kod szkolenia: ES_PRO / PL AA 2d

Elasticsearch: Professional

Przeznaczenie szkolenia

Szkolenie skierowane do administratorów, programistów, testerów oraz wszystkich osób związanych z Elasticsearch chcących nauczyć się lub podnieść kompetencje. Szkolenie wymaga od uczestników znajomości produktu Elasticsearch.

Korzyści wynikające z ukończenia szkolenia

Jest to kolejne szkolenie z cyklu szkoleń o Elasticsearch, będące następnym etapem w budowaniu kompetencji z zakresu zarządzania, konfigurowania oraz używania produktu Elasticsearch. Szkolenie porusza bardzo specyficzne elementy produktu pozwalające na zaawansowany tuning oraz optymalizację klastrów BigData w środowiskach produkcyjnych. Po zakończeniu szkolenia uczestnicy są w stanie efektywnie optymalizować produkcyjne klastry Elasticsearch przetwarzające setki albo i tysiące TB danych.

Oczekiwane przygotowanie słuchaczy

- Znajomość zagadnień wchodzących w skład szkoleń Elasticsearch Essentials oraz Elasticsearch Intermediate
- Podstawowa wiedza o serwisach RESTful
- Bash, w tym użycia curl (ewentualnie innego oprogramowania do wysyłania requestów HTTP jak Postman, Fiddler itp.)
- Znajomość wyrażeń regularnych (regex), popularnych języków skryptowych (Javascript, painless, mustache itp.) oraz Java

Język szkolenia

Szkolenie: polski

Czas trwania

2 dni / 14 godzin

Agenda szkolenia

1. Elasticsearch cluster
   - 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,
   - Skalowanie klastra Elasticsearch,
   - Cykle życia oraz stany indeksów,
   - Pluginy,
   - Fault tolerance – odporność na awarie
   - Optymalizacja klastra – parametry konfiguracyjne, custom allocation, tags, rack_id, zone itp.
2. Optymalizacja istniejącej architektury systemu
   - Sposoby na wdrożenie Elasticsearch do aktualnej architektury,
   - Massive ingestion,
   - Wykorzystanie systemów kolejkowych do optymalizacji ingestu,
3. Elasticsearch – agregaty
   - Przegląd wszystkich dostępnych typów agregacji,
   - Ryzyko używania agregat,
   - Optymalizacje agregat,
4. Elasticsearch – algorytm alokacji
   - Używanie routingu do optymalizacji wydajności klastra,
   - Wpływanie na algorytm alokacji celem poprawienia wydajności klastra,
5. Elasticsearch – Scoring
   - Domyślny scoring TF*IDF vs BM25
   - Custom scoring w praktyce – wpływanie na scoring oraz używanie custom scoringu
6. Logstash
   - Klaster Logstash,
   - Transformacje danych w Logstash,
   - Filtry grok
7. Machine Learning

- Temporal Anomaly Detection vs Population Anomaly Detection
- Przegląd funkcji Anomaly Explorer’a
- Zasada działania algorytmu kategoryzacji,

8. Kibana
   - Przykłady wykorzystania GeoIP oraz wizualizacji (Maps)
   - Instrumentacja kodu z użyciem APM server
9. Performance testing
   - Metody na testowanie wydajności klastra,
   - Budowanie testów z użyciem JMeter,
   - Analiza ustawień klastra pod kątem wydajności,
   - Analiza danych oraz mappingu pod kątem wydajności,
10. Security
    - Uruchomienie modułu Security,
    - Zarządzanie uprawnieniami