



INŻYNIERIA OPROGRAMOWANIA SYSTEMÓW
ROZPROSZONYCH

DOKUMENTACJA

Kaflog

Autorzy:

Łukasz OPIOŁA
Bartosz POLNIK
Michał JANIEC

Prowadzący:

Dr. Inż. Marcin JARZĄB

24 czerwca 2014

1 Opis tematu

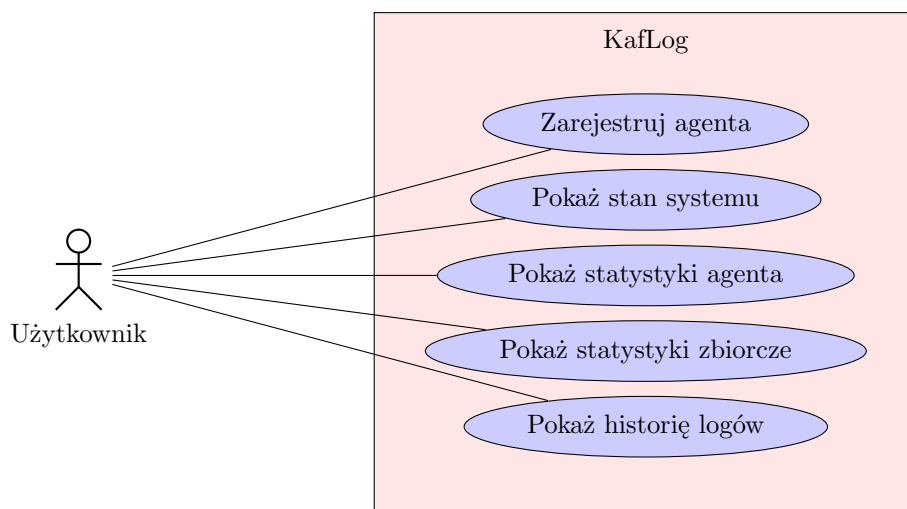
Celem projektu jest utworzenie systemu agregacji logów z rozproszonych węzłów obliczeniowych wspierających maszyną skalowalność w oparciu o bibliotekę Kafka. Na węzłach uruchomieni są agenci (producenci) śledzący zmiany w plikach zdarzeń i wysyłający informacje do brokera Kafka, który następnie wpisuje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

2 Technologie

Projekt zostanie zrealizowany z wykorzystaniem następujących technologii:

- Cloudera - do konfiguracji i zarządzania clustrem
- Hadoop Yarn (MR2) - do importu danych z brokera Kafka
- Hive - do przechowywania danych umożliwiających generowanie widoków
- HBase - do przechowywania widoków
- Kafka - do zbierania logów od agentów
- Impala - do łączenia widoków warstwy usługowej
- JMX - do monitorowania agentów

3 Przypadki użycia



4 Podział zadań

ŁUKASZ OPIOŁA	BARTOSZ POLNIK	MICHAŁ JANIEC
utworzenie szablonu środowiska	zarządzanie zespołem	zbieranie logów w celu przetwarzania real-time
tworzenie use-case	utworzenie szablonu projektu i podział na moduły	przetwarzanie logów w czasie rzeczywistym
utworzenie producenta	zapisywanie logów w celu generacji widoków w hdfsie	utworzenie inkrementacyjnych widoków na podstawie przetwarzanych danych
generowanie statystyk przez agentów	utworzenie widoków warstwy batch	utworzenie widoku statystyk w gui
widok w gui do monitorowania agentów	łączenie widoków	dodanie komentarzy do kodu
widok w gui aktualnie generowanych logów	testowanie platformy	refactoring kodu
testy wydajnościowe	dopasowanie generacji widoków real-time i batch celem integracji z Apache Impala	optymalizacja działania aplikacji do monitorowania agentów
zapewnienie bezpieczeństwa platformie	dodanie bezpieczeństwa do aplikacji monitorującej agentów	dodanie informacji o przedziale czasowym statystyk w gui
opis poszczególnych modułów	utworzenie przykładów dla gui	umożliwienie uruchamiania storma lokalnie
utworzenie diagramu architektury	uruchamianie cykliczne generacji widoków warstwy batch	poprawa wizualna widoku statystyk