

Bizgres, JasperReports en iReport op weg naar volwaardig BI-platform

# Consolidatieslag in open source BI

Robbert Hoeffnagel

**Ook op het gebied van Business Intelligence laat de open source-gemeenschap van zich horen. Een zoekactie op Internet levert al gauw een aantal interessante projecten op, die vaak maar een beperkte bekendheid hebben. Denk aan Bizgress of iReport. Inmiddels is ook in dit marktsegment een consolidatieslag op gang gekomen, waardoor bedrijven zich meer en meer met deze projecten gaan bemoeien; bijvoorbeeld als sponsor. De aantrekkelijkheid van de ontwikkelde programmatuur neemt hierdoor vaak alleen nog maar toe.**

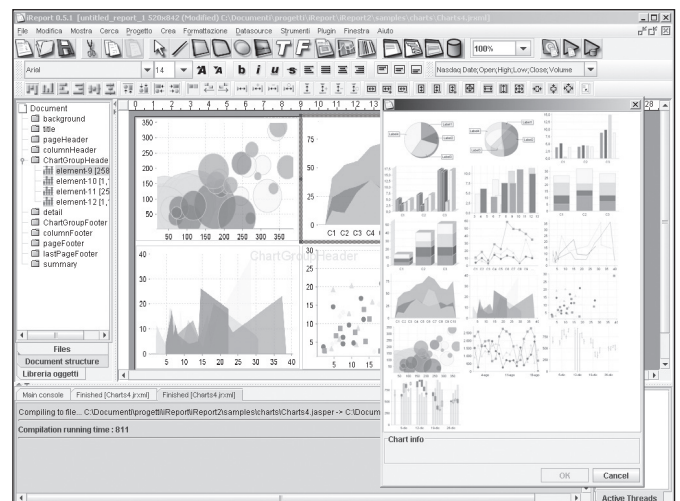
In de open source-gemeenschap wordt er nog wel eens over geklaagd: de toenemende invloed van het bedrijfsleven op projecten. Het bekendste voorbeeld is natuurlijk Linux, waar grote concerns als IBM en Novell niet alleen in belangrijke mate een bijdrage aan leveren, maar waar zij ook de nodige commerciële belangen te verdedigen hebben. Sterker nog, een bedrijf als Novell heeft zijn toekomst zelfs min of meer aan Linux opgehangen.

## Sponsor van project

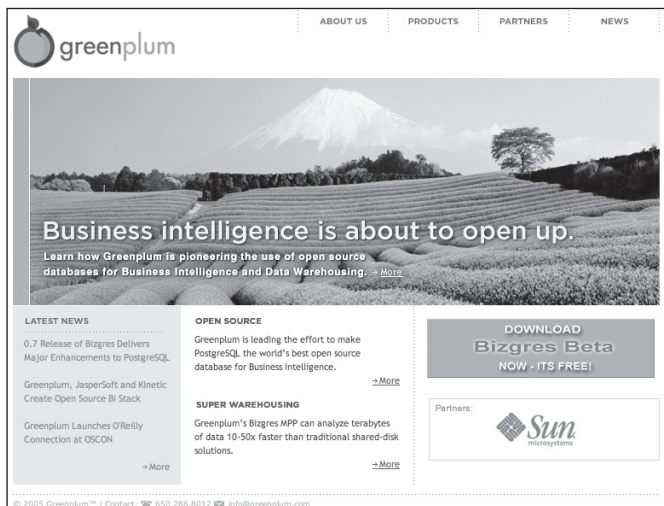
Op een veel bescheidener schaal zien we een vergelijkbare ontwikkeling bij de open source-projecten die zich richten op Business Intelligence. Een van de bekendste projecten is ongetwijfeld Bizgres dat eerder dit jaar van start ging. Dit is een project om op basis van de bekende open source database PostgreSQL tot een volwaardig BI-platform te komen en dat zich bovendien kan meten met alles wat het bedrijfsleven aan commerciële producten te bieden heeft. De leiding over het project is in handen van een vrij brede groep van personen. Sommigen van hen behoren tot de groep die oorspronkelijk op het idee van Bizgres is gekomen. Anderen zijn in dienst van Greenplum, een bedrijf dat zich bezig houdt met BI op basis van open source-producten. Daarnaast spelen mensen uit de PostgreSQL-community een rol, evenals medewerkers van andere bedrijven die een technische bijdrage leveren. De rol die Greenplum speelt, is typerend voor de wijze waarop bedrijven steeds vaker betrokken raken bij open source-projecten. Het bedrijf treedt namelijk op als sponsor. Hier heeft het bedrijf een aantal redenen voor. Doordat de onderneming de eigen producten en diensten baseert op – met name – Bizgres is het er veel aan gelegen dat de snelheid waarmee nieuwe functionaliteit wordt ontwikkeld zo hoog mogelijk is. Door een aantal diensten aan de community aan te bieden en bijvoorbeeld eigen ontwikkelaars beschikbaar te stellen, hoopt men dit doel te

bereiken. Daarnaast is het voor het bedrijf van groot belang dat PostgreSQL uitgroeit tot hét standaard platform voor open source BI. Dit is natuurlijk een puur commerciële reden en is ingegeven door het feit dat de expertise van het bedrijf hierdoor alleen maar meer waard wordt.

Bizgres is – ook door de relatief korte tijd dat het project pas bestaat – nog niet erg bekend als BI-platform. Het project heeft tot doel om lijn te brengen in de ontwikkelingen rond PostgreSQL. Het gaat hier om een project dat is afgeleid van het bekende Postgres-project van Michael Stonebraker van de Universiteit van Californië in Berkeley. Dit is wat de community zelf een ‘object-relacioneel’ database management-systeem noemt, waaraan inmiddels meer dan tien jaar wordt gewerkt. Postgres zelf is ontwikkeld met geld van de Amerikaanse overheid. Sinds 1986 zijn er tal van implementaties gerealiseerd, die in 1989 hebben



Afbeelding 1: Werken met iReports als front-end voor Jasper Reports.



**Afbeelding 2:** Greenplum houdt zich bezig met BI en datawarehousing op basis van Bizgres.

geleid tot een eerste 'product' dat aan een selecte groep van gebruikers beschikbaar is gesteld. Illustration Information Technologies heeft als eerste een commercieel product op deze code gebaseerd. Deze firma is overigens later onderdeel geworden van Informix, dat zelf inmiddels is overgenomen door IBM.

## Postgres95 en PostgreSQL

In 1994 is aan Postgres een SQL-interpreter toegevoegd, waarna het project onder een nieuwe naam bekend werd: Postgres95. Postgres95 is volledig in ANSI C geschreven en kent een code base die een kwart kleiner is dan zijn voorgangers. Bovendien is de interne structuur flink onder handen genomen. Het gevolg daarvan was dat het prestatieniveau gemeten op basis van de Wisconsin Benchmark dertig tot vijftig procent kon worden verbeterd. Een andere verandering was het feit dat de query-taal PostQUEL werd vervangen door SQL dat in de server is geïmplementeerd. Subquery's werden niet ondersteund, maar konden wel door middel van door de gebruiker gedefinieerde SQL-functies worden geïmiteerd. Verder werd psql toegevoegd voor interactieve SQL-query's waarbij gebruik wordt gemaakt van Readline. Al deze veranderingen hebben uiteindelijk geleid tot de komst van PostgreSQL.

PostgreSQL stelt geen beperkingen aan de omvang van de database. Een tabel mag maximaal 32 TB groot zijn, met rijen van maximaal 1,6 TB en velden van ten hoogste 1 GB. Aan het aantal rijen per tabel wordt wederom geen beperking gesteld, terwijl afhankelijk van het type maximaal 250 tot 1600 kolommen per tabel mogelijk zijn. Er kan verder gewerkt worden met een onbeperkt aantal indexen per tabel.

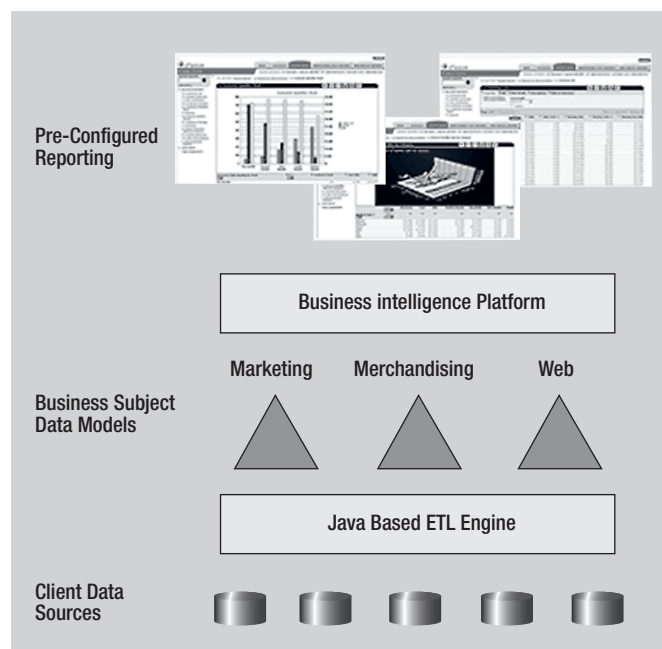
In 2003 ging een bedrijf genaamd Greenplum van start. Het doel van deze firma was het ontwikkelen van BI en datawarehouse-oplossingen op basis van open source-producten. De basis voor deze producten wordt geleverd door PostgreSQL. In april van dit

jaar heeft dit geleid tot een formele sponsor-relatie met de PostgreSQL-community om tot een specifiek voor BI bedoelde variant op PostgreSQL te komen: Bizgres. Bizgres kan het beste gezien worden als een 'distributie' van PostgreSQL, net als SuSe of Red Hat distro's van Linux bieden. Het biedt naast PostgreSQL een reeks van extra hulpmiddelen die nodig zijn om van deze open source-database een goede basis voor BI te maken. Denk wat add-ons betreft aan JDBC, de JavaLoader, autovacuum en bijvoorbeeld een programma als de Configurator. Daarnaast is een eenvoudige installer ontwikkeld die het mogelijk maakt om alle tools te installeren.

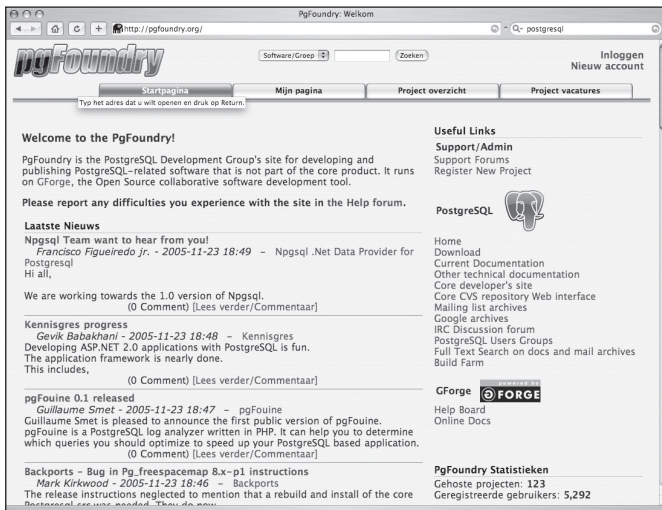
## Jasper Reports en iReport

Afgelopen zomer maakte Greenplum bekend dat het nauw gaat samenwerken met een bedrijf dat luistert naar de naam JasperSoft. Samen met een derde partij – Kinetic Networks – heeft men een zogeheten 'development stack' voor BI gelanceerd. Het gaat hier om een reeks van op Bizgres gebaseerde tools waarmee ontwikkelaars zelf BI-applicaties kunnen ontwikkelen. In dit geheel levert JasperSoft de rapportage-voorzieningen, terwijl Kinetic de benodigde en op Java gebaseerde ETL-voorzieningen aandraagt. Om aan te geven hoe met deze development stack in de praktijk kan worden gewerkt, hebben de drie bedrijven een voorbeeldapplicatie ontwikkeld: Bizgres Clickstream waarmee een zeer gedetailleerd inzicht kan worden verkregen in de activiteiten van bezoekers van een website.

JasperSoft is een bedrijf dat – net als Greenplum – zijn business-model heeft gebaseerd op open source-producten. Het levert zelf Jasper Reports, dat een open source-alternatief voor bijvoorbeeld Crystal Reports of Cognos wil zijn. Van de begin oktober



**Afbeelding 3:** De ETL-architectuur van Kinetic Networks.



Afbeelding 4: PostgreSQL; community speelt grote rol.

gelanceerde versie 1.0 van dit programma zijn volgens het bedrijf inmiddels 750.000 kopieën gedownload. Hoewel het product zich volgens analisten van bijvoorbeeld onderzoeksbureau Quandt Analytics wat betreft functionaliteit nog niet kan meten met bijvoorbeeld Crystal Reports zijn de aanwezige voorzieningen voor veel toepassingen wel degelijk voldoende. Bovendien is voor een goede integratie met Bizgres gezorgd, wat uiteraard een belangrijke voorwaarde is voor acceptatie van dit soort BI-tools. JasperSoft volgt een strategie die vergelijkbaar is met die van Greenplum. Sinds kort treedt namelijk ook dit bedrijf op als sponsor van een open source-project. Het gaat om iReport. Het gaat hier om een hulpmiddel dat bovenop Jasper Reports draait en gebruikt kan worden voor het ontwerpen van rapporten. In feite dient het als een gebruikersinterface om op grafische wijze rapporten te bouwen. Het bedrijf geeft aan dat een groot aantal gebruikers – 65 procent is de schatting – dit hulpmiddel sowieso al in gebruik heeft. Dat JasperSoft sponsor van dit project is geworden, betekent onder andere dat nu toch tot een nadrukkelijke integratie van de twee programma's kan worden gekomen. Tot nu toe liep iReport altijd een paar maanden achter op Jasper Reports, omdat men eerst moest wachten op het beschikbaar komen van de nieuwe code voordat men iReport hierop kon aanpassen. In het vervolg zullen nieuwe versies van tools gelijktijdig beschikbaar komen.

## 'Open source gorilla'

De samenwerking tussen bedrijven en een open source-community's wordt steeds nauwer. Analisten verwachten dat op termijn hieruit een of enkele zogeheten 'open source gorilla's' zullen ontstaan. Zoals beide BI-voorbeelden die hier genoemd zijn laten zien, biedt samenwerking vaak belangrijke voordelen. Er komt bijvoorbeeld wat meer lijn in de richting waarin projecten zich ontwikkelen, waardoor de inspanningen van alle betrokkenen sneller en tot een beter zichtbaar resultaat leiden. Toch zal er nog lange tijd een zekere spanning bestaan tussen de

traditionele open source-ontwikkelaars en de bedrijven die bij hun projecten aanhaken. Dat is ook goed zichtbaar op veel websites van open source-projecten. Neem als voorbeeld de 'frequently asked questions' (faq) op de site van Bizgres. Hier staat letterlijk als een van de vragen: 'Is it true that some companies are part of Bizgres?' Het bevestigende antwoord is omzichtig geformuleerd. Zo stelt men onder andere: 'There are a number of companies that are interested in making PostgreSQL the best open source database for BI.' Tegelijkertijd wordt aangegeven dat het iedereen is toegestaan om op basis van PostgreSQL zelf een commercieel product te ontwikkelen, zolang men zich maar houdt aan de regels die zijn vastgelegd in de BSD-licentie.

## PostgreSQL stelt geen beperkingen aan de omvang van de database

Voor bedrijven die overwegen om met open source business intelligence aan de slag te gaan, is de hele discussie over de invloed van bedrijven op open source-projecten echter maar in beperkte mate interessant. Veel relevanter is dat de betrokkenheid van bedrijven de continuïteit van dit soort projecten zeker kan stellen. Geen onderneming wil uiteraard voor een cruciaal hulpmiddel als BI afhankelijk zijn van een klein en veelal onzichtbaar groepje ontwikkelaars, waarop men geen enkele grip heeft. Hoewel het karakter van de open source-wereld er ongetwijfeld enigszins door is veranderd, komt de betrokkenheid van het bedrijfsleven het uiteindelijke succes van veel open source-projecten alleen maar ten goede.

**Robbert Hoeffnagel** is freelance journalist.

## Meer weten?

Bizgres: [www.bizgres.org](http://www.bizgres.org)

PostgreSQL: [www.postgresql.org](http://www.postgresql.org)

iReport: <http://ireport.sourceforge.net>

Greenplum: [www.greenplum.com](http://www.greenplum.com)

JasperSoft: [www.jaspersoft.com](http://www.jaspersoft.com)

Kinetic Networks: [www.kineticnetworks.com](http://www.kineticnetworks.com)