



INŻYNIERIA OPROGRAMOWANIA SYSTEMÓW
ROZPROSZONYCH

DOKUMENTACJA

**System agregacji i przetwarzania
zdarzeń systemowych**

Autorzy:
Łukasz OPIOŁA
Bartosz POLNIK

Prowadzący:
Dr. Inż. Marcin JARZĄB

18 marca 2014

1 Opis tematu

System agregacji logów z rozproszonych węzłów obliczeniowych wspierających maszyną skalowalność w oparciu o bibliotekę Kafka. Na węzłach uruchomieni są agenci (producenci) śledzący zmiany w plikach zdarzeń i wysyłający informacje do brokera Kafka, który następnie wpisuje dane do Hadoopa. Celem jest dostarczenie określonych raportów odnośnie kategorii zdarzeń (INFO, ERROR, DEBUG), częstotliwości ich występowania etc.

2 Technologie

Projekt zostanie zrealizowany z wykorzystaniem następujących technologii:

- Cloudera
- Hadoop Yarn (MR2)
- Hive
- HBase
- Oozie
- Kafka

3 Podział zadań

Łukasz Opiola	Bartosz Polnik
4	5
7	8