

BADANIE WYDAJNOŚCI ZŁĄCZEŃ I ZAGNIEŹDZEŃ DLA SCHEMATÓW ZNORMALIZOWANYCH I ZDENORMALIZOWANYCH

Konfiguracja sprzętowa i programowa, na której przeprowadzono badanie:

CPU: Intel Core i7 3610QM 2.30 GHz

RAM: 8GB

System: 64-bitowy Windows 10

Karta graficzna: GeForce 635M 2GB

HDD: 750GB, 7200 RPM

System zarządzania bazami danych:

SQL Server 2019 15.0.18369.0

Opis analizy:

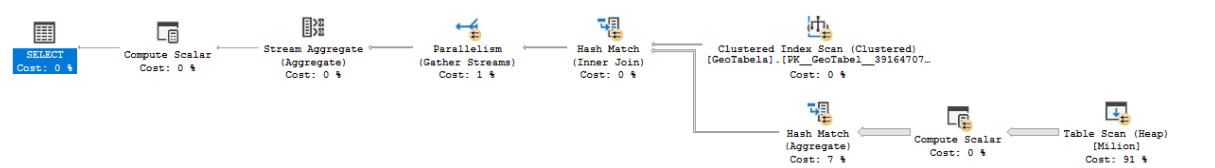
Badanie miało na celu przeprowadzenie analizy wydajności złączeń tabel z indeksami oraz bez indeksów. Tabele wykorzystane do badania są tabelami znormalizowanymi i zdenormalizowanymi. Tworząc tabele wzorowano się na budowie tabeli geochronologicznej, która jest schematem obrazującym przebieg historii Ziemi na podstawie następstwa procesów geologicznych i układu warstw skalnych. Wzorując się na wymienionej wcześniej tabeli przygotowano dwie wersje: pierwszą wersją jest tabela w wersji znormalizowanej (schemat płatka śniegu), czyli jest zaprojektowana zgodnie z modelem relacyjnej bazy danych, drugą wersją jest tabela w wersji zdenormalizowanej (schemat gwiazdy), której główną cechą jest centralna tabela, nazywana tabelą faktów, z którą połączone są tabele wymiarów.

Następnym krokiem było stworzenie tabeli Dziesięć i bazującej na niej tabeli Milion, wypełnionej kolejnymi liczbami naturalnymi od 0 do 999 999. Tabele te były niezbędne do przeprowadzenia badania.

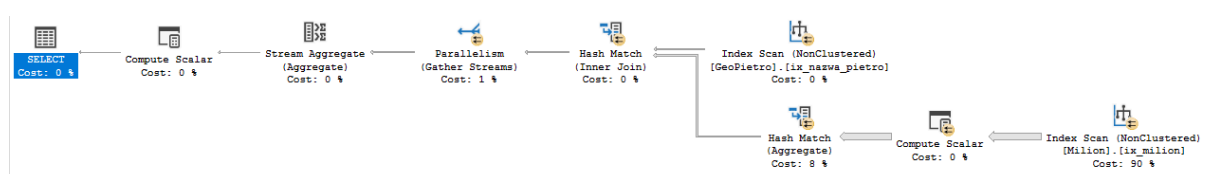
Do wykonania analizy przeprowadzono odpowiednie testy, które polegały na użyciu czterech zapytań. Testy zostały przeprowadzone w dwóch wariantach: pierwszym wariantem było użycie zapytań bez nałożonych indeksów na kolumny danych, a drugim wariantem było użycie zapytań z nałożonymi indeksami. Zapytanie 1 miało na celu złączenie syntetycznej tablicy z milionem wyników z tabelą geochronologiczną w postaci zdenormalizowanej z warunkiem dopasowującym zakresy wartości złączanych kolumn, zapytanie 2 miało na celu złączenie syntetycznej tablicy miliona wyników z tabelą geochronologiczną w postaci znormalizowanej, przy czym tabela geochronologiczna jest reprezentowana poprzez złączenie pięciu tabel, celem zapytania 3 było złączenie syntetycznej tablicy miliona wyników z tabelą geochronologiczną w postaci zdenormalizowanej, przy złączeniu poprzez zagnieżdżenie skorelowane, natomiast zapytanie 4 miało na celu złączenie syntetycznej tablicy miliona wyników z tabelą geochronologiczną w postaci znormalizowanej, przy złączeniu wykonywanym poprzez zagnieżdżenie skorelowane.

Wyniki analizy bez indeksów:

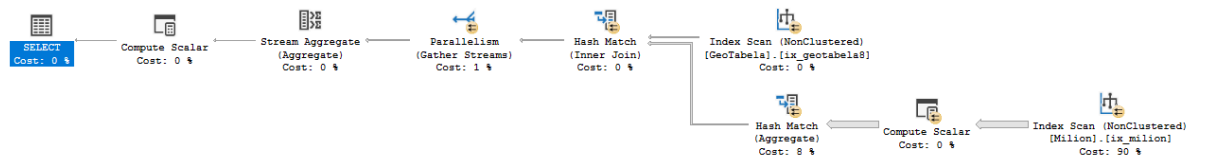
Zapytanie 1 (1 ZL):



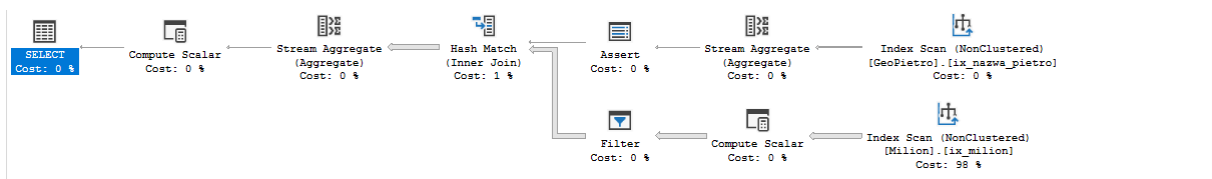
Zapytanie 2 (2 ZL):



Zapytanie 3 (3 ZG):

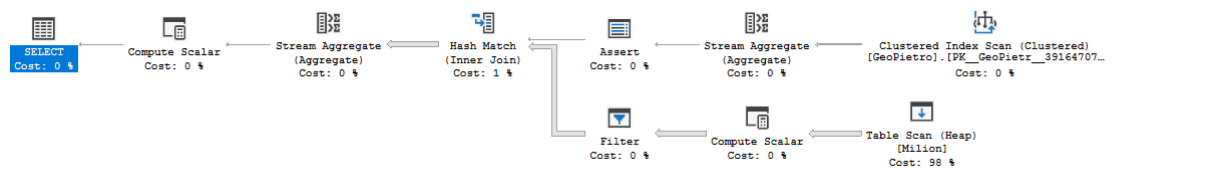


Zapytanie 4 (4 ZG):

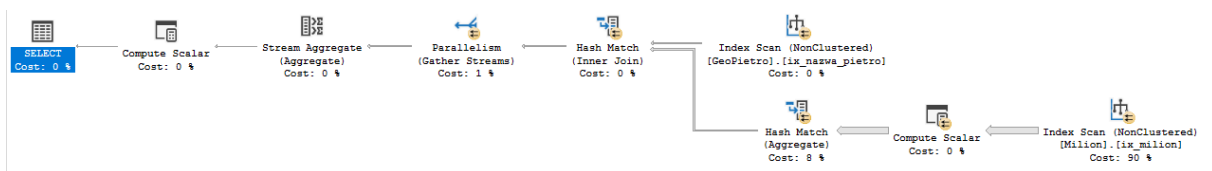


Wyniki analizy z indeksami:

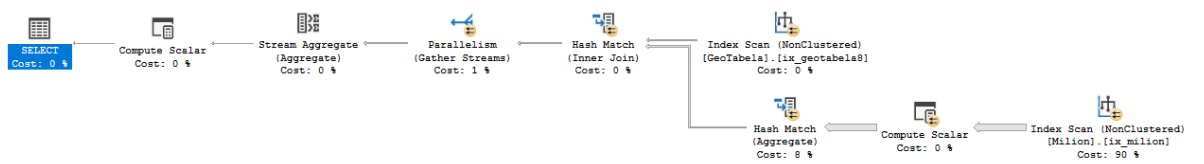
Zapytanie 1 (1 ZL):



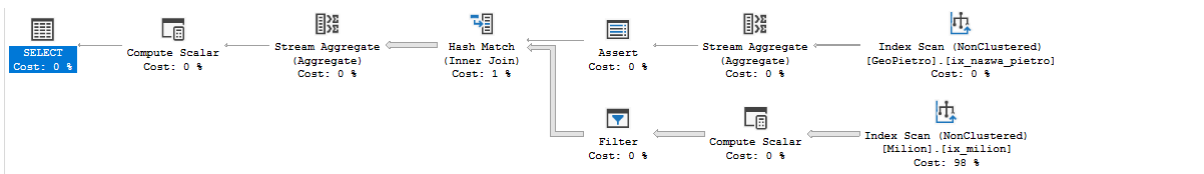
Zapytanie 2 (2 ZL):



Zapytanie 3 (3 ZG):



Zapytanie 4 (4 ZG):



Czasy wykonania zapytań:

	1 ZL		2 ZL		3 ZG		4 ZG	
Bez indeksów	MIN	ŚR	MIN	ŚR	MIN	ŚR	MIN	ŚR
Czas [ms]	92	110	103	114	98	118	109	124
Z indeksami								
Czas [ms]	74	95	92	104	92	118	95	115

Wyniki analizy i wnioski przeprowadzonego badania:

Po przeprowadzonej analizie można zauważyć, że indeksacja wpłynęła pozytywnie na wydajność wykonywania poszczególnych zapytań. Różnica pomiędzy postacią znormalizowaną, a zdenormalizowaną jest niewielka. Z przeprowadzonego badania można wywnioskować, że operacje działające na postaci zdenormalizowanej są szybsze od operacji bazujących na postaci znormalizowanej. Mimo to, postać znormalizowana ma istotną zaletę nad postacią zdenormalizowaną w postaci porządku i przejrzystości informacji jaki ona wprowadza.

Bibliografia:

WYDAJNOŚĆ ZŁĄCZEŃ I ZAGNIEŹDŹEŃ DLA SCHEMATÓW

ZNORMALIZOWANYCH I ZDENORMALIZOWANYCH- Łukasz JAJEŚNICA, Adam PIÓRKOWSKI