

Program szkolenia:

Apache Kafka

Informacje:

Nazwa:	Apache Kafka
Kod:	BigDataML-Kafka
Kategoria:	BigData, streaming i Machine Learning
Grupa docelowa:	developerzy
Czas trwania:	2 dni
Forma:	50% wykłady / 50% warsztaty

Szkolenie jest praktycznym warsztatem, podczas którego uczyliśmy się jak zbudować aplikację, która korzysta z klastra Kafka.

W części wykładowej uczestnicy zrozumieją jaką rolę pełni Kafka we współczesnym rozproszonym pipeline, z jakich komponentów się składa oraz jak efektywnie korzystać z API.

Zalety szkolenia:

- Przykłady o realnym poziomie komplikacji
- Możliwość kontaktu z doświadczonym ekspertem

Szczegółowy program:

1. Introduction to Apache Kafka

1.1. The LinkedIn Problem

1.2. Why Kafka

1.3. Basic terminology

2. Getting Started

2.1. Installing Kafka

2.2. Writing Messages

2.3. Reading Messages

3. More about writes and reads

3.1. Configuring producers

3.2. More about Serializers

3.3. The Pool Loop

3.4. Commits and Offsets

3.5. More about Deserializers

3.6. Configuring Consumers

4. Reliability

4.1. Fallacies of distributed computing

4.2. Reliability Guarantees

4.3. Reliable systems (and ways to validate it)

4.4. Data Pipelines

4.4.1. Kafka Connect vs Producers and Consumers

4.4.2. Other alternatives

5. Internals

5.1. Cluster membership

5.2. Replication

5.3. Request Processing

5.4. Physical Storage

6. Monitoring and Administration