

Optymalizacja zapytań

Czas: 2 × 45 min

praca indywidualna

Cel: Znalezienie wolnych zapytań za pomocą dziennika i planu wykonania oraz przepisanie ich na szybsze odpowiedniki: stronicowanie bez OFFSET, złączenie zamiast podzapytania skorelowanego, jedno zapytanie zamiast N+1.

Po zajęciach potrafisz:

- włącza dziennik wolnych zapytań i odczytuje z niego czas oraz liczbę przeczytanych wierszy
- porównuje plan szacowany (EXPLAIN) z rzeczywistym (ANALYZE / EXPLAIN ANALYZE)
- zastępuje stronicowanie przez OFFSET stronicowaniem po kluczu (keyset)
- przepisuje podzapytanie skorelowane na złączenie i uzasadnia różnicę czasu
- rozpoznaje problem N+1 w kodzie PHP i usuwa go jednym zapytaniem z JOIN

Sprzęt i materiały

- komputer z XAMPP (MariaDB 10.4, PHP 8) lub MySQL 8.4
- baza `demo` z zadania 15–16 i baza `sklep`
- plik `r2_17_zamowienia_n1.php` z folderu `Pliki_do_zadan`

Oprogramowanie

- **XAMPP — konsola mysql, Apache i PHP** — dziennik wolnych zapytań w `C:\xampp\mysql\data`
apachefriends.org/download.html
- **VS Code** — edycja skryptu PHP i podgląd pliku dziennika
code.visualstudio.com

Przygotowanie przed zajęciami

- Sprawdź, że baza `demo` ma indeks `idx_strona_czas` (strona, czas) z zadania 16. Jeśli nie — zaimportuj `r2_15_demo.sql` i utwórz go.
- Przypomnij sobie z prezentacji 2.10 kolumny EXPLAIN i pojęcie problemu N+1.

Bezpieczeństwo

- Zmiany `SET GLOBAL` dotyczą całego serwera — wykonuj je tylko na swoim XAMPP i na koniec przywróć wartości domyślne (dziennik wyłączony, `long_query_time = 10`).
- Dziennik wolnych zapytań zapisuje pełną treść zapytań (także dane klientów). Na serwerze produkcyjnym chroń ten plik i nie wysyłaj go nikomu.
- Skrypt PHP uruchamiaj tylko lokalnie (`localhost`) — zawiera dane logowania do bazy.

Przebieg zadania

1 Dziennik wolnych zapytań

W MariaDB nowa wartość `slow_query_log` obowiązuje dopiero w nowych połączeniach — po `SET GLOBAL` wyjdź z konsoli (`exit`) i połącz się ponownie. W MySQL 8.4 zmiana działa od razu.

1. Z1. Wyświetl zmienne `slow_query_log`, `long_query_time`, `slow_query_log_file` i `log_output` (`SHOW VARIABLES WHERE Variable_name IN (...)`).
2. Z2. Włącz dziennik (`slow_query_log = 1`), ustaw zapis do pliku i tabeli (`log_output = 'FILE, TABLE'`) oraz próg `long_query_time = 0.1` s. Połącz się ponownie i sprawdź wartości.
3. Z3. Wykonaj w bazie `demo` trzy zapytania: `SELECT SLEEP(0.3);`, liczenie odwiedzin z 1 marca warunkiem `DATE(czas) = '2026-03-01'` oraz to samo zapytanie z zakresem czasu. Odczytaj tabelę `mysql.slow_log` (`start_time`, `query_time`, `rows_examined`, `sql_text`) i otwórz plik dziennika w edytorze. Które zapytania trafiły do dziennika?

2 Plan szacowany i rzeczywisty

MariaDB: `ANALYZE SELECT ...` (kolumny `r_rows`, `r_filtered`). MySQL 8.4: `EXPLAIN ANALYZE SELECT ...` (actual rows, actual time).

1. Z4. Wykonaj `ANALYZE` (lub `EXPLAIN ANALYZE`) dla zapytania liczącego odwiedzinę strony `'/koszyk'` z 1 marca. Porównaj `rows` z `r_rows`.
2. Z5. Wykonaj `ANALYZE` dla liczenia odwiedzin z `czas_ms BETWEEN 100 AND 110`. Ile wierszy serwer przeczytał (`r_rows`), a jaki procent przeszedł przez warunek (`r_filtered`)? Co to mówi o potrzebie indeksu?
3. Z6. Porównaj czas i `EXPLAIN` zapytań `SELECT *` oraz `SELECT id, czas` dla strony `'/koszyk'` w marcu (po trzy pomiary). Wyjaśnij różnicę w kolumnie Extra.

3 Stronicowanie i podzapytania

1. Z7. Wyświetl „stronę” 10 odwiedzin od pozycji 900 001 w kolejności id przez `LIMIT 10 OFFSET 900000`. Zmierz czas i pokaż `EXPLAIN (rows)`.
2. Z8. Zapisz to samo jako stronicowanie po kluczu: warunek `id > ostatnie_id_z_poprzedniej_strony` i `LIMIT 10`. Porównaj czas i plan.
3. Z9. Stronicowanie po czasie: kolejne 10 odwiedzin strony `'/koszyk'` po chwili `2026-06-01 12:00:00` (`ORDER BY czas`). Pokaż `EXPLAIN` — który indeks i jaki typ dostępu?
4. Z10. Podzapytanie skorelowane: dla odwiedzin strony `'/koszyk'` z 1 marca między 0:00 a 1:00 wyświetl id, id_klienta i liczbę wszystkich odwiedzin tego klienta (podzapytanie w liście `SELECT`). Zmierz czas i pokaż `EXPLAIN (DEPENDENT SUBQUERY)`.
5. Z11. Przepisz Z10 na złączenie z tabelą pochodną (`SELECT id_klienta, COUNT(*) ... GROUP BY id_klienta`). Sprawdź, że wynik jest taki sam, i porównaj czas.
6. Z12. W bazie `sklep` napisz dwiema metodami (podzapytanie w `SELECT` oraz `LEFT JOIN` z `GROUP BY`) liczbę zamówień każdego klienta, także klientów bez zamówień. Porównaj wyniki i plany.

4 Problem N+1 w aplikacji

1. Z13. Uruchom `r2_17_zamowienia_n1.php` (w przeglądarce z `htdocs` albo w konsoli `php r2_17_zamowienia_n1.php`). Zapisz liczbę zapytań i czas. Wskaż w kodzie, skąd bierze się ta liczba zapytań.
2. Z14. Zapisz kopię jako `r2_17_zamowienia_join.php` i zastąp pętlę z zapytaniami jednym zapytaniem z `JOIN` i `GROUP BY` (klient i wartość zamówienia). Porównaj wynik, liczbę zapytań i czas.
3. Z15. Wyłącz dziennik wolnych zapytań i przywróć `long_query_time = 10` oraz `log_output = 'FILE'`. Zapisz polecenia SQL w pliku `r2_17_optymalizacja.sql`.

W sprawozdaniu umieść

- ☐ tabela zapytań Z1–Z15 z wynikami i czasami
- ☐ zrzut ekranu tabeli `mysql.slow_log` lub fragment pliku dziennika z Z3
- ☐ porównanie OFFSET i keyset (czas, rows) oraz podzapytania i złączenia (czas)
- ☐ kod zapytania z `r2_17_zamowienia_join.php` i porównanie liczby zapytań
- ☐ pliki `r2_17_optymalizacja.sql` i `r2_17_zamowienia_join.php`
- ☐ odpowiedzi na pytania kontrolne

Pytania kontrolne

1. Dlaczego `LIMIT 10 OFFSET 900000` jest wolne, choć zwraca tylko 10 wierszy?
2. Czym różnią się kolumny rows i r_rows? Kiedy duża różnica jest sygnałem problemu?
3. Dlaczego podzapytanie skorelowane bez indeksu potrafi być tysiące razy wolniejsze od złączenia?
4. Co to jest problem N+1 i dlaczego jest groźny, gdy baza działa na innym serwerze niż PHP?

Kryteria oceny

Kryterium	Punkty
Dziennik wolnych zapytań — włączenie, odczyt, przywrócenie (Z1–Z3, Z15)	4
Analiza planów rzeczywistych i SELECT * (Z4–Z6)	3
Stronicowanie keyset i przepisanie podzapytań (Z7–Z12)	6
Usunięcie problemu N+1 w PHP (Z13–Z14)	4
Odpowiedzi na pytania kontrolne	3
Razem	20

Sprawozdanie przygotuj na wzorze dołączonym do zadania (plik .docx o tym samym numerze).