

Raport z wydajności aplikacji

w ramach projektu z przedmiotu Zaawansowane Programowanie Obiektowe i Funkcyjne

Aleksandra Kulczycka, Piotr Kawa

1 Wstęp

Naszym zadaniem było stworzenie aplikacji służącej do generowania dużych ilości danych osobowych (co najmniej 1 mln) powstałych z kombinacji istniejących imion, nazwisk, miast, ulic, prawidłowych numerów pesel, dat urodzenia, numerów dowodów osobistych. Zasadnicze elementy wymagane w rozwiązaniu:

- Odczytanie danych z plików zawierających imiona, nazwiska i inne “substraty”.
- Wygenerowanie listy “fikcyjnych osób”, których dane stanowią losową kombinację.
- Wygenerowanie losowych relacji pomiędzy osobami (pokrewieństwa, małżeństwa)
- Eksport do pliku xls lub xlsx.
- Minimalna ilość wygenerowanych osób - 1 milion.
- Przygotowanie raportu z wydajności aplikacji w stosunku do rozmiaru zadania

Poniżej przedstawiamy raport z wydajności naszej aplikacji.

2 Pomiary

Środowisko: Java

Parametry wydajności: Czas generowania danych

Metodologia: Pomiaru dokonano przez uruchomienie aplikacji kilka razy dla różnych parametrów i zmierzenie czasu wygenerowania danych

Do eksportu wygenerowanych rodzin do pliku xlsx posłużyła nam klasa `Exporter`, korzystająca z `SXSSF-Workbook`, która jest częścią biblioteki Apache POI. `SXSSFWorkbook` jest odpowiednikiem klasy `XSSFWorkbook`, ale pozwala na utworzenie arkuszy w sposób, który jest bardziej wydajny dla dużych zbiorów danych.

Przy pomocy statycznej metody klasy `System` w języku Java, tzn. `System.currentTimeMillis()`, która zwraca aktualny czas w milisekundach. (Czas jest reprezentowany przez liczbę milisekund od epoki, czyli od 1 stycznia 1970 roku 00:00:00 UTC.) wykonaliśmy testy dla liczby danych osobowych w przedziale 5 tys. - 10 mln.

Wyniki ukazane zostały w tabeli poniżej (czas przedstawiony jest już w sekundach).

##	liczba_osob	czas_generowania_s
## 1	5,000	1.367
## 2	10,000	1.734
## 3	50,000	3.683
## 4	100,000	5.116
## 5	500,000	17.090
## 6	800,000	25.635
## 7	1,000,000	32.016
## 8	5,000,000	138.698
## 9	8,000,000	239.685
## 10	10,000,000	311.843

Dodatkowo zaprezentowaliśmy wyniki pomiarów na wykresie.



3 Podsumowanie

Wyniki pomiarów czasowych pokazują, że aplikacja działa zgodnie z oczekiwaniami i charakteryzuje się wysoką wydajnością, co minimalizuje czas oczekiwania i zapewnia lepsze doświadczenie użytkownika. Czas wykonania zadania jest zadowalający, tzn czas działania aplikacji rośnie proporcjonalnie do rozmiaru danych wejściowych.