

Goed gereedschap is het halve werk

Gebruiksimpresie van OracleDBA-tools (5)

Rondom het Oracle RDBMS leveren diverse leveranciers hulpmiddelen die het leven, maar vooral het werk van een DBA helpen te veraangenamen. Een eerste blik in de Oracle Toolgids in het decembernummer van Optimize leert dat er tientallen tools worden vermeld die op één of andere manier onder deze definitie vallen. Daarnaast zijn er ongetwijfeld nog tools die een bespreking waard zijn, maar die niet in Rubriek 5 (databasebeheer) van de Oracle Toolgids staan, al is het alleen maar omdat blijkt dat DBA's soms voor ontwikkelaars bedoelde producten graag 'misbruiken' voor het beheer.

Deze keer kijk ik naar DataBee van Net2000 Ltd, een tool voor het samenstellen van uitgebalanceerde testsets. Met name in omgevingen met databases van honderden Gigabytes of meerdere Terabytes is dit van belang: de hardware om een paar kopiëren te maken is niet beschikbaar. Gevolg is dat ontwikkelaars en testers hun werk niet optimaal kunnen doen.

Historie

DBATool is geschreven door Dale Edgar. Ik ontmoette hem tijdens het UKOUG (Britse Oracle User Group, red.) event in Birmingham waar hij me over DataBee vertelde. Toen hij de eerste versie schreef (ik weet niet meer wanneer) had hij te maken met een groeiende productie-database, waar herhaaldelijk een beperkte testset van moesten worden afgeleid. Dat is in de praktijk een lastige klus. Immers, een paar klanten uit een database kopiëren gaat nog wel, maar dan moeten daarbij een wel aantal facturen van die klanten worden meegekopieerd. Die facturen verwijzen weer naar artikelen. Die artikelen weer naar magazijnposities, leveringen en prijzen. Kortom, het is absoluut geen sinecure om op basis van een productiedatabase een verantwoorde en integere subset samen te stellen. Met DataBee geeft Dale Edgar op deze vraag een antwoord.

Op zijn website (www.databee.com) is ook een uitgebreide whitepaper te downloaden, waarin wordt ingegaan op de theoretische aspecten van het opzetten van verantwoorde testsets. Het lijkt

erg eenvoudig, maar iedereen die het wel eens geprobeerd heeft met een wat grotere en/of complexe database heeft ervaren dat het 'even maken van een testdatabase' vies tegenvalt.

Productopbouw

DataBee bestaat uit een reeks tools die kunnen worden gebruikt voor het maken van testdatabases. Er zijn vier aparte programma's: De Set Planner voor het plannen van de testset, de Set Designer voor o.a. het vastleggen van de relaties tussen de tabellen in de testset, de Set Extractor die de rowid's van de testset bepaalt en gecompriemd opslaat, en tenslotte de Set Loader, die verantwoordelijk is voor de gegevenstransfer van de productie database naar de testomgeving.

Naast deze vier programma's staat DBATool. DBATool is gratis, wat ons Hollanders natuurlijk bijzonder aanspreekt. DBATool is een 'standalone' programma, dat buiten de rest van DataBee om kan worden gebruikt. Het heeft geen verbinding met de Oracle database nodig, sterker, dat kan niet eens.

De beschikbare documentatie is uitgebreid. In een 'Quick Start Guide' van 80 pagina's wordt stap voor stap behandeld hoe één en ander in z'n werk gaat. Er is ook een voorbeeldschema beschikbaar. Helaas is de handleiding niet helemaal actueel.

Verantwoording

Deze serie tool-besprekingen is gebaseerd op mijn persoonlijke ervaringen met de betreffende tools. Ieder besproken tool wordt geïnstalleerd in de omgeving waarvoor het is bedoeld en de software wordt met de nodige testcases uitgeprobeerd. Het artikel kan worden gezien als een 'gebruiks-impresie'. De weergave van de resultaten is daarom per definitie subjectief en voor discussie vatbaar. Ik stel me graag open voor discussie. Opmerkingen, vragen en suggesties (ook voor eventueel in volgende afleveringen te bespreken tools) zijn welkom op mijn email-adres. Op uw reacties kom ik graag terug.

Installatie

Installatie van DataBee en DBATool is een fluitje van een cent. Plaats de zelf-uitpakkers in een directory, start ze op, beantwoordt een paar vragen over de directory waar ze geplaatst moeten worden en je kan aan het werk.

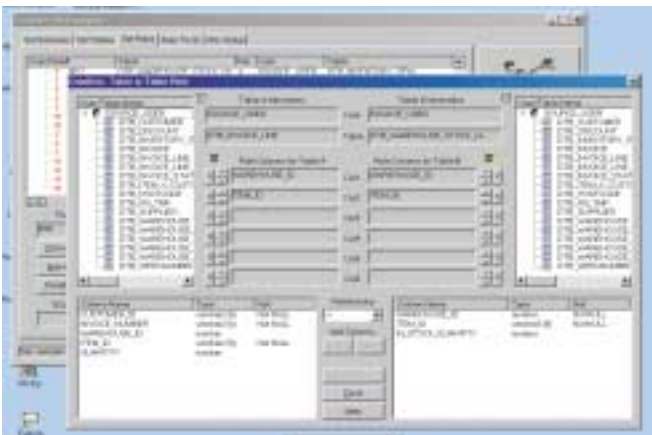
Registratie

DBATool is zoals gezegd een gratis programma. Wel is registratie nodig om het programma te activeren. Dat gaat heel snel via internet, na het invullen van een vragenlijst met NAW gegevens en het accepteren van de licentievooraarden wordt via internet verbinding met de DataBee website gezocht en wordt de pagina met de automatisch gegenereerde sleutel getoond. Deze kan met knippen/plakken in het licentievenster worden ingevuld en vanaf dan is DBATool te gebruiken.

Voor de DataBee programma's werkt de registratie bijna op dezelfde manier, ook hiervoor wordt de vragenlijst ingevuld. De sleutel wordt echter niet automatisch gegenereerd, die wordt na menselijke tussenkomst later per email toegezonden. Dat kan dus één of twee dagen duren.



Afbeelding 1. Met de wizard een plan voorbereiden in de Set Planner



Afbeelding 2. Rules vastleggen in de Set Extractor

DataBee

De vier programma's die samen DataBee vormen hebben een grafische user-interface, die prettig werkt. Het vinden van alle functies is in het begin even wennen, maar dat komt meer doordat de functionaliteit van het tool geen alledaagse kost is dan doordat het niet logisch in elkaar zit. Binnen de programma's kunnen de verbindingen met de schema's in de database worden vastgelegd en met een aparte knop worden getest. De uitgebreide handleiding is apart van de website te downloaden.

Set Planner

DataBee laat de gebruiker op een gestructureerde manier omgaan met het aanmaken van een testset. De eerste stap is dan ook het maken van een plan van aanpak. De Set Planner ondersteunt dat proces door de informatie over het datamodel, inclusief eventuele foreign keys, indexen, omvang van de tabellen e.d. te verzamelen. De Set Planner bevat hiervoor een wizard, waarmee in een paar stappen de benodigde gegevens kunnen worden verzameld. Daarbij is het mogelijk om gegevens van meerdere schema's te combineren.

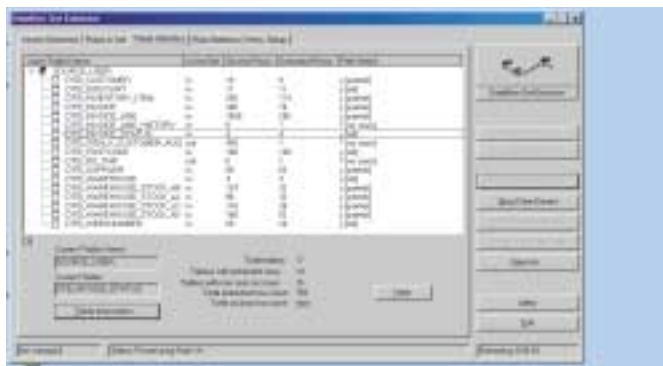
Van de verzamelde (tabel)gegevens kunnen in de Set Planner de onderlinge relaties worden geraadpleegd. Op basis van deze relaties en de doelstellingen van de testset wordt dan per tabel aangegeven of de rijen allemaal, gedeeltelijk of helemaal niet worden overgenomen naar de testset. Bovendien wordt vastgelegd welke tabel als basis wordt gebruikt voor het samenstellen van de subset. Tenslotte kan de set worden opgeslagen voor verwerking in de volgende fase.

Set Designer

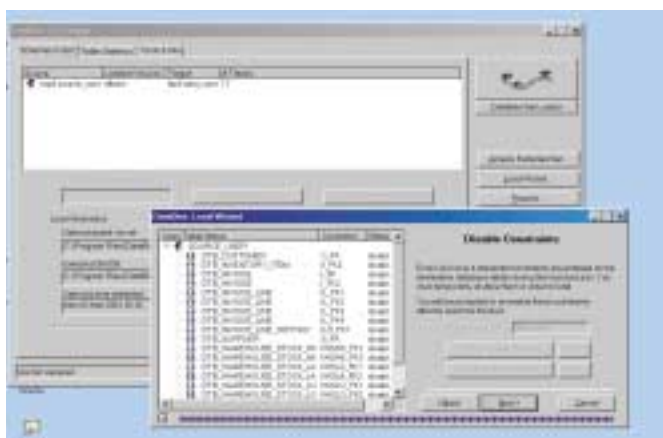
Nadat met de Set Planner de basis is gelegd wordt de Set Designer gebruikt om relaties tussen de tabellen te definiëren. Kan dat dan niet met de constraints die al in de database zitten? Natuurlijk kan dat, maar helaas zijn die er vaak niet. ERP-Systemen als PeopleSoft, JDEdwards en Baan hebben vaak duizenden tabellen, zonder in de database vastgelegde constraints. Verder is het voor de lookup-tabellen noodzakelijk om regels vast te leggen waarmee alle rijen worden meegenomen. Het lijkt wat omslachtig, maar het wordt al snel duidelijk dat het maken van een testset door het volgen van de constraints niet voldoende is, zelfs al zijn ze er. De Set Extractor bevat verder nog een aantal tools die het vinden van regels vereenvoudigen.

Set Extractor

Het samenstellen van de set met behulp van de Set Designer wordt samen met de Set Extractor in een iteratief proces uitgevoerd. De regels worden net zolang verfijnd totdat de Extractor de gewenste set gegevens oplevert. De Extractor loopt de regels niet recht-toe, recht-aan af maar evalueert de gegevensset en regels zo vaak totdat aan alle regels is voldaan.



Afbeelding 3. Rules vastleggen in de Set Extractor



Afbeelding 4. Uitschakelen van constraints voor het laden van de gegevens

Als een bepaalde regel het toevoegen van rijen aan een tabel veroorzaakt, worden de relaties met die tabel opnieuw geëvalueerd om te onderzoeken of andere tabellen niet ook extra rijen nodig hebben, enzovoort. Dubbele rijen worden automatisch geëlimineerd.

Set Loader

Met de Set Loader worden tenslotte de gegevens van de productie-database naar de testdatabase gekopieerd. Het is niet efficiënt om gegevens van de ene naar de andere database kopiëren via de SQL*Net verbindingen tussen de PC en de databases. Ook een export biedt niet de gewenste mogelijkheden. DataBee slaat alleen de rowid's op tijdens het samenstellen van de set. De gegevens zelf gaan niet via de PC naar de bestemming, maar via een databaselink tussen het test-schema en het productie-schema. Dat vereist wel wat voorzichtigheid: het is nu eenmaal meestal niet wenselijk dat ontwikkelaars toegang hebben tot de productie-gegevens. Het kopiëren vindt in stappen plaats. Dit vermijdt problemen met te kleine rollback-segmenten en maakt het bovendien mogelijk om het laden door te starten als het proces onderweg strandt. De Set Loader geeft de DBA de gereedschappen om de foreign key constraints en triggers uit te schakelen gedurende het laden van de gegevens.

DBATool

DBATool valt een beetje buiten de bovengenoemde programma's. Het heeft uitgebreide mogelijkheden om exportfiles te bekijken en op basis daarvan DDL te genereren. De exportfiles mogen dan geen gegevens bevatten, alleen de datastructuur. DBATool is vooral handig om de schema's in de testomgeving aan te maken op basis van de structuur van de productieomgeving. Daarbij kunnen op een handige manier de opslagparameters van de tabellen en indexen worden aangepast.

Ook los van het gebruik in samenhang met DataBee komt de DBA-Tool regelmatig van pas. Hoe vaak moet er niet het één en ander worden aangepast aan de storage parameters van een set tabellen voordat deze kunnen worden geïmporteerd in een testdatabase, of bij een migratie van de gegevens naar andere hardware? In de praktijk wordt dan vaak een (beperkte) DDL-file gecreëerd met het gebruik van de imp-optie INDEXFILE=<bestandsnaam>. Het resulterende bestand ondergaat vervolgens enige bewerkingen met een editor naar keuze, toch blijft dit behelpen. Voor dit soort situaties vormt het (gratis) DBATool een uitkomst.

Samenvatting

DataBee vormt voor organisaties met grote databases een onmisbaar gereedschap als ze hun diverse ontwikkel- en teststadia serieus nemen. Met DataBee kan ook de ontwikkelaar voor een grote database beschikken over een representatieve en samenhangende testset, zonder dat daarvoor dezelfde opslagcapaciteit als het productiesysteem benodigd is. Ondanks de stevige prijs (omgerekend zo'n € 7.500,00) verdient het zich al snel terug door de besparing op server-hardware. Het creëren van een testset gaat met DataBee nog steeds niet met je ogen dicht, maar als de regels eenmaal zijn vastgelegd is het herhaaldelijk verversen van testsets een fluitje van een cent. DBATool is gewoon erg handig voor het aanpassen van opslagparameters van een schema voordat de gegevens worden geïmporteerd.

Producten	DataBee / DBATool
Leverancier	Net200 Ltd. www.databee.com
Prijs	£ 5000,00 per named user
Positief	Zeer doordacht concept Uitgebreide handleiding Grote besparingen mogelijk
Negatief	Handleiding niet helemaal actueel.

Carel-Jan Engel

werkt als onafhankelijk Oracle consultant en is directeur van DBA!ert. Hij is vanaf 1985 in verschillende rollen bezig met ontwikkeling en beheer van systemen op basis van Oracle.
E-mail: cjpengel.dbalert@xs4all.nl