

RTBHOUSE =

Big Data Engineer

Team Core

TWOJE ZADANIA:

- rozwój i utrzymanie rozproszonych systemów przetwarzających duże ilości danych (większość real-time) na potrzeby naszej platformy RTB
- optymalizacja tworzonego oprogramowania pod kątem wydajności i zużycia zasobów
- zapewnienie niezawodności i skalowalności budowanych rozwiązań
- tworzenie testów wydajnościowych i poprawnościowych dla nowych elementów systemu
- analiza nowych technologii pod kątem możliwości ich zastosowania w warunkach produkcyjnych
- rozwój narzędzi monitorujących i analizujących pracę produkcyjnego systemu
- ciągła optymalizacja istniejących narzędzi i procesów

WYBRANE WYKORZYSTYWANE TECHNOLOGIE:

- Java, Python
- Hadoop, Kafka
- Kafka Streams, Flume, Logstash
- Docker, Jenkins, Graphite
- Aerospike, PostgreSQL
- Google Big Query, Elastic

PRZYKŁADOWE TEMATY:

Wybrane zagadnienia, którymi zajmowaliśmy się w ostatnim czasie:

- wymiana frameworka w komponencie przetwarzającym dane (przejście ze Storm-a na Kafka Streams)
- stworzenie mergera strumieni danych w oparciu o Kafka Client API
- stworzenie synchronizatora profili użytkowników między DC-kami w oparciu o Kafka Streams
- stworzenie komponentu wyliczającego agregaty w oparciu o Kafka Client API i filtry Blooma
- wdrożenie Logstash-a do ładowania i Elastic-a do odpytywania indeksowanych danych (przejście z Flume + Solr)
- stworzenie end-to-end monitoringu poprawności i opóźnienia danych
- wymiana komponentu streamującego dane do BigQuery i HDFS (z Flume na własne rozwiązanie oparte o Kafka Client API) – w trakcie
- ciągłe utrzymanie systemu, wykrywanie i rozwiązywanie problemów wydajnościowych, a także skalowanie ze względu na rosnącą ilość danych

NASZE OCZEKIWANIA:

- biegłość w programowaniu
- doskonałe rozumienie tego, jak działają złożone systemy informatyczne (od poziomu sprzętu, poprzez software, aż po algorytmikę)
- dobra znajomość podstawowych metod tworzenia programów współbieżnych i systemów rozproszonych (od poziomu wątków do poziomu kontynentów)
- praktyczna umiejętność obserwowania, monitorowania i analizowania działania produkcyjnych systemów (oraz wyciągania z tego wartościowych wniosków)
- umiejętność krytycznej analizy tworzonych rozwiązań pod kątem wydajności (od szacowania teoretycznej wydajności projektowanych systemów do wykrywania i usuwania faktycznych problemów wydajnościowych na produkcji)
- gotowość do pracy w modelu DevOps

DODATKOWYMI ATUTAMI BĘDĄ:

- doświadczenie w tworzeniu systemów rozproszonych
- dobra znajomość wybranych technologii Big Data takich jak Hadoop, Kafka, Storm, Spark lub Flink
- znajomość metod i narzędzi profilowania aplikacji (najlepiej Java, zarówno z poziomu JVM-a jak i Linux-a)

OFERUJEMY:

- atrakcyjne wynagrodzenie
- pracę w zespole pasjonatów, którzy chętnie dzielą się swoją wiedzą i doświadczeniem
- wyjątkowo elastyczne warunki pracy – nie mamy core hours, nie mamy limitów na urlopy, można w dużej części pracować zdalnie
- dostęp do najnowocześniejszych technologii i możliwość realnego wykorzystywania ich w projekcie o dużej skali i wysokiej dynamice
- sprzęt i oprogramowanie, jakiego potrzebujesz

Więcej informacji: techblog.rtbhouse.com/jobs/

Kontakt: it-jobs@rtbhouse.com