

Kod szkolenia: **CASSANDRA**

Tytuł szkolenia: **Bazy danych NoSQL - Cassandra**

Dni: 3

## Opis:

### Adresaci szkolenia:

Szkolenie jest adresowane do programistów aplikacji internetowych, którzy chcą tworzyć systemy oparte o skalowalne bazy danych, a w szczególności osób dla których wydajność opracowywanych rozwiązań ma wysoki priorytet. Szkolenie jest także kierowane do osób zajmujących się obecnie tworzeniem relacyjnych baz danych, które chcą uzupełnić swoją wiedzę o technologii alternatywne, które stopniowo wypierają bazy relacyjne z poszczególnych obszarów zastosowań.

### Cel szkolenia:

Uczestnicy szkolenia zdobędą ogólną wiedzę dotyczącą baz typu NoSQL, ich funkcjonalnościach, zastosowaniach i ograniczeniach. Dzięki temu będą mogli wybrać właściwą bazę danych dla swoich projektów. Szkolenie, poza ogólnym wprowadzeniem do baz nierelacyjnych, skupia się na bazie danych Cassandra, stworzonej pierwotnie przez Facebook'a.

W szczególności uczestnicy:

- poznają model danych typu BigTable oraz jego realizację w bazie Cassandra,
- dowiedzą się jak dane są przechowywane i jak fizycznie realizowany jest dostęp do nich,
- poznają mechanizmy zapewniające wysoką dostępność i wydajność oraz kompromisy jakie się z tym wiążą,
- poznają również różnych klientów do łączenia się bazą Cassandra, w szczególności dla języka Java (Pelops, Hector, Kundera) oraz uniwersalnych (Thrift, CQL).

### Mocne strony szkolenia:

Program obejmuje zarówno ogólne wprowadzenie w tematykę nierelacyjnych baz danych jak i całościowe przedstawienie bazy danych Cassandra wraz z najważniejszymi jej klientami dla języka Java.

Szkolenie jest unikalne, gdyż tematyka poruszana w jego trakcie nie jest wyczerpująco ujęta w dostępnej literaturze, a wiedza na ten temat jest mocno rozproszona.

Program jest ciągle uaktualniany ze względu na szybki rozwój rozwiązań, których dotyczy szkolenie.

## Wymagania:

Od uczestników szkolenia wymagana jest umiejętność programowania w języku Java (do poznania na kursie J/JP).

## Parametry szkolenia:

3\*8 godzin (3\*7 godzin netto) wykładów i warsztatów (z wyraźną przewagą warsztatów).

W trakcie warsztatów, oprócz prostych ćwiczeń, uczestnicy projektują bazę danych i implementują dostęp do niej.

Wielkość grupy: maks. 8-10 osób.

## Program szkolenia:

1. Wprowadzenie do baz NoSQL
  - I. Przyczyna powstania baz chmurowych
  - II. Spójność, Dostępność, Odporność na partycjonowanie
  - III. Twierdzenie CAP
  - IV. Co różni bazy NoSQL od baz relacyjnych
  - V. Podstawowe parametry baz NoSQL
  - VI. Klasyfikacja i przegląd baz NoSQL (Cassandra, Hbase, Mongo, Riak, CouchDB, Tokyo Cabinet, Voldemort, etc.)
  - VII. Gdzie warto wykorzystać bazę Cassandra
  - VIII. Instalacja i uruchomienie Cassandra
  - IX. Wbudowany interfejs
2. Struktura danych
  - I. Bazy danych, tabele, wiersze, kolumny
  - II. Typy danych
  - III. Grupowanie wierszy: klucz partycjonujący a klucz grupujący
  - IV. Kolekcje
  - V. Kolumny licznikowe
  - VI. Indeksy pomocnicze
  - VII. Przykłady tworzenia bazy
  - VIII. Przykłady konwersji bazy relacyjnej do bazy NoSQL
3. Praca z danymi i interfejsy programistyczne
  - I. CQL3:
    - ustawianie poziomu spójności
    - dodawanie / nadpisywanie danych
    - praca z kolekcjami
    - hurtowy zapis wielu danych
    - niepodzielny zapis warunkowy (lekkie transakcje)

- zapytania po kluczu głównym
- zapytania wykorzystujące indeksy
- triggery
- II. Oficjalny sterownik do Cassandra dla języka Java
  - podłączani e się do klastra, sesje
  - wykonywanie zapytań i odbieranie wyników
  - wywołania asynchroniczne
- III. Pozostałe interfejsy: Thrift, JDBC, Pelops, Hector
- IV. Współpraca z infrastrukturą Hadoop i mechanizm Map-Reduce
- 4. Wewnętrzna struktura danych i mechanizmy
  - I. Mechanizm Read-Repair
  - II. Mechanizm Hinted-Handoff
  - III. Commit-Log
  - IV. Flushing
  - V. Compaction
  - VI. SSTables
  - VII. Kompresja
  - VIII. Repair
- 5. Konfiguracja i optymalizacja
  - I. Konfiguracja Cassandra
  - II. Rozdzielenie log'u od danych
  - III. Konfiguracja klastra
  - IV. Partycjonowanie
  - V. Monitorowanie
  - VI. Autentykacja i autoryzacja
  - VII. Backup
  - VIII. Konfiguracja cache i pamięci
- 6. Dodatkowe źródła wiedzy
- 7. Demo i przegląd możliwości DataStax Enterprise.

