

## Streszczenie pracy doktorskiej

mgr Anny Sasak-Okoń

### **p.t. Spekulatywne zrównoleglenie jako technika wspomagania przetwarzania kwerend w relacyjnych bazach danych.**

Niniejsza rozprawa została zainspirowana ideą wykonania spekulatywnego, w której system komputerowy wykonuje pewne zadania zanim stwierdzona jest możliwość ich naturalnego, równoległego wykonania. Jeżeli do działającego Systemu Zarządzania Bazą Danych (SZBD) spływają z wielu źródeł kwerendy, w takiej liczbie, że formują kolejkę przekraczając możliwości SZBD wykonywania kwerend, to jest to miejsce na wprowadzenie działania Spekulatywnego. Jeżeli w tej kolejce narzucona jest pewna kolejność kwerend, wówczas wykonując pewne kwerendy lub ich fragmenty z wyprzedzeniem, uzyskuje się analogię do Spekulacji Przepływu Sterowania. Jeżeli dodatkowo w kolejce mogą się znajdować kwerendy wpływające na Stan Danych w bazie poprzez ich modyfikację, można mówić o Spekulacji Wartości Danych. Co więcej, charakterystyka zapytań nadchodzących do baz danych, określanych mianem „sklepowych”, odgrywa bardzo dużą rolę. Bazy takiego typu posiadają wiele danych produktowych przeglądanych przez użytkowników. Swoje zapytania budują oni za pośrednictwem udostępnionych aplikacji, narzucając własne kryteria przeglądania. W ten sposób, dana baza otrzymuje zawsze kwerendy o określonej budowie, różniące się jedynie wartościami nakładanych warunków. Wszystkie podobieństwa między zakolejkowanymi kwerendami przemawiają na korzyść Spekulacji Wartości Danych.

Na podstawie tych obserwacji, w niniejszej pracy postawiono sobie za cel zbudowanie Modelu Wykonania Spekulatywnego Kwerend SQL, który dla kolejki kwerend oczekujących na wykonanie przez SZBD, poprzez odpowiednio dobrane Kwerendy Spekulatywne, skróci czas wykonywania poszczególnych kwerend, a co za tym idzie także całkowity czas oczekiwania na odpowiedź od SZBD. Na podstawie reprezentacji grafowych poszczególnych kwerend nadchodzących do bazy danych, oraz grup tych kwerend przeprowadzania jest Analiza Spekulatywna. Jej celem jest wygenerowanie, a następnie wykonanie zbioru Kwerend Spekulatywnych. Wyniki wykonanych Kwerend Spekulatywne zapisywane są w formie relacji w Bazie Wyników Spekulacji rezydującej w pamięci podręcznej. Wykonywana Kwerenda Wejściowa, zamiast do bazy danych na dysku, może odwołać się do wyników Kwerendy Spekulatywnej z Bazy Wyników Spekulacji, uzyskując znacznie krótszy czas dostępu do danych oraz krótszy czas wykonania, gdyż zwykle wynik spekulacji będzie zawierał mniejszą liczbę danych do przeszukania niż podstawowa relacja w bazie na dysku. Całościowa

struktura pracy obejmuje osiem rozdziałów o następującej treści. Rozdział 1 to wprowadzenie w tematykę obliczeń równoległych w ogólności. Rozdziały 2 i 3 poświęcone zostały Relacyjnym Bazom Danych i zagadnieniom związanych z wprowadzaniem w nich wykonania równoległego. Rozdział 4 został poświęcony obliczeniom spekulatywnym w ogólności oraz w Relacyjnych Bazach Danych. Przedstawiona została ogólna idea, która leży u podstaw wykonania instrukcji w sposób spekulatywny. Dokonano także przeglądu literatury, który pokazuje w jaki sposób obliczenia spekulatywne znajdują zastosowanie w optymalizacji procesów bazodanowych.

Rozdział 5 opisuje motywacje, którymi kierowano się przy realizacji niniejszej pracy, główną ideę i ogólne założenia dla opracowanego modelu wspomaganego wykonania kwerend w relacyjnych bazach danych za pomocą programowej Warstwy Spekulatywnej. W rozdziale tym postawiono tezę badawczą i nakreślono główne kierunki prac zmierzające do realizacji i weryfikacji tej tezy.

Rozdział 6 to zasadnicza część pracy. Rozpoczyna się ona od wprowadzenia podstawowych pojęć wykorzystywanych w dalszej części pracy, związanych z Systemem Zarządzania Bazą Danych i budową kwerend SQL. Następnie określono budowę obsługiwanych zapytań SQL zgodną z wymaganiami SQLite, tak aby wyrażała większość typowych zapytań do Bazy Danych sklepu internetowego.

Dla relacyjnej bazy danych i pojedynczych kwerend o założonej budowie, oraz w oparciu o literaturę, sformułowano zbiór reguł wykorzystywanych do tworzenia ich reprezentacji grafowych — nazywanych grafami RAW (Relacja—Atrybut—Wartość). Następnie opracowano sposób reprezentacji zbiorów kwerend za pomocą multigrafu, powstającego przez łączną projekcję grafów poszczególnych kwerend ze zbioru.

Zdefiniowano i opisano trzy typy Kwerend Spekulatywnych: typu Parametr Spekulatywny, typu Podgraf Spekulatywny i typu Spekulatywny Stan Danych. Określono sposób reprezentacji kwerend spekulatywnych zdefiniowanych typów w multigrafie kwerend za pomocą krawędzi spekulatywnych oraz opisano strategię wprowadzania tych krawędzi.

Zaprezentowano schemat budowy i komunikacji programowej Warstwy Spekulatywnej realizującej spekulatywny model wspomaganego wykonania kwerend. Jest to postać ogólna jednofazowego algorytmu wykonania spekulatywnego, która dla grup kwerend, określanych Oknem Spekulacji, dokonuje analizy ich reprezentacji grafowej, na podstawie której generowane są Kandydujące Kwerendy Spekulatywne. W tej wersji algorytmu generowane są wszystkie możliwe Kwerendy Spekulatywne wynikające z analizy spekulatywnej multigrafu

kwerend dla danego Okna Spekulacji. To ile z nich zostanie wykonane ograniczone jest jedynie ilością dostępnej mocy obliczeniowej.

Na końcu rozdziału zdefiniowano zbiór miar służących do weryfikacji jakości wykonywanych Kwerend Spekulatywnych i opisano strategię zarządzania Bazą Wyników Spekulacji przechowującą wyniki wykonanych Kwerend Spekulatywnych.

Rozdział 7 pokazuje na wstępie weryfikację eksperymentalną Spekulatywnego Modelu Przetwarzania Kwerend, wdrożonego jako Algorytm Jednofazowy, za pomocą Warstwy Spekulatywnej. Na podstawie uzyskanych wyników rozważono możliwość ograniczenia liczby generowanych kwerend spekulatywnych, gdyż wykonywanie Kwerend Spekulatywnych przy ograniczonej liczbie zasobów obliczeniowych pozostawiało dużą ich część bez wykonania i wykorzystania. Zaprezentowano ideę ulepszanego, Dwufazowego Algorytmu Spekulatywnego Wspomagania Wykonania Kwerend. W tym algorytmie bardzo silnie ograniczono liczbę generowanych Kwerend Spekulatywnych generowanych z Wyprzedzeniem dla Okna Spekulacji, uzyskując możliwość elastycznego wykorzystania mocy obliczeniowej. Uwolnione zasoby wykorzystano do wykonywania odpowiednio dobranych, bardziej ogólnych Kwerend Spekulatywnych generowanych na Żądanie. Wprowadzony nowy typ kwerend generowanych na żądanie, wykorzystuje zarejestrowane informacje o charakterystyce dotychczasowo zrealizowanych Kwerend Wejściowych z użyciem wyników Kwerend Spekulatywnych.

Po opisie ulepszanego Dwufazowego Algorytmu Wspomagania Wykonania Kwerend, ponownie przeprowadzono szereg badań eksperymentalnych weryfikując m.in. liczbę wykonań Kwerend Wejściowych z wykorzystaniem wyników Kwerend Spekulatywnych z Wyprzedzeniem oraz na Żądanie. Dokonano też przeglądu wykonanych Kwerend Spekulatywnych obu rodzajów z uwzględnieniem ich budowy i wartości zdefiniowanych miar, weryfikujących ich jakość wykonania i stopień wykorzystania.

W rozdziale 8 podsumowano wyniki uzyskane w ramach niniejszej rozprawy i nakreślono kierunki dalszych badań.