

‘Business Intelligence for the masses’

# Hoofdrol voor PowerAnalyzer in toekomstplannen Informatica

Robbert Hoeffnagel

**M**et de aankondiging van het Business Activity Platform proberen Informatica en WebMethods invulling te geven aan de langverwachte trend, die zegt dat ‘Enterprise Application Integration’ (EAI) en ‘Extract Transform and Load’ (ETL) operaties met elkaar zullen convergeren tot één methodiek voor het integreren van data en applicaties. Een van de twee pilaren waarop dit platform rust is PowerAnalyzer van Informatica. Deze software speelt een centrale rol in de toekomstplannen van het softwarehuis. Het moet tevens de weg openen voor een brede verspreiding van BI-tools onder niet-traditionele eindgebruikers.

Al jaren zien analisten die de ETL-markt volgen, twee belangrijke trends ontstaan. De eerste is de verbreding van het productenprogramma van menig ETL-aanbieder in de richting van analytische applicaties. In plaats van het aanbieden van ‘alleen maar’ een verbinding tussen data uit tal van bronnen en de toepassingen waarmee deze gegevens bewerkt kunnen worden, zag men het als een logische ontwikkeling dat bedrijven als Acta, Cognos en Informatica zouden proberen om een plek hoger in de ‘food chain’ te zoeken.

De tweede ontwikkeling heeft te maken met het naar elkaar toegroeien van EAI en ETL. In dit geval dienen de ETL-tools te kunnen aansluiten op *message brokers* in bijvoorbeeld MQ Series van IBM of Tibco’s RendezVous. In beide gevallen draait het om integratie, in het eerste geval om het integreren van data die afkomstig is uit tal van vaak zeer verschillende bronsystemen, in het tweede geval gaat het om het aan elkaar knopen van toepassingen. In plaats van het batch-gewijs ophalen, bewerken en laden van gegevens, draait het bij EAI juist om het record voor record of in kleine ‘bursts’ beschikbaar maken van gegevens die afkomstig zijn uit verschillende bronnen en in ‘near real-time’.

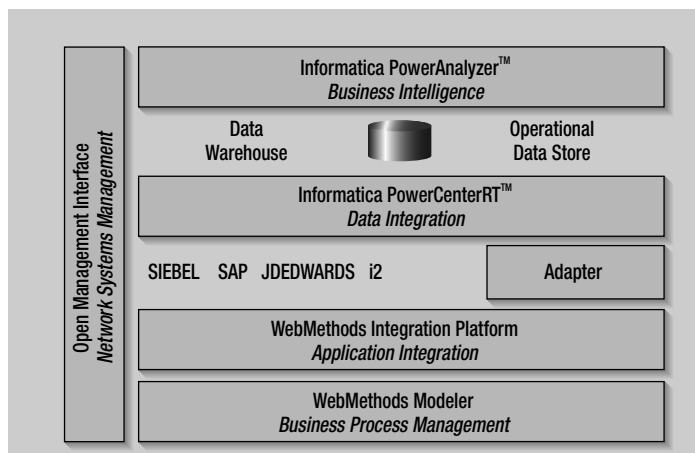
## EAI + ETL

Met name over de tweede trend is de afgelopen jaren veel geschreven, zonder dat het tot een daadwerkelijke convergentie is gekomen. Daar lijkt met een recente aankondiging, zoals die van Informatica en WebMethods van het zogenaamde Business Activity Platform (BAP),

verandering in te komen. De twee bedrijven combineren het EAI-platform van WebMethods met Informatica’s PowerCenter RT-platform en de analytische applicatie PowerAnalyzer 4. Dit laatste product speelt bovendien een centrale rol bij de pogingen van Informatica om analytische programma’s voor een breed publiek aantrekkelijk te maken.

Centraal in deze aanpak staat een term van Gartner: Business Activity Monitoring of BAM. Het idee is dat de enige manier om werkelijk inzicht te krijgen in de stand van zaken binnen de organisatie, het aan elkaar koppelen is van alle relevante applicaties en de gegevens die hiermee worden gegenereerd of verwerkt. Dat dient bovendien zowel in real-time te kunnen gebeuren of met een snelheid die daar dicht tegenaan ligt, als ook in bulk.

Het draait hierbij dus zowel om de applicaties als om de data. Integratie van toepassingen kan plaatsvinden via de messaging-faciliteiten die door EAI-servers van leveranciers als IBM, Tibco, Vitria of WebMethods worden aangeboden. Hierbij wordt op basis van berichten die via een ‘broker’ worden doorgegeven, gebruik gemaakt van de functionaliteit van andere applicaties. Om ook gegevens te kunnen integreren is echter meer nodig: ETL en dan bij voorkeur ETL-functionaliteit die zowel bulk- als real-time gegevenstransport mogelijk maakt.



AFBEELDING 1: DE OPBOUW VAN HET BUSINESS ACTIVITY PLATFORM (BAP) VAN INFORMATICA EN WEBMETHODS.

## KOPPELEN

De aankondiging van Informatica en WebMethods van het Business Activity Platform (BAP) dient precies dat doel: het koppelen van messaging, ETL-faciliteiten en het gebruik van metadata om de bruikbaarheid van de beschikbare gegevens te optimaliseren gekoppeld aan analytische functionaliteit. Aan de ene kant biedt BAP de faciliteiten voor het lezen en transformeren van EAI-data en het in real-time integreren van EAI-berichten in een datawarehouse of een operationele data store. Aan de andere kant biedt de combinatie analytische faciliteiten waarmee de aldus beschikbaar gekomen gegevens kunnen worden bekeken en geanalyseerd.

BAP bestaat uit een reeks van bestaande componenten (zie afbeelding 1). Allereerst gaat het om het Integration Platform en de Modeler van WebMethods voor respectievelijk het integreren van applicaties en het beheren van businessprocessen. Van Informatica afkomstig is PowerCenterRT voor het in real-time integreren van data en het voor Business Intelligence bedoelde PowerAnalyzer. Daarnaast is een adapter (officieel de 'Business Activity Platform Adapter' geheten) beschikbaar die het mogelijk maakt om op het niveau van metadata tot integratie tussen deze componenten te komen.

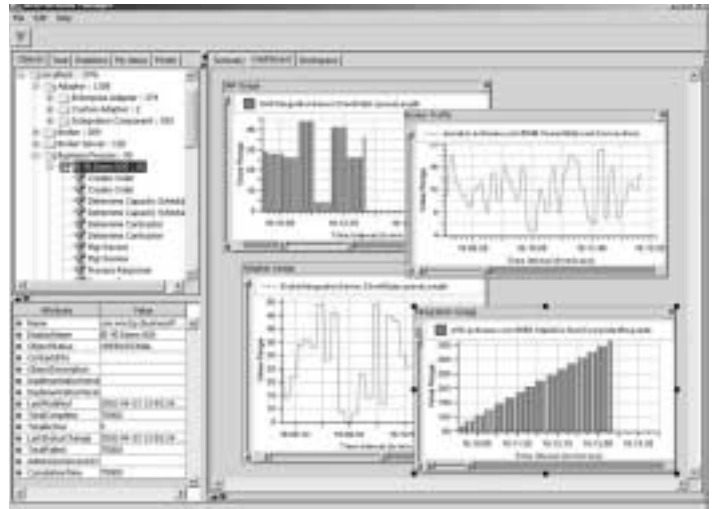
PowerCenterRT maakt het opschonen, transformeren en integreren mogelijk van massale hoeveelheden gegevens, die afkomstig kunnen zijn uit bronnen als ERP-systemen, databases, operationele en legacy-systemen en dergelijke. Het biedt een centrale repository waarin metadata is vastgelegd waardoor het gebruik en hergebruik van gegevens wordt ondersteund. Het integratieplatform van WebMethods behoort tot de bekende integratieservers op basis waarvan legacy applicaties, standaard pakketten, databases en work-flows kunnen worden geïntegreerd.

Beide platformen kennen grafisch georiënteerde ontwikkelomgevingen. PowerCenterRT biedt een zogeheten 'code-less' omgeving voor het opzetten van integratiestructuren, terwijl de Modeler van WebMethods gebruikt kan worden voor het vormgeven van integratieprocessen (zie afbeelding 2). Ook hierbij geldt dat er in principe geen programmering nodig is.

## INTEGRATIE

Het voordeel van deze integratie zo stellen de twee bedrijven in een 'solution brief', getiteld 'The Business Activity Platform; First Best-of-Breed Enterprise Integration Platform', is dat BAP zowel het integreren van applicaties mogelijk maakt, waarbij eerst in bulk data beschikbaar moet worden gemaakt, als die vormen van integratie waarbij gegevens record voor record of in kleine hoeveelheden opgehaald en beschikbaar gesteld dienen te worden.

Hoe werkt dit in de praktijk? In de genoemde solution brief worden twee voorbeelden aangehaald. Zo kan een financiële instelling één gemeenschappelijk business-proces samenstellen voor het op geautomatiseerde wijze behandelen van aanvragen voor leningen. In dit geval kan PowerCenterRT uit de diverse bronsystemen alle gegevens ophalen die voor de behandeling noodzakelijk zijn en deze gegevens



**AFBEELDING 2: DE WEBMETHODS MANAGER VOOR HET BEHEREN EN VOLGEN VAN BUSINESS-PROCESSEN.**

beschikbaar stellen aan het integratieplatform van WebMethods.

In een tweede voorbeeld gaat het om een validatie van een verzekeringsmaatschappij die transacties wil controleren op basis van per klant vastgelegde gegevens. In dit geval wordt PowerCenterRT geabonneerd op alle messages die door het integratieplatform van WebMethods lopen en die betrekking hebben op het te controleren type transactie. Vervolgens valideert PowerCenterRT deze messages op basis van gegevens die vastliggen in klantenbestanden of die in een operationele datastore zijn verzameld.

Naast deze vorm van integratie van gegevens en applicaties biedt BAP zoals gezegd ook analytische functionaliteit. Hiervoor wordt PowerAnalyzer van Informatica gebruikt. Dit product is een goed voorbeeld van de trend die analisten al geruime tijd zien en die aangeeft dat ETL-aanbieders ook analytische faciliteiten willen aanbieden.

## NIEUWE RELEASE

Van PowerAnalyzer is onlangs de nieuwe release versie 4 verschenen, onder andere bedoeld om Business Intelligence ook binnen het handbereik van de niet-traditionele BI-gebruiker te brengen. Het bedrijf claimt ten opzichte van eerdere versies op vijf gebieden verbeteringen te hebben aangebracht: de gebruiksmogelijkheden zijn uitgebreid, de integratie met Excel is verbeterd, er zijn faciliteiten toegevoegd voor zogeheten 'guided analyses', men heeft het product voorzien van hulpmiddelen voor performance management en er is 'real-time awareness' toegevoegd.

De verbeteringen die Informatica ten aanzien van de gebruiksmogelijkheden ziet, zijn het gevolg van een andere manier van ontwikkelen van zakelijke software. Men heeft een methodiek gebruikt – men spreekt van 'persona driven' – die veelvuldig bij software voor gebruik door consumenten wordt toegepast. Door op basis van interviews met personen uit de beoogde gebruikersgroepen zogeheten 'user personality profiles' op te stellen, claimt men in staat te zijn geweest om de benodigde analytische functies zodanig in te richten dat deze goed op de diverse doelgroepen aansluiten. Daarbij stond men name het idee



**AFBEELDING 3: EEN DASHBOARD IN POWERANALYZER 4 VAN INFORMATICA.**

centraal dat de onderliggende complexiteit van bewerkingen en dergelijke voor de eindgebruiker verborgen dient te blijven (afbeelding 3).

Door de integratie met Excel kan een eindgebruiker binnen de gebruikersinterface van PowerAnalyzer over de volledige functionaliteit van Excel kan beschikken. Data kan 'live' vanuit PowerAnalyzer in Excel-templates of pivot-tabellen worden geplaatst. Deze vorm van integratie betekent dat Excel nu ook zonder problemen met behulp van een browser kan worden gebruikt.

### ANALYTISCHE WORKFLOWS

Achter het begrip 'guided analyses' gaat een techniek schuil die ook wel 'Analytic Workflow' wordt genoemd. Doordat de gebruiker van PowerAnalyzer vooral een dashboard gebruikt voor het presenteren van gegevens en analyseresultaten, is er een methode nodig om de relatie tussen een via het dashboard gesignaleerd probleem en de onderliggende oorzaken te kunnen bekijken. Door middel van een analytic workflow kan de gehele *decision tree* zichtbaar worden gemaakt. Hierin is precies te zien op welke manier bewerkingen op de gegevens zijn toegepast. Gebruikers kunnen deze decision trees zelf creëren en vastleggen waardoor ook de onderlinge samenwerking binnen de organisatie verbeterd kan worden. Andere eindgebruikers zijn immers altijd in staat de onderliggende bewerkingen in te zien die tot een in een dashboard gepresenteerd resultaat hebben geleid.

PowerAnalyzer beschikt verder over faciliteiten voor scorecarding, dashboards, six sigma metrics, metrics management, ad-hoc query's, analysis en bursting. Bij een aantal concurrenten is het voor een dergelijke brede reeks van analysefaciliteiten noodzakelijk dat meerdere

producten worden gebruikt. Dat kan een extra inspanning voor de ICT-afdeling opleveren, aangezien dan meerdere softwarepakketten geïnstalleerd en onderhouden dienen te worden.

De rol die PowerAnalyzer speelt in het Business Activity Platform wordt duidelijk door de mogelijkheden die het product biedt voor real-time bewerkingen. Informatica spreekt overigens liever van 'real-time' en contextgevoelige 'right-time'. De software kan hiertoe gekoppeld worden aan Power-CenterRT. Het ondersteunt onder andere Java Messaging Services (JMS) en kent een faciliteit die 'federated query's' wordt genoemd. Hiermee kunnen eindgebruikers via een dashboard real-time bewegende indicatoren bekijken.

### METADATA

Informatica en WebMethods doen met de introductie van BAP een poging invulling te geven aan de trend naar convergentie van EAI en ETL. In het rapport 'Market Overview Update: ETL noemt onderzoeksbureau Giga Information Group deze integratie weliswaar 'compelling', omdat een ICT-afdeling niet langer noodgedwongen meerdere technologieën voor datatransport hoeft aan te schaffen en te ondersteunen, maar wijst men er ook op dat dit idee al geruime tijd bestaat maar tot nu toe niet echt van de grond is gekomen. Nu de interfaces tussen ETL-tools en messaging-systemen beter worden, verwacht Giga dat er vooruitgang geboekt kan worden, maar ziet men toch nog drie redenen waarom de convergentie van EAI en ETL relatief traag zal verlopen.

De eerste reden is gelegen in het feit dat er ten aanzien van met name ERP- en legacy systemen zoveel verschillende datastores ondersteund dienen te worden, zodat van één technische oplossing vooralsnog geen sprake kan zijn. Een tweede probleem acht men de metadata. Message brokers kunnen hier onvoldoende mee uit de voeten en dat blijft volgens Giga voorlopig wel even. Hierdoor kunnen ontwikkelaars en databasebeheerders niet beschikken over de gegevens die zij nodig hebben voor een optimale vorm van hergebruik van data.

Een laatste obstakel voor een snelle acceptatie van de EAI/ETL-convergentie is volgens Giga tenslotte het fundamentele verschil tussen query-intensieve en transactie-intensieve systemen. Deze verschillen hebben altijd uitgenodigd – en doen dat nog steeds – tot verschillende benaderingen van enerzijds het in bulk uitvoeren van ETL-operaties en anderzijds de snelle berichtenuitwisseling die bij EAI wordt toegepast.

Toch heeft het samengaan van EAI- en ETL-technieken de toekomst, menen de analisten van Giga. Niemand wil twee of meer tools kopen die hetzelfde doel dienen. Het ook wel 'zero latency data update' genoemde fenomeen is van groot belang voor iedere organisatie waar men continu een beeld wil hebben van de stand van zaken van de operatie. Daarvoor is het echter wel van groot belang dat er een metadata-oplossing komt die over de gehele organisatie heen kan worden gelegd en het gebruik en hergebruik van data, ongeacht de toegepaste integratietechniek, ondersteunt. Zover is het echter nog niet. ●

Robbert Hoeffnagel is freelance journalist.