

Administracja i programowanie pod Microsoft SQL Server 2000

Paweł Rajba

pawel@ii.uni.wroc.pl
<http://www.kursy24.eu/>

Zawartość modułu 9

- Optymalizacja zapytań
- Pobieranie planu wykonania
- Indeksy i wydajność

Optymalizacja zapytań

- Zadania optymalizatora zapytań
 - Ustalenie indeksów i ich przydatności dla wykonywanego zapytania
 - Ustalenie indeksów, które mogą zmniejszyć ilość wierszy do przejścia
 - Ustalenie optymalnej kolejności złączeń tabel
 - Przegląd różnych planów i wybór tego o najniższym koszcie
 - Utworzenie kolumny ze statystykami
- W efekcie powstaje plan wykonania zapytania

Optymalizacja zapytań

- Wyznaczanie kosztu
 - jest na podstawie liczby operacji WE/WY i czas wykorzystania procesora
- Celem jest redukcja
 - liczby zwróconych wierszy
 - liczby przeczytanych stron
 - całego czasu wykonania poprzez minimalizację operacji WE/WY i czasu wykorzystania procesora
 - optymalizator nie zwraca większej uwagi na wykorzystanie pozostałych zasobów

Optymalizacja zapytań

- Uwaga:
 - przy maszynach wieloprocessorowych optymalizator może rozdzielić wykonanie zapytania na kilka procesorów
 - przy takim rozdziale może dojść do wykorzystania większej ilości zasobów, niż w przypadku wykonania zapytania na jednym procesorze

Optymalizacja zapytań

- Etapy wykonania zapytania
 - parsowanie
 - sprawdzana jest składnia
 - wynikiem tego etapu jest drzewo rozbioru zapytania
 - standaryzacja
 - przetworzenie do postaci wygodnej dla optymalizatora
 - wszystkie wykryte zbędne klauzule są usuwane
 - wynikiem jest ustandaryzowane drzewo rozbioru zapytania
 - optymalizacja
 - wybierany jest jeden spośród wielu planów wykonania
 - są trzy fazy: analiza, wybór indeksów, kolejność złączeń

Optymalizacja zapytań

- Etapy wykonania zapytania
 - kompilacja
 - zapytanie jest kompilowane do kodu wykonywalnego
 - algorytm dostępu do danych
 - wybierane są optymalne sposoby dostępu do danych (skanowanie tabeli lub wykorzystanie indeksów)

Optymalizacja zapytań

- Fazy optymalizacji zapytań
 - Analiza zapytania
 - optymalizator ustala kryteria wyszukiwania (WHERE) i złączenia (JOIN)
 - przez minimalizację wyszukiwań, redukuje liczbę koniecznych do przejścia wierszy
 - Wybór indeksów
 - ustalane jest istnienie indeksów dla znalezionych klauzul
 - sprawdzana jest przydatność, czyli m.in. ilość wierszy, która zostanie zwrócona (ustalane na podstawie statystyk)
 - obliczany jest koszt różnych metody dostępu do szukanych wierszy

Optymalizacja zapytań

- Fazy optymalizacji zapytań
 - Wybór kolejności złączeń
 - w przypadku samo-złączeń lub złączeń opartych na większej ilości tabeli obliczana jest kolejność, w której złączenia mają zostać wykonane

Optymalizacja zapytań

- Buforowanie planu zapytań
 - obszar procedure cache służy do pamiętania planów wykonania zapytań
 - dla danego zapytania istnieją co najwyżej dwa plany wykonania
 - jeden dla wykonania seryjnego
 - jeden dla wykonania równoległego (niezależnie od stopnia równoległości)
 - kontekstem wykonania zapytania są dane specyficzne dla danego wykonania (np. parametry)

Optymalizacja zapytań

- Buforowanie planu zapytań
 - przy odpaleniu zapytania, SQL Server najpierw sprawdza, czy nie ma już przygotowanego planu wykonania
 - jeśli jest korzysta z niego, oszczędzając na kompilacji
 - jeśli nie ma, tworzy go i umieszcza w buforze

Optymalizacja zapytań

- Rekompilacja planu wykonania
 - Wykonywana jest wtedy, gdy plan wykonania przestaje być przydatny
 - Sytuacje, w których dojdzie do rekompilacji planu
 - zmiana struktury tabeli lub widoku (ALTER)
 - odświeżenie statystyk (ręcznie lub automatycznie)
 - usunięcie indeksu wykorzystywanego w planie
 - wywołanie procedury sp_recompile
 - duże zmiany w zawartości tabeli
 - w przypadku tabel z wyzwalaczami: duża zawartość tabel inserted i deleted

Optymalizacja zapytań

- Maksymalny czas wykonania zapytania
 - domyślnie każde zapytanie wykonuje się do skutku
 - można ustawić opcję query governor cost limit, która uniemożliwia wykonanie zapytań, które wykonywałyby się za długo
 - procesor zapytań oszacowuje czas wykonania i jeśli jest on większy, niż dopuszczalny, zapytanie nie zostanie wykonane

Optymalizacja zapytań

- Ustawienie maksymalnego czasu wykonania zapytania
 - dla wszystkich połączeń wartość opcji ustawiamy za pomocą procedury `sp_configure`
 - po ustawieniu nowa wartość jest dostępna od razu bez konieczności restartu serwera
 - dla danego połączenia używamy polecenia
 - `SET QUERY_GOVERNOR_COST_LIMIT`
 - ustawienie wartości 0 wyłącza limit

Pobieranie planu wykonania

- Opcje STATISTICS
 - Do wykonania każdego zapytania może zostać dołączona dodatkowa informacja
 - Ustawienie jest w ramach połączenia
 - Do ustawienia używamy polecenia
 - SET STATISTICS rodzaj { ON | OFF }
 - Są trzy rodzaje
 - STATISTICS TIME
 - STATISTICS PROFILE
 - STATISTICS IO

Pobieranie planu wykonania

- Opcje STATISTICS
 - STATISTICS TIME
 - podaje liczbę milisekund sparsowania, kompilacji i wykonania zapytania
 - STATISTICS PROFILE
 - podawany jest profil zapytania, czyli szczegóły dotyczącego jego wykonania

Pobieranie planu wykonania

- Opcje STATISTICS

- STATISTICS IO

- podaje aktywność operacji wejścia/wyjścia

- w wyniku są następujące kolumny

- Table: nazwa tabeli

- scan count: ile razy tabela była przeglądana

- logical reads: liczba stron pobrana z bufora (cache)

- physical reads: liczba stron pobrana z dysku

- $\text{cache hit ratio} = (\text{logical reads} - \text{physical reads}) / \text{logical reads}$

- read-ahead reads: liczba stron umieszczona w buforze (podczas zapytania)

Pobieranie planu wykonania

- Opcje SHOWPLAN
 - Opcje te włączają tryb wyświetlania szczegółowych informacji dotyczących zapytania
 - W wyniku otrzymujemy strukturę:
 - wyniki umieszczone w wierszach
 - wyniki przybierają strukturę drzewa
 - wyniki reprezentują kroki wykonane przez optymalizator
 - prezentowane wartości oszacowaniami na podstawie statystyk, a nie rzeczywiste wartości z wykonania

Pobieranie planu wykonania

- Opcje SHOWPLAN
 - Wyniki zawierają następujące informacje:
 - który indeks wykorzystany z którą tabelą
 - kolejność złączeń tabel
 - inne...
 - Do ustawienia mamy dwie opcje:
 - SET SHOWPLAN_ALL { ON | OFF }
 - SET SHOWPLAN_TEXT { ON | OFF }
 - Wersja ALL zawiera trochę więcej informacji

Pobieranie planu wykonania

- Graficzna prezentacja planu w Query Analyzer
 - składa się z różnych ikon reprezentujących kroki do wykonania zapytania
 - strzałki określają kolejność wykonania operacji

Pobieranie planu wykonania

- Bookmark lookup
 - operator wewnętrzny optymalizatora
 - przy przeglądaniu wierszy, potencjalni kandydaci do wyniku są zapisywani na boku (tworzone są do nich zakładki)
 - odzyskiwanie wierszy odbywa się poprzez
 - RID wiersza
 - lub klucz do indeksu zgrupowanego
 - używany zwykle, gdy w klauzuli WHERE jest operator IN, OR, AND

Indeksy i wydajność

- Indeksy można utworzyć tak, żeby przy zapytaniu nie było potrzeby odwoływać się do stron z danymi
 - oczywistą zaletą jest redukcja odwołań WE/WY
- Zasady dotyczące takich indeksów
 - indeks musi być niezgrupowany
 - indeks musi zawierać wszystkie kolumny wyniku
 - zastosowanie indeksowanych widoków może prowadzić do indeksowania częściowych sum itp, co z kolei może dać duży wzrost wydajności