**WYDAJNOŚĆ ZŁĄCZEŃ I ZAGNIEŻDŻEŃ DLA SCHEMATÓW ZNORMALIZOWANYCH I ZDENORMALIZOWANYCH**

Wykonanie: Kacper Sobczyk

Celem badania jest sprawdzenie czy forma znormalizowana jest szybsza czy wolniejsza od formy zdenormalizowanej i jaki wpływ na prędkość wykonania zadania mają zapytania zagnieżdżone.

1. Do wykonania badania stworzono tabelę stratygraficzną w postaci znormalizowanej.

Stworzono następujące tabele: GeoEon, GeoEra, GeoOkres, GeoEpoka oraz GeoPietro, tak, że tabele łączyły się ze sobą za pomocą kluczy obcych.



Tabela GeoEon

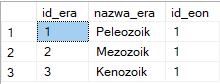


Tabela GeoEra

Obraz zawierający tekst, biały

Opis wygenerowany automatycznie

Tabela GeoOkres

Obraz zawierający stół

Opis wygenerowany automatycznie

Tabela GeoEpoka

Obraz zawierający stół

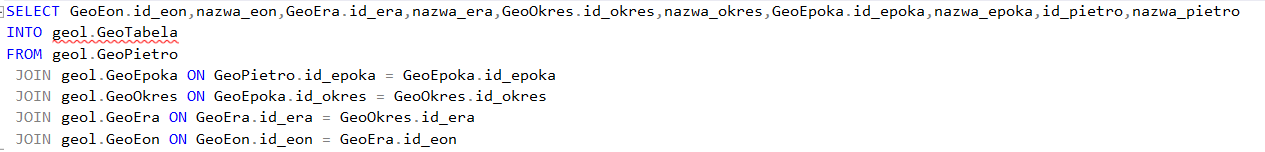
Opis wygenerowany automatycznieObraz zawierający stół

Opis wygenerowany automatycznie

Tabela GeoPietro

1. Ze schematu znormalizowanego powstał schemat zdenormalizowany

(schemat gwiazdy), który powstał za pomocą połączenia wszystkich tabel za pomocą komendy:



1. Do wykonania badania zostały wykonane dodatkowe dwie tabele: tabela Dziesiec

i Milion, która powstała na podstawie auto złączenia z tabelą Dziesiec. Tabela Dziesiec zawierała wszystkie cyfry oraz ilość bitów jaka była potrzebna do zapisu tej cyfry.

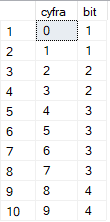
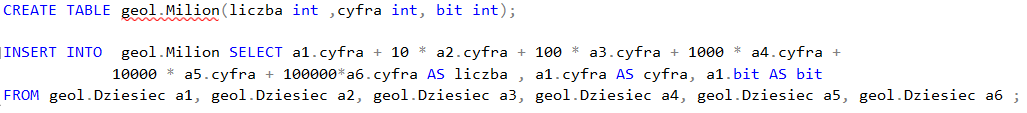


Tabela Milion posiadała liczby od 0 do 999999 i powstała za pomocą komendy:



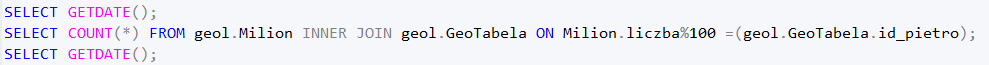
1. Kolejnym krokiem w badaniu było wykonywanie pomiarów przed i po wykonaniu obliczenia.

Pomiary zostały wykonane na komputerze o następujących parametrach:

* Procesor: Intel(R) Core (TM) i7-8550U CPU @ 1.80GHz 1.99 GHz
* Pamięć RAM: 8GB DDR4
* Dysk twardy: SSD 256 GB
* System operacyjny: Windows 10
* Używane środowisko: SQL Server v18.8

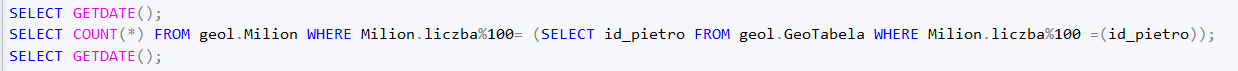
Każdy pomiar został wykonany 10 razy – z pomiarów wyciągnięto średnią oraz minimum.

Zostały wykonane pomiary dla następujących zapytań:

*  Zapytanie 1 (1 ZL), którego celem jest złączenie syntetycznej tablicy miliona wyników z tabelą geochronologiczną w postaci zdenormalizowanej, przy czym do warunku złączenia dodano operację modulo, dopasowującą zakresy wartości złączanych kolumn.
* Zapytanie 2 (2 ZL), którego celem jest złączenie syntetycznej tablicy miliona wyników z tabelą geochronologiczną w postaci znormalizowanej, reprezentowaną przez złączenia pięciu tabel.

Obraz zawierający tekst

Opis wygenerowany automatycznie

* Zapytanie 3 (3 ZG), którego celem jest złączenie syntetycznej tablicy miliona wyników z tabelą geochronologiczną w postaci zdenormalizowanej, przy czym złączenie jest wykonywane poprzez zagnieżdżenie skorelowane.
* Zapytanie 4 (4 ZG), którego celem jest złączenie syntetycznej tablicy miliona wyników z tabelą geochronologiczną w postaci znormalizowanej, przy czym złączenie jest wykonywane poprzez zagnieżdżenie skorelowane, a zapytanie wewnętrzne jest złączeniem tabel poszczególnych jednostek geochronologicznych.

Obraz zawierający tekst

Opis wygenerowany automatycznie

Po wykonaniu obliczeń zostały naniesione indeksy na klucze główne tabel, w tabeli nieskorelowanej GeoTabela na wszystkie kolumny z id oraz na liczby z tabeli Milion.

Obraz zawierający tekst

Opis wygenerowany automatycznie

1. Z wykonanych obliczeń otrzymano następujące wyniki:
2. Wnioski

Ze stworzonych wyników można wysunąć następujące wnioski:

* Postać zdenormalizowana jest minimalnie wolniejsza od postaci znormalizowanej dla połączeń (Badanie ZL1 i ZL2 dla NO IND). Postać zdenormalizowana jest szybsza, jeśli wykorzystujemy indeksowanie kolumn (Porównanie ZL 1 i ZL2 dla IND).
* Postać zdenormalizowana jest minimalnie szybsza od postaci znormalizowanej dla zagnieżdżeń skorelowanych zarówno przed indeksowaniem jak i po indeksowaniu kolumn (Badanie ZG 3 i ZG 4 dla IND i NO IND).
* Dla każdej postaci indeksowanie kolumn przyspiesza wykonywanie zapytań.