



POZNAN UNIVERSITY OF TECHNOLOGY

Tematy projektów Edycja 2017

Robert Wrembel
Poznan University of Technology
Institute of Computing Science
Robert.Wrembel@cs.put.poznan.pl
www.cs.put.poznan.pl/rwrembel



Reguły

- ⇒ **Projekty zespołowe** ⇒ max. 3 osoby w zespole
- ⇒ **Zaliczenie studium** ⇒ dyplom
 - 60% egzamin pisemny
 - 40% projekt
- ⇒ **Zaliczenie projektu**
 - **prezentacja końcowa (obrona): ocenie podlega**
 - zawartość
 - struktura
 - sposób prezentacji
 - **raport techniczny: ocenie podlega**
 - jakość produktu projektu
 - zawartość raportu
 - struktura raportu
 - język (czytelność prezentacji)



Reguły

➤ Realizacja projektu

- metodyka (Scrum, Prince2, ...)
- manager
- zespół
- cykliczne prezentacje na zajęciach
 - postęp
 - problemy



Temat 1

- Projekt i implementacja systemu analitycznego opartego **Netezza**
- Wymagania
 - dane dot. lotów w USA
 - symulator Netezza
 - ETL: dowolne
 - 3 osoby



Temat 1

⇒ Wymagania cd.

- aplikacje analityczne: tech. dowolna
 - liczby odwołanych lotów w zadanym czasie (wskazany dzień, miesiąc, kwartał, rok), z podziałem na linie i typy samolotów
 - sumaryczne opóźnienia godzinowe lotów (wskazany dzień, miesiąc, kwartał, rok), z podziałem na linie i typy samolotów
 - linia lotnicza z najkrótszym sumarycznym czasem opóźnienia w zadanym kwartale lub miesiącu
 - najczęstsze trasy z podziałem na linie lotnicze

⇒ Zespół

- Aleksander Cybulski
 - Maciej Appelt
 - Jakub Nowicki
-



Temat 2

⇒ DB2 Design Advisor i Oracle Database Advisor w optymalizacji zapytań analitycznych

⇒ Wymagania

- Star Schema Benchmark
- zdefiniowanie struktur fizycznych (indeksy, MV) w oparciu o DB2 DA i Oracle DA
- ocena wydajności hurtowni danych przed i po zastosowaniu ww. struktur
- 3 osoby



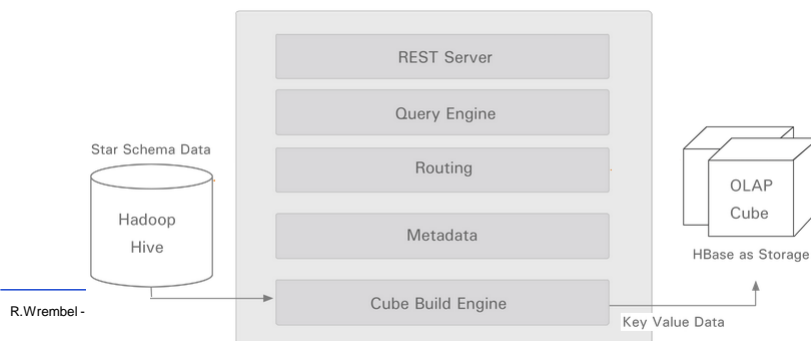
Temat 3

- **Row-store i Column-store w SQL Server: prównanie wydajności**
- **Wymagania**
 - **Star Schema Benchmark**
 - **ocena wydajności obu rozwiązań dla pojedynczych zapytań benchmarku i całościowo - dla obciążenia 13 zapytaniami**
 - **3 osoby**
- **Zespół**
 - **Julita Wyrwał**
 - **Marysia Prokopowicz**



Temat 4

- **Projekt i implementacja systemu analitycznego opartego o Apache Kylin**
- **Wymagania**
 - **źródło danych: np. Twitter**
 - **zbudowanie kostki**
 - **3 osoby**





Temat 4

⇒ Wymagania cd.

- analiza średnich długości tweetów (słowa, znaki) z podziałem na języki
- analiza liczby tweetów z podziałem na lokalizacje (np. kraje)
- histogram długości tweetów (słowa, znaki) dla wskazanych 5-ciu języków



Temat 5

- Projekt i implementacja systemu analitycznego opartego o **Apache Drill**
- Wymagania
 - źródło danych: np. Twitter
 - zbudowanie kostki
 - analiza średnich długości tweetów (słowa, znaki) z podziałem na języki
 - analiza liczby tweetów z podziałem na lokalizacje (np. kraje)
 - histogram długości tweetów (słowa, znaki) dla wskazanych 5-ciu języków
 - 3 osoby
- ⇒ Zespół
 - Karol Kaźmierczak
 - Piotr Markwitan
 - Łukasz Dłużewski



Temat 6

- **Projekt i implementacja systemu analitycznego opartego o [Apache Impala](#)**
- **Wymagania**
 - źródło danych: np. Twitter
 - zbudowanie kostki
 - analiza średnich długości tweetów (słowa, znaki) z podziałem na języki
 - analiza liczby tweetów z podziałem na lokalizacje (np. kraje)
 - histogram długości tweetów (słowa, znaki) dla wskazanych 5-ciu języków
 - 3 osoby
- ⇒ **Zespół**
 - Marcin Kurek
 - Artur Mielniczek
 - Marcin Borowski



Temat 7

- ⇒ **Analiza kursów walut w [Spark](#)**
- ⇒ **Wymagania**
 - źródło danych: kursy walut NBP za ostatnie 10 lat
 - składnica danych: HBase
 - analizy
 - wyznaczenie trendu wskazanej waluty
 - wizualizacja średnich miesięcznych kursów za wskazany okres dla wskazanych walut
 - znalezienie min i max ceny wskazanej waluty we wskazanym okresie
 - 3 osoby
- ⇒ **Zespół**
 - Tomasz Pianka
 - Marcin Matczuk
 - Wojciech Knysz