

39

Tworzenie i wykorzystanie indeksów

EFEKTY KSZTAŁCENIA Z PODSTAWY PROGRAMOWEJ:

- PKZ(E.b)(13) stosuje programy komputerowe wspomagające wykonywanie zadań;
- E.13.2(7) modyfikuje i rozbudowuje struktury baz danych.

W TYM ROZDZIALE:

- przypomnisz sobie, co to jest indeks;
- dowiesz się, jakie są typy indeksów;
- nauczysz się tworzyć i usuwać indeksy;
- dowiesz się, jak stosować indeksy;
- nauczysz się nawiązywać połączenie z bazą danych.

Wprowadzenie

W celu podniesienia wydajności bazy danych podczas jej przeszukiwania można zastosować indeks. Indeks to posortowana struktura danych używana do kojarzenia wartości z wierszami. Indeksy nałożone na atrybut organizują dane, dzięki czemu zmniejsza się liczba odczytywanych danych i operacje są wykonywane szybciej. Każda modyfikacja danych w obrębie indeksu powoduje konieczność jego aktualizacji i ponownego segregowania danych, co zwiększa obciążenie systemu. Większość systemów baz danych automatycznie indeksuje klucz główny. Stosowanie dodatkowych indeksów powinno być poprzedzone zebraniem i analizą informacji dotyczących zapytań, ich częstotliwości i obciążenia systemu. Nadmiarowy indeks powoduje straty wydajności spowodowane niepotrzebnym obciążeniem systemu.

W PostgreSQL są stosowane następujące typy indeksów:

- **B-tree (B-drzewa)** – używany najczęściej dla zapytań zawierających takie operatory porównania, jak: większe od, mniejsze od, większe lub równe, mniejsze lub równe, różne, pomiędzy (BETWEEN), IS NULL, IS NOT NULL;
- **indeks haszowany** – stosuje funkcję przekształcającą ciągi wartości klucza o dowolnej długości na wartość o ustalonej długości. Wykorzystuje się go w bazach danych, gdy wyszukiwanie jest oparte na równości klucza;
- **indeks GIST (Generalized search Tree)** – jest stosowany przede wszystkim do bardziej skomplikowanych typów danych i operacji spoza podstawowych porównań oraz określania zakresów (głównie za pomocą geometrycznych typów danych);
- **indeks GIN (Generalized Inverted Indexes)** – jest wykorzystywany, gdy zachodzi potrzeba zmapowania wielu wartości w jednym wierszu, np. przy tablicowych typach danych.

Aby utworzyć indeks, należy użyć polecenia **CREATE INDEX**. Domyślnie jest tworzony indeks typu B-tree. Aby utworzyć inne typy indeksów, używa się dodatkowych opcji, np. **USING hash(id_pracowni)**, aby utworzyć indeks haszowany.

Indeks wielokolumnowy jest stosowany, gdy na kolumnach często są używane zapytania zawierające warunek **AND**. Indeksy mogą być nakładane również na wyrażenia i funkcje wyliczane z pojedynczych kolumn i grup kolumn. Indeks **UNIQUE** jest nakładany na kolumny lub grupy kolumn w celu wymuszenia unikatowości wartości w nich danych. Indeks **częściowy (partial index)**, nakładany jedynie na części tabeli, jest stosowany, gdy wśród kwerend bazy danych często zdarzają się zapytania na określonym fragmencie danych.

Aby usunąć indeks, używa się polecenia **DROP INDEX**.

Indeksy można tworzyć i kasować również za pomocą narzędzi graficznych, np. pgAdmin III.

LITERATURA

- P. Domka, *Bazy danych i systemy baz danych*, WSiP, Warszawa 2013:
 - rozdział 37, s. 289 – Indeksy;
 - rozdział 38, s. 291 – Indeksy PostgreSQL.

SPRAWDŹ SWOJE UMIEJĘTNOŚCI

ZADANIE 1.

Skorzystaj z wyszukiwarki internetowej lub dokumentacji MySQL i znajdź informacje o składni poleceń związanych z zarządzaniem indeksami. Nawiąż połączenie z bazą MySQL i uruchom przykładową bazę **firma**. Utwórz index na kolumnie **nazwisko**. W edytorze tekstu wpisz odpowiednie informacje zgodnie z poniższą formatką. Zapisz dokument.

| | |
|--|--|
| Polecenie do utworzenia indeksu | |
| Polecenie do wyświetlenia listy indeksów | |
| Zrzuty ekranu potwierdzające wykonanie zadania | |

ZADANIE 2.

Skorzystaj z wyszukiwarki internetowej lub dokumentacji MySQL i znajdź informacje o składni poleceń związanych z zarządzaniem indeksami. Nawiąż połączenie z bazą MySQL i uruchom przykładową bazę **firma**. Utwórz index na kolumnie **nazwisko i imie**. W edytorze tekstu wpisz odpowiednie informacje zgodnie z poniższą formatką. Zapisz dokument.

| | |
|--|--|
| Polecenie do utworzenia indeksu | |
| Polecenie do wyświetlenia listy indeksów | |
| Zrzuty ekranu potwierdzające wykonanie zadania | |

ZADANIE 3.

Skorzystaj z wyszukiwarki internetowej lub dokumentacji PostgreSQL i znajdź informacje o składni poleceń związanych z zarządzaniem indeksami. Nawiąż połączenie z bazą PostgreSQL i uruchom przykładową bazę **firma**. Utwórz haszowany index na kolumnie **nazwisko**. W edytorze tekstu wpisz odpowiednie informacje zgodnie z poniższą formatką. Zapisz dokument.

| | |
|--|--|
| Polecenie do utworzenia indeksu | |
| Polecenie do wyświetlenia listy indeksów | |
| Zrzuty ekranu potwierdzające wykonanie zadania | |

ZADANIE 4.

Skorzystaj z wyszukiwarki internetowej lub dokumentacji PostgreSQL i znajdź informacje o składni poleceń związanych z zarządzaniem indeksami. Nawiąż połączenie z bazą PostgreSQL i uruchom przykładową bazę **firma**. Utwórz unikatowy index na kolumnie **nazwisko i imie**. W edytorze tekstu wpisz odpowiednie informacje zgodnie z poniższą formatką. Zapisz dokument.

| | |
|--|--|
| Polecenie do utworzenia indeksu | |
| Polecenie do wyświetlenia listy indeksów | |
| Zrzuty ekranu potwierdzające wykonanie zadania | |

ZADANIE 5.

Nawiąż połączenie z bazą PostgreSQL i uruchom przykładową bazę **firma**. Utwórz unikatowy index na kolumnie **nazwisko**. Wyjaśnij przyczynę niepowodzenia. Co zrobić, aby utworzyć unikatowy indeks? Odpowiedź zapisz w edytorze tekstu. Zapisz dokument.

Rozwiązania zadań zapisz w pliku pod nazwą **BD_39_nazwisko.doc**. Przedstaw do oceny nauczycielowi.

PODSUMOWANIE

TEST 39. Część pisemna egzaminu zawodowego**Zadanie 1.**

Indeksy stosuje się w celu

- A. podniesienia wydajności bazy danych podczas jej przeszukiwania.
- B. podniesienia wydajności bazy danych podczas wstawiania lub modyfikowania danych.
- C. ułatwienia kasowania rekordów.
- D. zapewnienia spójności danych.

Zadanie 2.

Który typ indeksu jest tworzony domyślnie?

- A. GIN.
- B. GIST.
- C. B-tree.
- D. haszowany.

Zadanie 3.

Indeks częściowy (*partial index*) jest nakładany

- A. jedynie na części tabeli.
- B. jedynie na wybrane kolumny.
- C. jedynie na wybrane kolumny klucza złożonego.
- D. jedynie na wybrane tabele.

Zadanie 4.

Nadmiarowy indeks powoduje

- A. zwiększenie szybkości przeglądania danych.
- B. zmniejszenie szybkości przeglądania danych.
- C. zwiększenie szybkości wprowadzania danych.
- D. zmniejszenie szybkości wprowadzania danych.

Zadanie 5.

Aby utworzyć indeks haszowany, używa się opcji

- A. USING hash.id_pracown.
- B. USING hash(id_pracown).
- C. HASHING(id_pracown).
- D. USING #(id_pracown).

ZADANIE EGZAMINACYJNE 1. Część praktyczna egzaminu zawodowego

Jesteś pracownikiem firmy zajmującej się tworzeniem oprogramowania, projektowaniem i wdrażaniem systemów baz danych. Do firmy zgłosił się klient, który od współpracownika otrzymał skrypt w języku SQL, tworzący bazę danych **firma**, oraz tabelę **pracownicy**, zawierającą kolumny: **id_pracown**, **nazwisko**, **imie**, **telefon**, **e-mail**). Baza zawiera bardzo dużo rekordów i klient chciałby przyspieszyć proces wyszukiwania danych. Najczęściej w tabeli wyszukiwane są rekordy na podstawie:

- nazwiska i imienia;
- adresu e-mail.

Twoim zadaniem jest:

- uruchomienie skryptu i utworzenie bazy danych i tabeli;
- wybór odpowiednich indeksów, które pozwolą na przyspieszenie przeszukiwania tabeli;
- utworzenie odpowiednich indeksów;
- sprawdzenie, czy indeksy zostały utworzone.

Wykonaj wszystkie polecenia na stanowisku wyposażonym w:

- serwer baz danych z zainstalowaną bazą danych MySQL i PostgreSQL;
- plik ze skryptem tworzącym bazę danych.

Rezultaty podlegające ocenie:

- uruchomienie skryptu i utworzenie bazy danych i tabeli;
- wybór odpowiednich indeksów, które pozwolą na przyspieszenie przeszukiwania tabeli;
- utworzenie odpowiednich indeksów;

PODSUMOWANIE

- sprawdzenie, czy indeksy zostały utworzone;
- przebieg prac zgodny z zasadami BHP, ergonomii i organizacji pracy.

Czas przeznaczony na wykonanie zadania wynosi 30 minut.

ZADANIE EGZAMINACYJNE 2. Część praktyczna egzaminu zawodowego

Jesteś pracownikiem firmy zajmującej się tworzeniem oprogramowania, projektowaniem i wdrażaniem systemów baz danych. Do firmy zgłosił się klient, który od współpracownika otrzymał skrypt w języku SQL tworzący bazę danych **firma** oraz tabelę **klienci**, zawierającą kolumny: **id_klienta**, **nazwisko**, **imie**, **telefon**, **e-mail**). Klient chciałby zabezpieczyć bazę przed wielokrotnym wprowadzaniem tego samego adresu e-mail i numeru telefonu, jednocześnie przyspieszając przeglądanie bazy.

Twoim zadaniem jest:

- uruchomienie skryptu i utworzenie bazy danych i tabeli;
- wybór odpowiednich indeksów, które zabezpieczą bazę przed wielokrotnym wprowadzaniem tego samego adresu e-mail i numeru telefonu, i pozwolą na przyspieszenie przeszukiwania tabeli;
- utworzenie odpowiednich indeksów;
- sprawdzenie, czy indeksy zostały utworzone;
- sprawdzenie, czy utworzone indeksy zabezpieczają bazę przed wielokrotnym wprowadzeniem tego samego adresu e-mail i numeru telefonu.

Wykonaj wszystkie polecenia na stanowisku wyposażonym w:

- serwer baz danych z zainstalowaną bazą danych MySQL i PostgreSQL;
- plik ze skryptem tworzącym bazę danych.

Rezultaty podlegające ocenie:

- uruchomienie skryptu i utworzenie bazy danych i tabeli;
- wybór odpowiednich indeksów, które zabezpieczą bazę przed wielokrotnym wprowadzeniem tego samego adresu e-mail i numeru telefonu, i pozwolą na przyspieszenie przeszukiwania tabeli;
- utworzenie odpowiednich indeksów;
- sprawdzenie, czy indeksy zostały utworzone;
- sprawdzenie, czy utworzone indeksy zabezpieczają bazę przed wielokrotnym wprowadzeniem tego samego adresu e-mail i numeru telefonu;
- przebieg prac zgodny z zasadami BHP, ergonomii i organizacji pracy.

Czas przeznaczony na wykonanie zadania wynosi 30 minut.

WNIOSKI