

Kraków, 25.05.2023r.



Bazy Danych

**Ćwiczenie 9 – Badanie wydajności złączeń i zagnieżdżeń dla schematów
znormalizowanych i zdenormalizowanych**

Weronika Chudzińska, nr albumu: 412356

Wstęp

Celem ćwiczenia było przeprowadzenie szeregu testów mających na celu ukazać wydajność złączeń oraz zapytań zagnieżdżonych. Na potrzeby testów stworzona została baza danych zawierająca informacje niezbędne do przygotowania tabeli geochronologicznej.

Specyfikacja techniczna komputera

- Procesor: Intel® Core(TM) i5-1035G1 CPU @ 1.00GHz 1.19 GHz
- RAM: 8,00 GB
- System: Windows 11
- SSD: WDC PC SN520 SDAPNUW-512G-1014
- GPU: Intel® UHD Graphics

Jako system zarządzania bazami danych wybrane zostało oprogramowanie PostgreSQL-15.3.

Przygotowanie badania

We wspomnianym wcześniej PostgreSQL stworzona została baza danych zawierająca tabelę geochronologiczną. Zawarte w niej dane podzielone zostały na:

- eony (*GeoEon*),
- ery (*GeoEra*),
- okresy (*GeoOkres*),
- epoki (*GeoEpoka*),
- piętra (*GeoPietra*).

Z poszczególnych mniejszych segmentów stworzona została główna tabela stratygraficzna – *TabelaGeo*.

Na potrzeby testów stworzone zostały również tabele pomocnicze: *dziesiec* oraz *milion*.

Całość pliku zawierającego wszystkie niezbędne fragmenty kodu umieszczona została w repozytorium.

Przebieg badania

Przygotowane zostały również zapytania, które miały definiować różnicę w wydajności poszczególnych złączeń.

Zapytanie 1 – jego głównym celem jest złączenie syntetycznej tablicy zawierającej milion wyników z tabelą geochronologiczną. Co ważne, odbywa się to w postaci zdenormalizowanej.

```
SELECT COUNT(*) FROM tab.milion INNER JOIN tab.TabelaGeo ON  
(mod(tab.milion.liczba, 77)=(tab.TabelaGeo.id_pietro));
```

Zapytanie 2 – jego celem jest złączenie syntetycznej tablicy zawierającej milion wyników z tabelą geochronologiczną w postaci znormalizowanej (złączenie wszystkich tabel).

```
SELECT COUNT(*) FROM tab.milion INNER JOIN tab.GeoPietro ON  
(mod(tab.milion.liczba, 77) = tab.GeoPietro.id_pietro)  
NATURAL JOIN tab.GeoEpoka  
NATURAL JOIN tab.GeoOkres  
NATURAL JOIN tab.GeoEra  
NATURAL JOIN tab.GeoEon;
```

Zapytanie 3 – jego celem jest złączenie syntetycznej tablicy miliona wyników z tabelą geochronologiczną w postaci zdenormalizowanej. Ma to miejsce przez zagnieżdżenie skorelowane.

```
SELECT COUNT(*) FROM tab.milion WHERE mod(tab.milion.liczba, 77) =  
(SELECT id_pietro FROM tab.TabelaGeo WHERE mod(tab.milion.liczba, 77) =  
(id_pietro));
```

Zapytanie 4 – jego celem jest złączenie syntetycznej tablicy miliona wyników z tabelą geochronologiczną w postaci znormalizowanej. Złączenie jest wykonywane przez zagnieżdżenie skorelowane, a zapytanie wewnętrzne jest złączeniem tabel poszczególnych jednostek geochronologicznych.

```
SELECT COUNT(*) FROM tab.milion WHERE mod(tab.milion.liczba, 77) IN  
(SELECT tab.GeoPietro.id_pietro FROM tab.GeoPietro NATURAL JOIN tab.GeoEpoka  
NATURAL JOIN tab.GeoOkres NATURAL JOIN tab.GeoEra NATURAL JOIN tab.GeoEon);
```

Każdy test został powtórzony dziesięciokrotnie. Z wyników zostały wyciągnięte wartości średnie oraz minimalne. Następnie pomiary czasu powtórzono, ale po ustaleniu indeksowania.

Wyniki

Porównanie średnich i minimalnych czasów wykonywania zapytań zostało przedstawione w *Tabeli 1*.

	1 ZL		2 ZL		3 ZG		4 ZG	
	średnia [ms]	min [ms]	średnia [ms]	min [ms]	średnia [ms]	min [ms]	średnia [ms]	min [ms]
BEZ INDEKSOWANIA								
PostgreSQL	230,4	189	457,2	360	15802,6	15192	250,2	178
Z INDEKSOWANIEM								
PostgreSQL	253,8	179	370	321	15935,2	15230	222,4	185

Tabela 1 wyniki pomiarów czasu wykonywania zapytań zarówno bez indeksowania, jak i z jego uwzględnieniem

Wnioski

Najdłuższym czasem wykonywania zadania charakteryzuje się zapytanie nr 3 (zagnieżdżenie skorelowane, zdenormalizowane,). Indeksowanie minimalnie wpłynęło na czas jego wykonywania wydłużając go. W przypadku zapytania nr 2 oraz nr 4 (znormalizowane) obecność indeksów korzystnie wpłynęła na średni czas ich wykonywania. Zauważalny jest również spadek wydajności, który nastąpił po optymalizacji zapytania (biorąc pod uwagę numery 1 oraz 2).