

Alena Martinková, MAR0702

DAIS 2

# **Úvod**

Pre stĺpcové uloženie a komprimáciu som si vybrala 4 dotazy. Pričom dotazy **s5, t7** a **t3** sú dotazy spomenuté v predchádzajúcich úlohách a dotaz **s6** je nový dotaz. Všetky dotazy sú priložené v sql súbore. Operácie **s5** a **t7** sú vhodné na stĺpcové uloženie, pretože obidve obsahujú v *selecte* len agregačné funkcie. Dotazy **t3** a **s6** sú nevhodné – **t3** selectuje všetky atribúty, popr. **s6** atribúty z viacerých tabuliek.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Operácia** | **Riadková tabuľka [ms]** | | **Stĺpcová tabuľka [ms]** | |
|  | **Bez komprimácie** | **S komprimáciou** | **Bez komprimácie** | **S komprimáciou** |
| **s5** | 50ms | 19ms | 21ms | 11ms |
| **t7** | 49ms | 10ms | 15ms | 10ms |
| **t3** | 16ms | 19ms | 32ms | 22ms |
| **s6** | 83ms | 22ms | 110ms | 30ms |

Operácie boli vykonávané nad tabuľkami Statistic a Team.

U vhodných pre stĺpcové uloženie došlo k zlepšeniu: 2,3x a 3,2x, takže môžeme vidieť, že stĺpcové uloženie je v týchto prípadoch výhodné. Dotazy nevhodné pre stĺpcové uloženie sa zhrošili 1,3x až 2x. Po komprimácii sa čas vykonávania operácii taktiež zlepšil, v niektorých prípadoch (t7) sa dokonca vyrovnal čas operácie v riadkovej tabuľke, času operácie v stĺpcovej tabuľke. Výsledok by som zhodnotila kladne – dotazy, ktoré mali byť v stĺpcovej tabuľke rýchlejšie naozaj rýchlejšie boli, a naopak.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tabuľka** | **Riadková tabuľka** | | **Stĺpcová tabuľka** | |
|  | Bez komprimácie | S komprimáciou | Bez komprimácie  (s default komprimáciou) | S komprimáciou |
| **Team** | 660 kB | 450 kB | 250 kB | 180 kB |
| 84 blokov | 58 blokov | 32 blokov | 23 blokov |
| **Statistic** | 4690 kB | 3380 kB | 5180 kB | 2660 kB |
| 600 blokov | 440 blokov | 663 blokov | 341 blokov |
| **Spolu** | 5350 kB | 3830 kB | 5430 kB | 2840 kB |
| 684 blokov | 498 blokov | 695 blokov | 364 blokov |

Čo sa týka kompresie, stĺpcové tabuľky boli vytvorené (SQL server) už defaultne COLUMNSTORE, skúsila som aj COLUMNSTORE\_ARCHIVE. Zlepšenie bolo nasledovné: 1,38x v stĺpcovej tabuľke Team a **1,94x** v stĺpcovej tabuľke Statistic.

Na kompresiu riadkovej tabuľky bola použitá pre tabuľku Team kompresia: DATA\_COMPRESSION = ROW tzn. kódovanie premennej dĺžky a pre tabuľku Statistic sa ukázala lepšia kompresia: DATA\_COMPRESSION = PAGE (zhruba o 150 kB) tzn. riadková, prefixová a slovníková komprimácia. U riadkových tabuliek došlo k zlepšeniu, konkrétne tabuľka Team: 1,3x a tabuľka Statistic: 1,38x.

Komprimáciu som použila taktiež na indexy v daných tabuľkách, veľkosť je zarátaná vo výsledku. Najväčší rozdiel, čo sa týka komprimácie indexov bol v stĺpcových tabuľkách – tam práve komprimácia indexu urobila najviac.