

Itzik Ben-Gan
Dejan Sarka
Adam Machanic
Kevin Farlee

Zapytania w języku T-SQL

w Microsoft SQL Server 2014
i SQL Server 2012

Przekład: Natalia Chounlamany
Marek Włodarz

APN Promise, Warszawa 2015

Spis treści

Przedmowa	xv
Wstęp	xvii
1 Logiczne przetwarzanie zapytań	1
Fazy logicznego przetwarzania zapytań	3
Krótkie omówienie faz logicznego przetwarzania zapytania	4
Przykładowe zapytanie oparte na scenariuszu z użyciem tabeli klientów i zamówień	7
Szczegółowe omówienie faz logicznego przetwarzania zapytania	9
Krok 1: Faza FROM	9
Krok 2: Faza WHERE	16
Krok 3: Faza GROUP BY	17
Krok 4: Faza HAVING	19
Krok 5: Faza SELECT	19
Krok 6: Faza ORDER BY	23
Krok 7: Zastosowanie filtra TOP lub OFFSET-FETCH	25
Pozostałe aspekty logicznego przetwarzania zapytań	30
Operatory tabeli	30
Funkcje okna	40
Operatory UNION, EXCEPT oraz INTERSECT	42
Podsumowanie	44
2 Optymalizowanie zapytań	45
Struktury wewnętrzne	46
Strony i fragmenty	46
Struktura tabel	48
Narzędzia do mierzenia wydajności zapytań	60
Metody dostępu	65
Skanowanie tabeli/nieuporządkowane skanowanie indeksu klastrowego	65
Nieuporządkowane skanowanie pokrywającego indeksu nieklastrowego	68
Uporządkowane skanowanie indeksu klastrowego	70
Uporządkowane skanowanie pokrywającego indeksu nieklastrowego	72
Skanowanie w wykonaniu aparatu magazynu	74
Przeszukanie indeksu nieklastrowego + skanowanie zakresu + wyszukania	92
Nieuporządkowane skanowanie indeksu nieklastrowego + operacje wyszukania	102

Operacja przeszukania indeksu klastrowego + skanowanie zakresu	105
Przeszukanie pokrywającego indeksu nieklastrowego + skanowanie zakresu	107
Szacowanie liczebności	110
Porównanie wersji komponentu do szacowania liczebności	111
Konsekwencje niedoszacowań i przeszacowań	112
Statystyki	115
Szacowania dla wielu predykatów	118
Problem rosnącego klucza	122
Niewiadome	125
Funkcje indeksowania	131
Indeksy malejące	131
Kolumny dołączone	135
Filtrowane indeksy oraz statystyki	136
Indeksy magazynu kolumn	139
Wbudowana definicja indeksu	148
Wybieranie zapytań do optymalizacji przy użyciu zdarzeń rozszerzonych	149
Informacje i statystyki dotyczące indeksów oraz zapytań	153
Obiekty tymczasowe	158
Porównanie rozwiązań bazujących na zbiorach i iteracji	170
Dostrajanie zapytań poprzez ich korektę	175
Równoległe wykonanie zapytania	180
Jak działa równoległe wykonywanie zapytania	181
Równoległość a optymalizacja zapytań	199
Wzorzec zapytania z równoległe wykonywaną operacją APPLY	207
Podsumowanie	212
3 Zapytania złożone	213
Podzapytania	213
Podzapytania niezależne	214
Podzapytania skorelowane	216
Predykat EXISTS	222
Niepoprawne podzapytania	229
Wyrażenia tabeli	233
Tabele pochodne	234
Wspólne wyrażenia tabeli	237
Widoki	242
Wbudowane funkcje zwracające tabele	245
Generowanie liczb	246
Operator APPLY	250
Operator CROSS APPLY	250
Operator OUTER APPLY	252
Niejawny operator APPLY	253

Wielokrotne wykorzystywanie aliasów kolumn	254
Złączenia	256
Złączenie krzyżowe (Cross Join)	256
Złączenie wewnętrzne	261
Złączenie zewnętrzne	263
Samozłączenie	264
Złączenia równościowe i nierównościowe	264
Zapytania z wieloma złączeniami	265
Złączenia oraz antyzłączenia częściowe	271
Algorytmy złączenia	273
Rozdzielanie elementów	281
Operatory UNION, EXCEPT oraz INTERSECT	285
Operatory UNION ALL oraz UNION	286
Operator INTERSECT	289
Operator EXCEPT	291
Podsumowanie	293
4 Grupowanie i przestawianie danych oraz funkcje okna	295
Funkcje okna	295
Agregujące funkcje okna	296
Rankingowe funkcje okna	319
Funkcje okna przesunięcia	324
Statystyczne funkcje okna	326
Luki i wyspy	330
Przestawianie danych	339
Przestawianie danych jeden-do-jednego	340
Przestawianie danych wiele-do-jednego	344
Odwrotne przestawianie danych	348
Odwrotne przestawianie danych przy użyciu CROSS JOIN oraz VALUES	349
Odwrotne przestawianie danych przy użyciu CROSS APPLY oraz VALUES	351
Zastosowanie operatora UNPIVOT	353
Niestandardowe agregacje	354
Wykorzystanie kursora	355
Wykorzystanie operacji przestawiania danych	357
Specjalizowane rozwiązania	358
Zestawy grupowania	370
Podklauzula GROUPING SETS	371
Klauzule CUBE oraz ROLLUP	375
Algebra zestawów grupowania	377
Materializowanie zestawów grupowania	378
Sortowanie	381
Podsumowanie	383

5 Filtry TOP i OFFSET-FETCH	385
Filtry TOP oraz OFFSET-FETCH	385
Filtr TOP	385
Filtr OFFSET-FETCH	389
Optymalizacja filtrów na przykładzie stronicowania	391
Optymalizacja filtra TOP	391
Optymalizacja filtra OFFSET-FETCH	399
Optymalizacja funkcji ROW_NUMBER	403
Wykorzystanie opcji TOP w modyfikacjach	406
TOP w modyfikacjach	406
Modyfikacje fragmentaryczne	407
Pierwszych N z każdej grupy	409
Rozwiązanie bazujące na funkcji ROW_NUMBER	411
Rozwiązanie oparte na klauzulach TOP oraz APPLY	412
Rozwiązanie bazujące na łączeniu (sortowanie z przenoszeniem)	413
Mediana	415
Rozwiązanie wykorzystujące funkcję PERCENTILE_CONT	417
Rozwiązanie wykorzystujące funkcję ROW_NUMBER	417
Rozwiązanie wykorzystujące klauzule OFFSET-FETCH oraz APPLY	418
Podsumowanie	420
6 Modyfikowanie danych	421
Wstawianie danych	421
SELECT INTO	421
Import zbiorczy	424
Mierzenie ilości rejestrowanych danych	426
Dostawca zbiorczych zestawów wierszy	428
Sekwencje	430
Cechy charakterystyczne i ograniczenia właściwości tożsamości	431
Obiekt sekwencji	432
Względy wydajnościowe	437
Podsumowanie porównania tożsamości z sekwencją	446
Usuwanie danych	447
TRUNCATE TABLE	447
Usuwanie duplikatów	451
Aktualizowanie danych	453
Aktualizowanie przy użyciu wyrażeń tabeli	454
Aktualizowanie z wykorzystaniem zmiennych	455
Scalanie danych	457
Przykłady zastosowania instrukcji MERGE	457
Zapobieganie konfliktom instrukcji MERGE	461

ON to nie filtr	462
USING przypomina FROM	464
Klauzula OUTPUT	465
Przykład z instrukcją INSERT i tożsamością	465
Przykład archiwizacji usuwanych danych	467
Przykład z instrukcją MERGE	468
Funkcja Composable DML	471
Podsumowanie	472
7 Przetwarzanie danych typu data i czas	473
Typy danych daty i czasu	473
Funkcje daty i czasu	477
Nowe funkcje daty i czasu	487
Wyzwania związane z przetwarzaniem daty i czasu	490
Literały	491
Identyfikowanie dni tygodnia	494
Obsługiwanie danych samej daty lub samego czasu przy użyciu typów DATETIME oraz SMALLDATETIME	497
Obliczanie pierwszej, ostatniej i kolejnej daty	498
Argumenty wyszukiwania	503
Problemy zaokrągleń	505
Zapytania dotyczące dat i czasu	507
Grupowanie według tygodni	507
Interwały	509
Podsumowanie	534
8 T-SQL dla praktyków BI	535
Przygotowywanie danych	536
Widok analizy sprzedaży	537
Częstości	538
Częstości bez użycia funkcji okna	538
Częstości z wykorzystaniem funkcji okna	539
Statystyki opisowe dla zmiennych ciągłych	542
Centra rozkładu	542
Rozproszenie rozkładu	546
Wyższe momenty populacji	551
Zależności liniowe	560
Dwie ciągłe zmienne	561
Tablice liczebności i chi-kwadrat	568
Analiza wariancji	573
Całkowanie oznaczone	576
Średnie ruchome i entropia	580

Średnie ruchome.....	580
Entropia.....	587
Podsumowanie.....	591
9 Obiekty programowalne.....	595
Dynamiczny kod SQL.....	595
Korzystanie z polecenia EXEC.....	596
Korzystanie z procedury składowanej <i>sp_executesql</i>	600
Dynamiczne przestawianie danych.....	601
Dynamiczne warunki wyszukiwania.....	606
Dynamiczne sortowanie.....	614
Funkcje definiowane przez użytkownika.....	619
Skalarne UDF.....	619
Wielowyrażeniowe funkcje tabeli.....	624
Procedury składowane.....	626
Kompilacje, rekompilacje i ponowne użycie planów wykonania.....	627
Typ tabeli i parametry o wartościach tabeli.....	647
EXEC ... WITH RESULT SETS.....	650
Wyzwalacze.....	653
Typy i stosowanie wyzwalaczy.....	653
Wydajne programowanie wyzwalaczy.....	659
Programowanie SQLCLR.....	664
Architektura SQLCLR.....	665
Skalarne funkcje CLR i tworzenie naszej pierwszej asemblacji.....	668
Strumieniowe funkcje o wartościach tabeli.....	679
Procedury składowane i wyzwalacze SQLCLR.....	687
Typy definiowane przez użytkownika w SQLCLR.....	700
Agregacje zdefiniowane przez użytkownika SQLCLR.....	712
Transakcje i współbieżność.....	718
Czym są transakcje.....	719
Blokady.....	722
Eskalacja blokad.....	729
Opóźniona trwałość.....	730
Poziomy izolacji.....	733
Zakleszczenia.....	746
Obsługa błędów.....	752
Konstrukcja TRY-CATCH.....	753
Błędy w transakcjach.....	757
Logika ponawiania.....	760
Podsumowanie.....	761

10 In-Memory OLTP	763
Przegląd technologii In-Memory OLTP	763
Dane zawsze są w pamięci	764
Natywna kompilacja	765
Architektura wolna od blokad i zatrząsków	766
Integracja z SQL Server	767
Tworzenie tabel zoptymalizowanych pamięciowo	768
Tworzenie indeksów w tabelach zoptymalizowanych pamięciowo	770
Indeksy klastrowe i nieklastrowe	770
Nieklastrowe indeksy	771
Indeksy skrótowe	775
Środowiska wykonawcze	786
Zapytania interaktywne	786
Natywnie skompilowane procedury	795
Ograniczenia obszaru powłoki	800
DDL dla tabel	800
DML	802
Podsumowanie	802
11 Grafy i zapytania rekurencyjne	803
Terminologia	803
Graf	804
Drzewa	804
Hierarchie	805
Scenariusze	805
Schemat organizacyjny	806
Zestawienie materiałowe (BOM)	808
System drogowy	812
Iteracja/rekurencja	815
Podgrafy/potomkowie	817
Przodkowie/ścieżka	828
Generowanie poddrzewa poprzez wyliczenie ścieżek	832
Sortowanie	835
Cykle	839
Zmaterializowane ścieżki	843
Przygotowanie danych	843
Odpytywanie	850
Materializowanie ścieżek przy użyciu typu danych HIERARCHYID	855
Utrzymywanie danych	858
Zapytania	866
Dalsze aspekty pracy z HIERARCHYID	870

Zbiory zagnieżdżone.....	882
Przypisywanie wartości lewo- i prawostronnych	883
Zapytania.....	890
Domknięcie przechodnie	893
Skierowany graf acykliczny.....	893
Podsumowanie.....	908
O autorach	911