00000000
0x00100150 2 0000: 80000000 00000000
```

Dynamická konfigurace B-Tree Scannerů

- *start count*
 - Přidání dalších btscanner vláken (max. 128 v jednom příkaze)
- *stop|kill count*
 - Zastavení btscanner vláken
- *threshold size*
 - Stanovení hranice pro umístění indexu do hot list
- *duration num*
 - Stanovení doby platnosti hot list seznamu v sec
- *rangesize size*
 - Stanovení velikosti indexu pro range scan



- *alice mode*
 - Stanovení ALICE módu (0-12)
- *compression value*
 - Stanovení úrovně, kdy se budou dvě částečně zaplněné indexové stránky slučovat
 - High, med (medium), low, default

Děkuji za pozornost

jan_musil@cz.ibm.com
