kod szkolenia: ES_ADM_ZARZ / PL AA 3d

Elasticsearch - administracja i zarządzanie

Przeznaczenie szkolenia

Szkolenie skierowane do DevOps, programistów, testerów oraz wszystkich osób związanych z Elasticsearch chcących usystematyzować aktualną wiedzę oraz rozwijać się dalej w kierunku administracji klastrami Elasticsearch w środowiskach Big Data.

Korzyści wynikające z ukończenia szkolenia

Szkolenie ma charakter warsztatowy i jest ukierunkowane na najciekawsze i najważniejsze zagadnienia związane z administracją klastrem Elasticsearch oraz całym Elastic stack (Beats, Logstash, Kibana, Elasticsearch). Uczestnicy mają możliwość ustematyzowania aktualnej wiedzy, jednak szkolenie jest głównie ukierunkowane na zdobywanie umiejętności pozwalających na samodzielną optymalizację klastra, poprawianie performance klastra, budowanie odpowiedniego monitoringu, czy odpowiedniego modelowania klastra (ilość nodów, indeksów, shardów) w systemach przetwarzających TB a nawet PB danych. Szkolenie pokazuje dobre praktyki jak również architektury i podejścia których należy unikać w pracy z Elasticsearch.

Oczekiwane przygotowanie słuchaczy

- Dowolne doświadczenie w pracy z produktem Elasticsearch
- Doświadczenie w pracy z RESTful API

Język szkolenia

- Szkolenie : polski

Czas trwania

3 dni / 21 godzin

Agenda szkolenia

1. Elasticsearch – podstawy
   - Analiza danych
   - Indeksy Lucene – co to jest i dlaczego mówimy o Lucene w kontekście Elasticsearch,
   - Architektura produktu Elasticsearch (cluster, node, index, primary shard, replica shard, mapping, document, JSON, index templates itp.)
   - Praca z NoSQL w środowiskach BigData,
2. Elasticsearch – cluster
   - Budowa klastra
   - Konfiguracja nodów w klastrze (ingest/data/master/ml) – dobór ilości masterów w klastrze, ilości nodów w klastrze, ilości indeksów oraz właściwy dobór ilości oraz wielkości shardów dla indeksu,
   - Parametry konfiguracyjne,
   - Mechanizm Discovery
   - Skalowanie klastra Elasticsearch (scale-out i scale-in) – skalowalność horyzontalna i wertykalna
   - ILM – Cykle życia oraz stany indeksów,
   - Fault tolerance – odporność na awarie
   - Optymalizacja klastra – parametry konfiguracyjne, custom allocation, routing, tags, rack_id, zone itp.
   - Optymalizacja mappingu
   - Split-brain effect, load balancing, transport client,
   - Zaawansowane konfiguracja klastra Elasticsearch – przegląd i wyjaśnienie parametrów konfiguracyjnych
   - Elasticsearch RESTful API (request/response, klucze _id)
   - Cross-cluster search
3. Ingest danychLogstash
   - Wprowadzenie do produktu Logstash,
     - Transformacje danych
     - Przykładowe użycie Logstash do importu danych testowych, danych z Twitter’a, enrichment danych
     - Massive ingestion,
   - Ingest Pipelines

- Beats
  - Wprowadzenie do produktu Beats,
  - Uruchomienie metricbeat lub filebeat. Przegląd dobrych praktyk podczas pracy z Beats (optymalizacja konfiguracji)

4. Query DSL
   - Kibana – Podstawy Query DSL z użyciem Kibany (nawigacja, query, tworzenie wizualizacji)
   - Tips & tricks optymalizacji zapytań
5. Support & maintenance
   - Backupy – backup i restore klastra Elasticsearch,
   - Monitoring
   - Poprawa performance klastra,
   - Rozwiązywanie typowych problemów dnia codziennego w pracy z Elasticsearch
   - Integracja ElasticSearch z istniejącymi systemami
   - Zarządzanie klastrem poprzez RESTful API