

PBD_DOCKER	Romaniak Hubert	Informatyka niestacjonarna III rok	Semestr zimowy 2024/25
------------	-----------------	---------------------------------------	---------------------------

Opis zadania

Zadanie polega na uruchomieniu w Dockerze dwóch kontenerów z bazami MySQL, mających identyczny schemat. Następnie, należy uzupełnić jedną z baz losowymi danymi.

Główną częścią zadania jest uruchomienie trzeciego kontenera z aplikacją, która przekopiuje rekordy z jednej bazy do drugiej. Algorytm przepisujący powinien być nieoptymalny. Należy zmierzyć czas przepisywania dla kilku różnych rozmiarów tabeli.

Realizacja

Algorytm, który został użyty do przepisania danych wygląda następująco:

1. Jeżeli ilość rekordów w bazach jest taka sama, zakończ algorytm
2. Wylosuj jeden z rekordów z 1. bazy
3. Sprawdź czy rekord znajduje się w 2. bazie
4. Jeżeli rekord nie znajduje się w 2. bazie, idź do kroku 1
5. Zapisz rekord w 2. bazie
6. Idź do kroku 1

Algorytm ten przypomina „problem kolekcjonera kuponów”, który polega na przewidzeniu, jak długo należy zbierać kupony (za każdym razem losując ze zwracaniem), aby zebrać każdy z nich z określonej puli przynajmniej raz. Złożoność algorytmiczna tego problemu to $O(n \log n)$.

Otrzymane wyniki (każdy pomiar to średnia z 5 uruchomień algorytmu):

<i>ilość wpisów</i>	czas wykonania [s]
100	0,293217134
200	0,980364513
300	1,896155119
400	3,440258312
500	4,587092113
600	7,61340332
700	9,473242426
800	13,47536521
900	15,92954822
1000	19,56992645
1100	25,1285409
1200	30,26942282
1300	37,17220469
1400	41,45215893
1500	46,1513907
1600	53,48855033
1700	63,18006034
1800	72,70423965
1900	70,6115078
2000	87,23112741

Po wprowadzeniu danych do skryptu w Pythonie, używając biblioteki `scipy.optimize`, ostateczna funkcja opisująca linię trendu ma postać $f(x) = 0,000000853 \cdot x \cdot \log_{x-1}(36521146 \cdot x)$.

Dane te przedstawiają się w następujący sposób na wykresie.

