



Cost-Based Oracle Fundamentals

Jonathan Lewis

Deze eerste boekbespreking wil ik het uitstekende boek over Oracle's Cost Based Optimizer van de hand van Jonathan Lewis behandelen. Het boek verscheen eind afgelopen jaar.

Over de auteur

Jonathan Lewis is een bekend en begenadigd spreker op vele internationale Oracle-conferenties. In 2002 werd hij uitgeno-

digd door Precise, om samen met Anjo Kolk in Nederland een Performance Masterclass te doceren. Ik mocht hem toen interviewen voor zusterblad DB/M. Sindsdien heb ik Jonathan nog vaak mogen ontmoeten, en regelmatig met veel plezier één van zijn sessies of keynotes kunnen bijwonen op een of ander congres. Ik heb Jonathan leren kennen als een zeer gedreven, deskundig en bovenal perfectionistisch onderzoeker, voortdurend bezig om de interne werking van Oracle te doorgronden. Eerder schreef Jonathan het boek 'Practical Oracle 8i - Building Efficient Databases'. Een beetje wrang was wel dat de uitgever stond op het vermelden van een versienummer op de kaft, en Oracle rondom het verschijnen van het boek juist Oracle 9i aankondigde. Zijn eerste boek behandelt weliswaar zaken die specifiek op Oracle 8i betrekking hebben, het overgrote deel is algemeen van toepassing op het inrichten van Oracle databases, en nog steeds actueel.

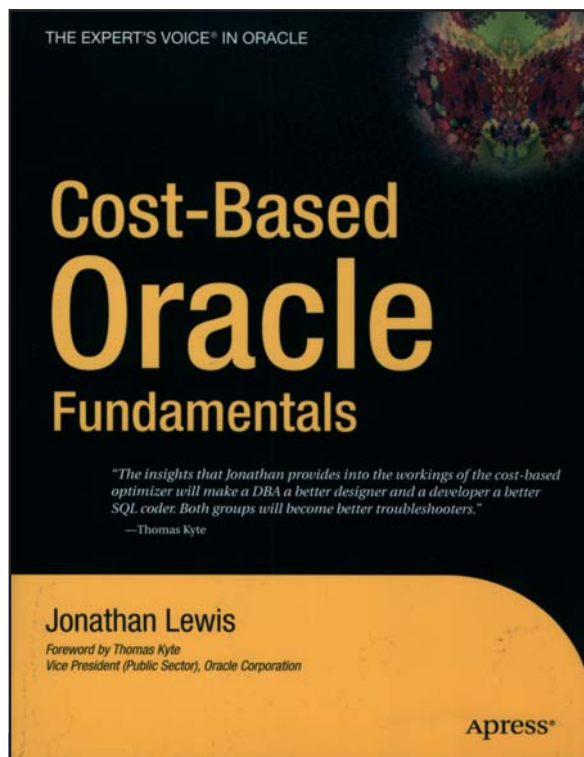
Jonathan heeft een website, www.jlcomp.demon.co.uk, waar hij regelmatig artikelen en white-papers publiceert rondom Oracle RDBMS performance. Verder is hij regelmatig actief met het beantwoorden van lastige performance-vragen op de mailinglist [oracle-l](mailto:oracle-l@freelists.org) (zie www.freelists.org).

Seminar

Er is een gratis exemplaar van het hier beschreven boek voor iedere deelnemer aan het seminar 'Optimising Oracle – Performance by Design', gepresenteerd door Jonathan Lewis op 21, 22 en 23 februari a.s. Dit OakTable seminar wordt georganiseerd door Lex de Haan. Er zijn nog slechts enkele plaatsen beschikbaar. Aanmelden kan nog tot 10 februari a.s. via www.naturaljoin.nl.

Het boek

Toen Jonathan begon met het schrijven van dit boek kon hij niet bevroeden dat het zo lang zou duren om het te schrijven. Dit deel kostte hem 22 maanden. Tijdens zijn presentaties lichtte hij regelmatig een tipje van de sluier op. Dat moest hij vervolgens weer haastig laten zakken, omdat Oracle met weer een update opnieuw een aantal veranderingen had aangebracht.



Boek: *Cost-Based Oracle Fundamentals*,
506 pagina's, begeleidende scripts downloadbaar via de website van de uitgever

Schrijver: Jonathan Lewis

Uitgever: Apress

ISBN: 1-59059-636-6



Zoals hij in de inleiding schrijft: op het moment dat het boek af was, werd de Linux-versie van Oracle 10gR2 vrijgegeven op OTN. Cost-Based Oracle Fundamentals is het eerste van drie delen waarin de Cost Based Optimizer uitgebreid wordt behandeld. In dit boek wordt de lezer de basiskennis bijgebracht die nodig is om de meer geavanceerde aspecten van de Cost Based Optimizer te kunnen begrijpen.

Hoofdstuk 1 begint heel basaal, met de definitie van *Cost*. De eenheid waarin de kosten tot uitdrukking worden gebracht door Oracle is niet altijd duidelijk. Met een aantal duidelijke voorbeelden (die de lezer zelf kan herhalen met de van de website van de uitgever downloadbare code) wordt uitgelegd hoe de kosten van I/O en de CPU rekenkundig in dit begrip worden ondergebracht.

Vervolgens wordt in hoofdstuk 2 de *full table scans* behandeld. Er wordt uitgebreid ingegaan op onder meer de *db_file_multi_block_read_count* parameter en de verschillende manieren waarop de diverse Oracle-versies de kosten van een *full table scan* benaderen.

Eén voor één worden de manieren waarop gegevens uit een tabel kunnen worden geselecteerd en gelezen behandeld. Eerst query's op één tabel, dan het gebruik van een B-tree index, vervolgens hoe de spreiding van gegevens over de blokken in de efficiency van de gegevenstoegang beïnvloedt. De auteur verloochent zijn ervaring als docent niet. Steeds weer wordt de lezer-leerling stap voor stap bij de hand genomen. Zonder belerend of kinderachtig te worden wordt de theoretische uitleg aan de hand van eenvoudig te reproduceren praktijkvoorbeelden geïllustreerd.

Nadat op deze manier in de eerste vijf hoofdstukken de eenvoudige gevallen zijn behandeld gaat hoofdstuk 6 een stukje dieper. In dit hoofdstuk behandelt Jonathan de redenen waarom de berekeningen van de selectiviteit van een query door de optimizer verkeerd worden ingeschat.

Daarna gaat het boek snel verder 'de diepte' in. In hoofdstuk 7 worden de *histograms* behandeld. Stap voor stap wordt het concept uitgelegd en worden de twee type histogrammen behandeld.

Hoofdstuk 8 behandelt de bitmap indexen.

In hoofdstuk 9 wordt beschreven in welke stappen de optimizer een query herschrijft in een misschien betere variant. De optimizer doet dat om beter te kunnen optimaliseren. Dit is een mechanisme dat nog sterk in ontwikkeling is. Bij een Oracle-upgrade kan

er gemakkelijk een andere werkwijze worden gevolgd, met soms betere, soms slechtere performance.

Hoofdstuk 10 t/m 13 gaan vervolgens uitgebreid in op joins. Kardinaliteit van joins, *nested loops*, *hash joins* en *merge joins* passen de revue.

Tenslotte wordt in hoofdstuk 14 het *trace event 10053* behandeld. Door dit *event* te activeren genereert de optimizer een bestand met de informatie waarop de beslissingen van de optimizer gebaseerd zijn (parameters en statistieken) en de verschillende berekeningen en tussenresultaten waarop de keuze voor het uiteindelijke *execution plan* worden gebaseerd. Het hoofdstuk bevat een *tracefile*, en de verschillende onderdelen daarvan worden per onderdeel beschreven.

Het boek wordt afgerond met twee appendices. De eerste somt de onderdelen van de optimizer op die bij een upgrade naar een nieuwe versie voor verrassingen kunnen zorgen. De tweede appendix somt de parameters (inclusief *hidden* parameters) op die het gedrag van de optimizer beïnvloeden.

Doelgroep

Het boek is bestemd voor zowel de DBA als de ontwikkelaar. De DBA die zich wil verdiepen in het optimaliseren van de performance van applicaties, of deze nu in huis gebouwd worden of worden gekocht, heeft aan dit boek een uitstekend leerboek om de werking van de optimizer te leren begrijpen. De ontwikkelaar die niet afhankelijk wil zijn van de DBA om goed performende SQL-statements te schrijven heeft met dit boek een prima middel in handen om de werking van de optimizer te doorgronden.

Conclusie

Cost-Based Oracle Fundamentals is een helder en boeiend geschreven boek, vol met verduidelijkende en door de lezer reproduceerbare voorbeelden. Jonathan Lewis slaagt er uitstekend in de kennis die hij heeft opgedaan in jarenlang onderzoek naar de interne werking van de Oracle-optimizer over te brengen op de lezer. De delen die nog volgen, zullen deze kennis nog verdiepen, naast bijvoorbeeld de aangekondigde behandeling van *partitioning*, *index organized tables*, *parallel processing* enzovoort. Ik raad dit boek sterk aan voor de DBA's en de ontwikkelaars die zich bezig houden met het tunen van query's.

Carel-Jan Engel werkt als onafhankelijk Oracle-consultant.

Hij is lid van het Oak Table Network.

E-mail: cjpengel.dbalert@xs4all.nl