

PostgreSQL – Strojenie wydajnościowe

Cele szkolenia

Kurs ma na celu zapoznanie uczestników z zaawansowanymi zagadnieniami z zakresu wykorzystania serwera baz danych PostgreSQL. W ramach szkolenia uczestnicy poznają techniki badania wydajności operacji wykonywanych na bazie danych oraz samego serwera. Omawiane są kwestie zapewniania wysokiej efektywności zapytań i modyfikacji zawartości bazy poprzez wykorzystanie indeksowania oraz optymalizacji poleceń. Poruszane są typowe problemy występujące przy dużych zbiorach danych oraz metody ich eliminowania, w tym techniki partycjonowania. Ćwiczenia praktyczne realizowane są na platformie GNU/Linux Ubuntu.

Profil słuchaczy

Kurs przeznaczony jest dla wszystkich, którzy pracują z środowisku PostgreSQL, mają styczność z jego optymalizacją oraz chcieliby rozszerzyć swoją wiedzę i umiejętności w celu efektywnego wykorzystywania możliwości tego systemu.

Wymagania wstępne

Od uczestników szkolenia wymagana jest znajomość relacyjnych baz danych, podstaw języka SQL, podstaw środowiska PostgreSQL oraz podstaw poruszania się w środowisku GNU/Linux.

Czas trwania

Cztery sesje szkoleniowe, 16 godzin lekcyjnych.

Czas sesji szkoleniowej: 180 minut (4 godziny lekcyjne) + przerwy. W trakcie dnia szkoleniowego realizowane są maksymalnie dwie sesje szkoleniowe. Sesja szkoleniowa składa się z modułów szkoleniowych.

Metoda realizacji szkolenia

Szkolenie realizowane jest w formie występujących po sobie sesji szkoleniowych podzielonych na moduły. W ramach modułu szkoleniowego przewidziana jest część teoretyczna wyjaśniająca dane zagadnienie, wspólnie oraz samodzielnie rozwiązane przez uczestnika ćwiczenie oraz blok pytań i odpowiedzi

Wiedza teoretyczna i praktyczna

→ Sesje szkoleniowe 1 (180 min)

Moduł 1:

Omówienie:

- Organizacji czasowe szkolenia
- Środowiska pracy
- Formy komunikacji

Przedstawienie celów i zakresu szkolenia

Wprowadzenie do tematyki strojenia bazy danych

Moduł 2:

Konteksty pracy parametrów serwera

Podstawowe parametry wydajnościowe

Sterowanie parametrami

Wspólne ćwiczenie
Ćwiczenie samodzielne
Omówienie rozwiązań
Podsumowanie modułu

Moduł 3:

Czyszczenie bazy
Martwe wiersze
Operacja Vacuum
Vacuum Full
Wspólne ćwiczenia
Ćwiczenia samodzielne
Omówienie rozwiązań
Podsumowanie modułu

Moduł 4:

Czyszczenie bazy – ciąg dalszy
Autovacuum
Koszty czyszczenia
Operacje update w trybie Hot
Wspólne ćwiczenia
Ćwiczenia samodzielne
Omówienie rozwiązań
Podsumowanie modułu

Moduł 5:

Podsumowanie sesji
Pytania i odpowiedzi
Przedstawienie planu kolejnej sesji

→ Sesje szkoleniowe 2 (180 min)

Moduł 1:

Indeksy
Zastosowani indeksów
Operacje na bazie a indeksy
Statystyki
Przygotowanie środowiska do testowania indeksów
Wspólne ćwiczenia
Ćwiczenia samodzielne
Omówienie rozwiązań
Podsumowanie modułu

Moduł 2:

Indeksy – ciąg dalszy
Czytanie statystyk po wykonaniu polecenia
Interpretacja statystyk
Podstawowe skanowania z indeksem
Wspólne ćwiczenia
Ćwiczenia samodzielne
Omówienie rozwiązań
Podsumowanie modułu

Moduł 3:

Indeksy – ciąg dalszy
Nieefektywne wykorzystanie indeksów
Łączenie indeksów
Wspólne ćwiczenia
Ćwiczenia samodzielne
Omówienie rozwiązań

Podsumowanie modułu

Moduł 4:

- Indeksy – ciąg dalszy
- Klastry
- Operacja klastrowania tabel
- Wspólne ćwiczenia
- Ćwiczenia samodzielne
- Omówienie rozwiązań
- Podsumowanie modułu

Moduł 5:

- Indeksy – ciąg dalszy
- Tworzenie indeksu
- Przebudowa indeksu
- Tworzenie współbieżne
- Indeksy hashowe
- Wspólne ćwiczenia
- Ćwiczenia samodzielne
- Omówienie rozwiązań
- Podsumowanie modułu

Moduł 6:

- Podsumowanie sesji
- Pytania i odpowiedzi
- Przedstawienie planu kolejnej sesji

→ Sesje szkoleniowe 3 (180 min)

Moduł 1:

- Wprowadzenie do optymalizacji zapytań
- Buforowanie danych
- Plany wykonania zapytań
- Tryby wyświetlania planu
- Analiza informacji z planu wykonania
- Obliczanie kosztu
- Wspólne ćwiczenia
- Ćwiczenia samodzielne
- Omówienie rozwiązań
- Podsumowanie modułu

Moduł 2:

- Optymalizacja zapytań – ciąg dalszy
- Węzły w planie zapytania
- Sort
- Limit i Offset
- Aggregate
- Wspólne ćwiczenia
- Ćwiczenia samodzielne
- Omówienie rozwiązań
- Podsumowanie modułu

Moduł 3:

- Optymalizacja zapytań – ciąg dalszy
- Append
- Group
- HashAggregate
- Index Only Scan
- Wspólne ćwiczenia

Ćwiczenia samodzielne
Omówienie rozwiązań
Podsumowanie modułu

Moduł 4:

Optymalizacja zapytań – ciąg dalszy
Łączenie tabel
Kolejność łączenia
Algorytmy łączenia
Sterowanie planem wykonania
Wspólne ćwiczenia
Ćwiczenia samodzielne
Omówienie rozwiązań
Podsumowanie modułu

Moduł 5:

Statystyki
Statistic Collector
Widoki systemowe
Dane na temat tabel i indeksów
Rozmiar obiektów
Wspólne ćwiczenia
Ćwiczenia samodzielne
Omówienie rozwiązań
Podsumowanie modułu

Moduł 6:

Podsumowanie sesji
Pytania i odpowiedzi
Przedstawienie planu kolejnej sesji

→ Sesje szkoleniowe 4 (180 min)

Moduł 1:

Blokady
Zakleszczenia
Wykrywanie
Unikanie
Wspólne ćwiczenia
Ćwiczenia samodzielne
Omówienie rozwiązań
Podsumowanie modułu

Moduł 2:

Przygotowanie poleceń do wykonania
Prepare
Wspólne ćwiczenia
Ćwiczenia samodzielne
Omówienie rozwiązań
Podsumowanie modułu

Moduł 3:

Tabele tymczasowe
Tabele bez logowania
Zastosowanie
Wspólne ćwiczenia
Ćwiczenia samodzielne
Omówienie rozwiązań
Podsumowanie modułu

Moduł 4:
Mechanizm Toast
Zastosowanie
Sterowanie
Wspólne ćwiczenia
Ćwiczenia samodzielne
Omówienie rozwiązań
Podsumowanie modułu

Moduł 5:
Partycjonowanie tabel
Zastosowanie
Metody
Ograniczenia
Wspólne ćwiczenia
Ćwiczenia samodzielne
Omówienie rozwiązań
Podsumowanie modułu

Moduł 6:
Ładowanie masowe
Przyspieszanie
Parametry
Wspólne ćwiczenia
Ćwiczenia samodzielne
Omówienie rozwiązań
Podsumowanie modułu

Moduł 7:
Dodatkowe pytania uczestników
Podsumowanie szkolenia
Test
Ankieta

Umiejętności

Po ukończeniu kursu uczestnik/czka będzie potrafił/a:

- Po ukończeniu kursu uczestnik powinien rozumieć problemy wydajnościowe pojawiające się na serwerze PostgreSQL.
- Sterować parametrami serwera w celu zwiększenia wydajności.
- Powinien znać i umieć zastosować w praktyce metody zwiększające efektywność operacji, w tym techniki indeksowania,
- Uczestnik potrafił będzie badać plan wykonania zapytań i wyciągać wnioski umożliwiające jak najefektywniejsze wykorzystanie możliwości serwera.
- Potrafić implementować partycjonowania.

Ścieżka rozwoju

Po zakończeniu szkolenia rekomendowane jest skorzystanie z:

- Szkolenie: „PostgreSQL – administracja bazą danych”
- <https://www.postgresql.org/docs/12/index.html>