

40

Zasady optymalizacji baz danych

EFEKTY KSZTAŁCENIA Z PODSTAWY PROGRAMOWEJ:

- PKZ(E.b)(13) stosuje programy komputerowe wspomagające wykonywanie zadań;
- E.13.2(3) projektuje i tworzy relacyjne bazy danych;
- E.13.2(9) zarządza bazą danych i jej bezpieczeństwem.

W TYM ROZDZIALE:

- przypomnisz sobie, co to jest optymalizacja baz danych;
- dowiesz się, jakie problemy mogą się pojawić w związku z optymalizacją;
- poznasz, na czym polega testowanie bazy danych;
- dowiesz się, co to jest audyt bazy danych;
- zrozumiesz, na czym polega optymalizacja złączeń;
- poznasz ograniczenia i standardy baz danych.

Wprowadzenie

Optymalizacja bazy danych polega na utrzymaniu bazy danych w stanie, w którym może ona zapewnić maksymalną szybkość działania przy wykorzystaniu minimalnych zasobów sprzętowych. Podczas optymalizacji należy uwzględnić: budowę i charakterystykę sieci komputerowej, w której baza będzie pracować, przeznaczenie bazy i jej obciążenie liczbą zapytań i liczbą danych. Optymalizacja bazy danych dotyczy:

- przestrzeni dyskowych systemów plików;
- szybkości dostępu do danych podyktowanej również uwarunkowaniami sprzętowymi;
- wydajności kwerend;
- efektywności i prawidłowości ustalenia indeksów.

Na optymalną pracę bazy danych mają wpływ:

- odpowiednio ustawione indeksy;
- zoptymalizowane zapytania – kwerendy;
- wykorzystanie narzędzi służących do monitorowania i optymalizacji bazy danych.

Do częstych problemów związanych z optymalizacją bazy danych należą:

- problemy z połączeniami z bazą danych;
- długi czas wykonywania kwerend;
- problemy z serwerem oferującym usługi;
- błędy podczas planowania i projektowania bazy prowadzące do niewłaściwego określenia przeznaczenia bazy, tabel i kolumn w tabelach, szacowanego obciążenia bazy i sieci, w której będzie pracować, przepływu informacji w bazie danych, typu danych w kolumnach.

Optymalizacja bazy danych rozpoczyna się już na etapie jej projektowania. Dobranie odpowiedniego typu i rozmiaru pola ma wpływ na wydajność systemu oraz rozmiar danych zapisywanych na dyskach i przesyłanych przez sieć. Warto również stosować nazwy kolumn niezawierające spacji i znaków specjalnych.

Podczas projektowania bazy danych należy:

- określić, jakie dane koniecznie chcemy przechowywać;
- wyznaczyć klucze główne i klucze obce oraz związki pomiędzy tabelami;
- zdefiniować właściwości typów kolumn i relacji;
- określić przepływy informacji na podstawie związków między tabelami;
- przeprowadzić normalizację tabel i związków między nimi.

Optymalizacji podlegają również takie elementy projektu, jak:

- informacje, które będą umieszczane w bazie danych;
- generowane raporty;
- przepływy informacji: kto wprowadza informacje i jak często, kto gromadzi informacje, kto i jak często będzie generować raporty, gdzie będzie używana baza danych (serwer w sieci lokalnej czy internetowej).

Testowanie bazy danych polega na zebraniu informacji o zachowaniu serwera w trakcie normalnej pracy, zwykłego i skrajnego przeciążenia oraz pojawiających się błędów. Testy mają odpowiedzieć na pytanie, jak szybko SZBD może przetwarzać zapytania z zachowaniem stabilności pracy. Mogą być wykonywane różne rodzaje testów baz danych:

- testy obciążenia;
- testy wydajności;
- testy obciążenia próbnego;
- testy przeciążenia (test warunków skrajnych);
- testy pojemnościowe;
- długoterminowy test przeciążenia skrajnego;
- test przeciążenia pojedynczej funkcji.

Audyt bazy danych pozwala na uzyskanie szczegółowych informacji dotyczących dostępu do danych, które powinny podlegać szczególnej ochronie. Polega on na obserwowaniu pod kątem bezpieczeństwa operacji przeprowadzanych przez użytkowników bazy danych.

Złączenia mają postać tymczasowych związków między tabelami przez zapytanie odwołujące się do danych, które nie zostały połączone związkami. Złączenia udostępniają informacje pochodzące z dwóch lub więcej tabel. W złączeniu dwóch tabel A i B, jeżeli kwerenda odczytuje jeden raz tabelę A, to tabela B jest odczytana tyle razy, ile krotek ma tabela A.

Integralność bazy danych gwarantuje, że dane umieszczone w bazie mają formę odpowiadającą rzeczywistości i spójną logicznie. Typy integralności bazy danych:

- **integralność encji** – gwarantuje, że wewnątrz tabeli nie mogą wystąpić dwie identyczne krotki;
- **integralność domeny** – gwarantuje, że wewnątrz kolumny mogą wystąpić wyłącznie dane, których typ jest taki sam, jak zdefiniowany typ kolumny;
- **integralność referencyjna** – opiera się na zachowaniu powiązanych informacji w tabeli połączonej za pomocą związku.

Powiązania między tabelami wymuszają integralność referencyjną, co oznacza, że nie można do jednej tabeli dodać rekordu, jeżeli nie będzie powiązanego z nim rekordu w drugiej tabeli.

Ograniczenia to reguły definiujące, jakie dane są poprawne dla pola kolumny. Klucz główny jednoznacznie identyfikuje krotki w tabeli. Klucz obcy to kolumna lub zbiór kolumn, których wartości są takie same, jak wartości klucza głównego w tabeli powiązanej. Klucz obcy określa, że dane umieszczone w nim muszą zgadzać się z danymi w kluczu głównym tabeli powiązanej. Kolumny będące kluczem głównym i obcym nie mogą zawierać wartości NULL. Klucz kandydujący ma pola, które się nie powtarzają i jest rozważany pod kątem użycia go jako klucza głównego.

Standardy pozwalają na zachowanie przejrzystości bazy danych i dotyczą takich elementów, jak:

- **kod** – określa, że dla wszystkich słów kluczowych SQL powinno się stosować wielkie litery, np. SELECT, DROP, ALTER, UPDATE, DELETE, natomiast dla identyfikatorów (nazw tabel, kolumn, baz danych) – małe litery;
- **nazewnictwo kolumn** – klucz główny tabeli powinien zawierać ID; należy unikać stosowania spacji w identyfikatorach (nazwach kolumn, tabel, baz danych) oraz słów kluczowych zarezerwowanych w języku SQL;
- **tabele** – każda z tabel powinna mieć klucz główny.

LITERATURA

- P. Domka, *Bazy danych i systemy baz danych*, WSiP, Warszawa 2013:
– rozdział 36, s. 275 – *Optymalizacja bazy danych*.

SPRAWDŹ SWOJE UMIEJĘTNOŚCI

ZADANIE 1.

Skorzystaj z wyszukiwarki internetowej i znajdź informacje o dwóch przykładowych firmach oferujących usługi audytu baz danych. W edytorze tekstu wpisz odpowiednie informacje zgodnie z poniższą formatką. Zapisz dokument.

Nazwa firmy	
Adres strony www	

SPRAWDŹ SWOJE UMIEJĘTNOŚCI

ZADANIE 2.

Przeprowadź testy pojemnościowe zainstalowanej na twoim komputerze bazy MySQL i PostgreSQL. W edytorze tekstu wpisz odpowiednie informacje zgodnie z poniższą formatką. Zapisz dokument.

	MySQL	PostgreSQL
Wersja systemu operacyjnego komputera		
Wersja bazy danych		
Maksymalna liczba jednocześnie zalogowanych użytkowników		

ZADANIE 3.

Podczas nauki korzystałeś z przykładowej bazy danych, np. **firma**. Zastanów się, jakie zapytania były wykonywane najczęściej. Zaproponuj przykładowe indeksy, które przyspieszą przeglądanie bazy danych. W edytorze tekstu wpisz odpowiednie informacje zgodnie z poniższą formatką. Zapisz dokument.

Nazwa tabeli	Kolumna / kolumny kluczowe

ZADANIE 4.

Dokonaj analizy wybranego połączenia przykładowej bazy danych, np. **firma**. W edytorze tekstu wpisz odpowiednie informacje zgodnie z poniższą formatką. Zapisz dokument.

	Tabela 1	Tabela 2
Nazwa tabeli		
Klucz główny		
Klucz obcy		
Typ danych klucza głównego		
Typ danych klucza obcego		
Typ relacji między tabelami		
Czy nazwa tabeli zgodna ze standardem		
Czy nazwy kolumn zgodne ze standardem		

ZADANIE 5.

Wyjaśnij własnymi słowami, na czym polega różnica między typem danych char i varchar. Odpowiedź zapisz w edytorze tekstu. Zapisz dokument.

Rozwiązania zadań zapisz w pliku pod nazwą **BD_40_nazwisko.doc**. Przedstaw do oceny nauczycielowi.

PODSUMOWANIE

TEST 40. Część pisemna egzaminu zawodowego**Zadanie 1.**

Utrzymanie bazy danych w stanie, w którym może zapewnić maksymalną szybkość działania przy wykorzystaniu minimalnych zasobów sprzętowych, to

- A. projektowanie.
- B. eksploatacja.
- C. optymalizacja.
- D. konserwacja.

Zadanie 2.

Który typ integralności gwarantuje, że wewnątrz tabeli nie mogą wystąpić dwie identyczne krotki?

- A. Encja.
- B. Domena.
- C. Referencyjność.
- D. Wszystkie wymienione w p. A–C.

Zadanie 3.

Który rodzaj testu wykazuje, jak dużo zapytań serwer może przetworzyć w określonym czasie?

- A. Test wydajności.
- B. Test obciążenia próbnego.
- C. Test przeciążenia.
- D. Test pojemnościowy.

Zadanie 4.

Uzyskanie szczegółowych informacji dotyczących dostępu do danych, które powinny podlegać szczególnej ochronie, to

- A. test wydajności.
- B. audyt.
- C. polityka bezpieczeństwa.
- D. ochrona danych.

Zadanie 5.

Jednoznacznie identyfikuje krotki (wiersze) w tabeli

- A. klucz prosty.
- B. klucz główny.
- C. klucz obcy.
- D. klucz złożony.

ZADANIE EGZAMINACYJNE 1. Część praktyczna egzaminu zawodowego

Nauczyciel informatyki z zaprzyjaźnionego liceum poprosił cię o zaprojektowanie bazy danych przechowującej informacje o klasach i uczniach szkoły. Baza powinna zawierać informacje pozwalające na określenie, do której klasy uczęszcza dany uczeń oraz:

- nazwisko i imię ucznia;
- datę urodzenia ucznia;
- płeć;
- rok szkolny utworzenia klasy;
- profil nauczania.

Twoim zadaniem jest:

- zaprojektowanie struktury bazy danych zgodnie ze standardami;
- dobór typów danych dla poszczególnych kolumn;
- dobór kluczy w tabelach;
- napisanie skryptu tworzącego bazę, tabele i indeksy, które zabezpieczą bazę przed wielokrotnym wprowadzeniem tego samego adresu e-mail i numeru telefonu oraz pozwolą na przyspieszenie przeszukiwania tabeli;
- wprowadzenie przykładowych danych (co najmniej dwa rekordy w każdej tabeli).

Wykonaj wszystkie polecenia na stanowisku wyposażonym w serwer baz danych z zainstalowaną bazą danych MySQL i PostgreSQL.

PODSUMOWANIE

Rezultaty podlegające ocenie:

- wybór odpowiednich indeksów, które zabezpieczą bazę przed wielokrotnym wprowadzeniem tego samego adresu e-mail i numeru telefonu, i pozwolą na przyspieszenie przeszukiwania tabeli;
- zaprojektowanie struktury bazy danych zgodnie ze standardami;
- poprawny dobór typów danych dla poszczególnych kolumn;
- poprawny dobór kluczy w tabelach;
- napisanie skryptu tworzącego bazę i tabelę;
- uruchomienie skryptu i utworzenie bazy danych i tabeli;
- wprowadzenie przykładowych danych (co najmniej dwa rekordy w każdej tabeli);
- przebieg prac zgodny z zasadami BHP, ergonomii i organizacji pracy.

Czas przeznaczony na wykonanie zadania wynosi 90 minut.

ZADANIE EGZAMINACYJNE 2. Część praktyczna egzaminu zawodowego

Jesteś pracownikiem firmy zajmującej się tworzeniem oprogramowania, projektowaniem i wdrażaniem systemów baz danych. Otrzymałeś zadanie polegające na przeprowadzeniu optymalizacji bazy danych. Baza danych zawiera trzy tabele:

- dane czytelników:
 - nazwisko;
 - imię;
 - adres;
- dane książek:
 - tytuł;
 - autor;
 - rok wydania;
 - wydawnictwo;
- wypożyczenia:
 - data wypożyczenia;
 - czytelnik;
 - książka.

Twoim zadaniem jest:

- dobranie nazw tabel i kolumn w tabelach;
- dobranie kluczy głównych i obcych w tabelach;
- utworzenie bazy danych i tabel;
- utworzenie kluczy w tabelach;
- utworzenie relacji między tabelami;
- utworzenie indeksu przyspieszającego wyszukiwanie czytelnika na podstawie nazwiska i imienia.

Wykonaj wszystkie polecenia na stanowisku wyposażonym w serwer baz danych z zainstalowaną bazą danych MySQL i PostgreSQL.

Rezultaty podlegające ocenie:

- dobranie nazw tabel i kolumn w tabelach;
- dobranie kluczy głównych i obcych w tabelach;
- utworzenie bazy danych i tabel;
- utworzenie kluczy w tabelach;
- utworzenie relacji między tabelami;
- utworzenie indeksu przyspieszającego wyszukiwanie czytelnika na podstawie nazwiska i imienia;
- przebieg prac zgodny z zasadami BHP, ergonomii i organizacji pracy.

Czas przeznaczony na wykonanie zadania wynosi 90 minut.

WNIOSKI