

Kod szkolenia: **NoSQL/360**

Tytuł szkolenia: **Wstęp do baz NoSQL**

Dni: 3

Opis:

Adresaci szkolenia:

Szkolenie jest adresowane do osób, które chcą zdobyć przekrojową wiedzę na temat baz NoSQL, ich zastosowań, możliwości i ograniczeń. Materiał szkolenia w szczególności przeznaczony jest dla programistów i architektów zajmujących się projektowaniem i tworzeniem systemów operujących na dużych zbiorach danych. Przegląd rozwiązań dokonywany podczas szkolenia pomaga w wyborze optymalnych technologii do rozwoju własnych systemów.

Cel szkolenia:

Uczestnicy szkolenia zdobędą ogólną wiedzę dotyczącą baz typu NoSQL, ich funkcjonalnościach, zastosowaniach i ograniczeniach.

W szczególności uczestnicy:

- poznają podstawy teoretyczne związane z bazami NoSQL
- poznają różne rodzaje baz NoSQL, ich wiodące implementacje, przykładowe zastosowania oraz ich wady i zalety
- dowiedzą się jak zapewniać spójność danych w bazach NoSQL oraz w jaki sposób implementować obsługę transakcji
- zapoznają się z narzędziami do masowego przetwarzania danych takimi jak: Apache Spark, czy Apache Hadoop

Mocne strony szkolenia:

Szkolenie jest prowadzone przez osoby, które na codzień pracują z NoSql i mają praktyczne i komercyjne doświadczenie w tym zakresie. Szkolenie pozwoli poznać świat nierelacyjnych baz, ich praw i ograniczeń oraz możliwości rozproszonego przetwarzania danych.

Wymagania:

Podstawowa znajomość baz SQL i modelowania danych. Mile widziana znajomość języka Java.

Specjalne wymagania techniczne

W cenie szkolenia uczestnik otrzymuje na własność obraz maszyny wirtualnej wykorzystaną na szkoleniu. Szkolenie musi być realizowane na maszynie wirtualnej, stąd wymagane jest min. 12GB RAM na komputerze szkoleniowym.

Parametry szkolenia:

3*8 godzin (3*7 godzin netto) wykładów i warsztatów. W trakcie warsztatów, oprócz prostych ćwiczeń, uczestnicy projektują bazę danych i implementują do niej proste przetwarzanie z wykorzystaniem Apache Spark.

Program szkolenia:

1. Wstęp do NoSQL
 - I. Przyczyny powstania baz NoSQL, historia NoSQL
 - II. Spójność, Dostępność, Odporność na partycjonowanie
 - III. Twierdzenie CAP
 - IV. Podstawowe parametry parametry baz NoSQL
 - V. Koncepcja Google Bigtable i Amazon DynamoDB
 - VI. Ogólna klasyfikacja baz NoSQL (klucz-wartość, kolumnowe, dokumentów, grafowe itd.)
 - VII. Spójność danych w bazach NoSQL
 - VIII. Rozproszone systemy przechowywania danych
2. Przegląd najpopularniejszych rozwiązań składowania danych wraz z przykładami użycia
 - I. Bazy typu klucz-wartość
 - i. Redis
 - ii. Memcached
 - II. Bazy kolumnowe
 - i. HBase (Hadoop database)
 - ii. Cassandra / Scylla
 - III. Bazy dokumentowe
 - i. MongoDB
 - ii. Couchbase / CouchDB
 - IV. Bazy grafowe
 - i. Neo4J
 - ii. JanusGraph (dawniej TitanDB)
 - V. Składowanie plików
 - i. Hadoop Distributed File System (HDFS)
 - VI. Przechowywanie strumieni danych
 - i. Kafka
 - VII. Szeregi czasowe
 - i. Druid
 - VIII. Silniki wyszukujące



- i. Elasticsearch
- 3. Przetwarzanie danych zgromadzonych w NoSQL
 - I. Wsadowe (batch)
 - i. Opis koncepcji
 - ii. Hadoop YARN/MR
 - iii. Spark
 - II. Strumieniowe
 - i. Opis koncepcji
 - ii. Spark Streaming
 - iii. Flink
 - iv. Storm

