

## Przetwarzanie asynchroniczne AMQP – ćwiczenie

Celem ćwiczenia jest praktyczne zapoznanie się z protokołem AMQP oraz narzędziami służącymi do asynchronicznego przetwarzania danych.

Rozproszony system klasyfikujący.

Należy zaprojektować i zaimplementować system rozproszony, który w sposób potokowy będzie filtrował i klasyfikował dane.

Do dyspozycji jest 6 węzłów, których funkcjonalność jest następująca.

1. Węzeł dystrybuujący dane – wczytuje on z pliku i wprowadza dane do systemu.
2. 4 węzły robocze – wykonują zadania obliczeniowe.
3. Węzeł zbierający dane - zbiera wyniki z węzłów roboczych i zapisuje je do pliku wyjściowego.

Każdy węzeł pracuje jako osobna aplikacja.

Należy zaprojektować i wykonać system, który będzie realizował zadanie klasyfikacji i filtrowania danych wykorzystujący węzły robocze.

System ładuje dane z pliku tekstowego i przesyła je do węzłów roboczych, które wykonują klasyfikację i filtrowanie danych po klasyfikacji (przy pomocy biblioteki WEKA).

W wyniku filtrowania usuwane są obiekty należące do jednej z klas (wybranej przez studenta).

Wyniki są zapisywane zgodnie z kolejnością ich przetworzenia.

Projektując system należy zwrócić uwagę na wydajność systemu (konfiguracja węzłów roboczych jest istotna!).

System należy zaprojektować w UML (tylko diagramy komponentów, wdrożenia i aktywności) i wykonać w technologii AMQP.

Jeden z węzłów ma być wykonany przy pomocy innej implementacji AMQP niż Apache QPID/Java.

Materiały:

AMQP: <http://www.amqp.org/>

Apache QPID: <http://qpido.apache.org/>

WEKA: <http://weka.wikispaces.com/Use+WEKA+in+your+Java+code>

Dane: <http://archive.ics.uci.edu/ml/datasets/Statlog+%28Shuttle%29>