

## BI Architectuur visie: het integratie platform

Steeds meer worden applicaties ontmanteld in services, doordat ondernemingen de voordelen inzien van service georiënteerde architecturen. Services die maar een deel van de data nodig hebben om goed te kunnen functioneren. De databases, onderliggend aan de applicaties, zullen dus modelmatig kleiner worden. Maar wat er gebeurt er dan met de data? Waar wordt deze in de toekomst bewaard?

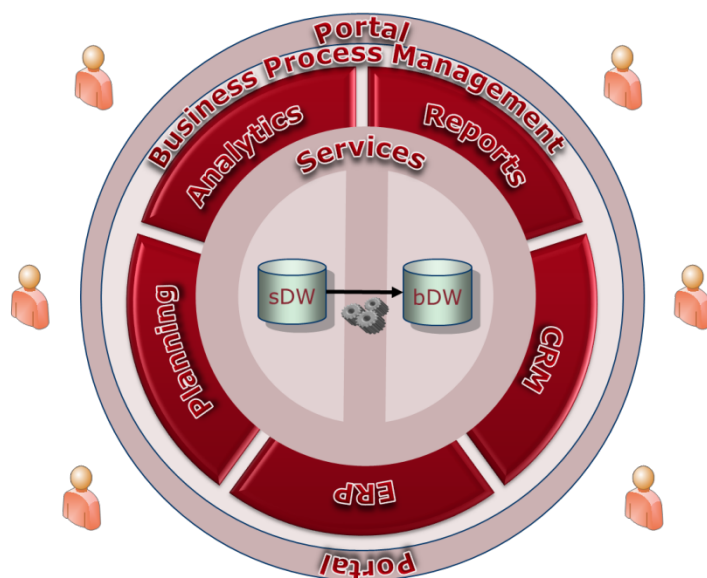
Daarnaast zien we dat ondernemingen steeds meer functioneren in een omgeving waarin het noodzakelijk is, bijna real-time, met andere partners te communiceren over data. Partners, leveranciers, klanten, de overheid et cetera. Dit noopt tot het opzetten van projecten, leidende tot de realisatie van bus architecturen, om data uit te kunnen wisselen op basis van abonnementservices (publish en subscribe).

Als laatste zijn organisaties natuurlijk volop bezig hun besluitvormingsprocessen te ondersteunen met adequate BI architecturen.

Een goed voorbeeld is hoe gemeenten in Nederland dit organiseren. Wij voerden een onderzoek uit in het zuiden van Nederland, en ontdekten daarbij dat grotere gemeenten bezig zijn met het opzetten van een Enterprise applicatie integratie architectuur. Specifiek bedoeld voor deze informatie uitwisseling met hun partners (politie, jeugdzorg, brandweer, belastingdiensten, rijksoverheid, geestelijke zorginstellingen et cetera). Het data model dat dit moet ondersteunen wordt het RSGB genoemd (Referentiemodel Stelsel van Gemeentelijke Basisgegevens). Dit data model is opgesteld door de rijksoverheid. Daarnaast zijn al deze gemeenten bezig met het realiseren van een data warehouse omgeving, om dezelfde data nog eens te verzamelen. Ditmaal voor rapportage vereisten. In veel gevallen zijn zelfs twee projectteams onafhankelijk van elkaar bezig.

Naar onze mening zijn deze projecten gelijkwaardig, en zouden ze niet afzonderlijk beschouwd moeten worden. Wij suggereren het opzetten van een integratieplatform of informatie hub. Hierbij vermijden wij expres de term: data warehouse, omdat het besmet is met de gedachte dat het alleen geschikt zou zijn voor rapportage doeleinden.

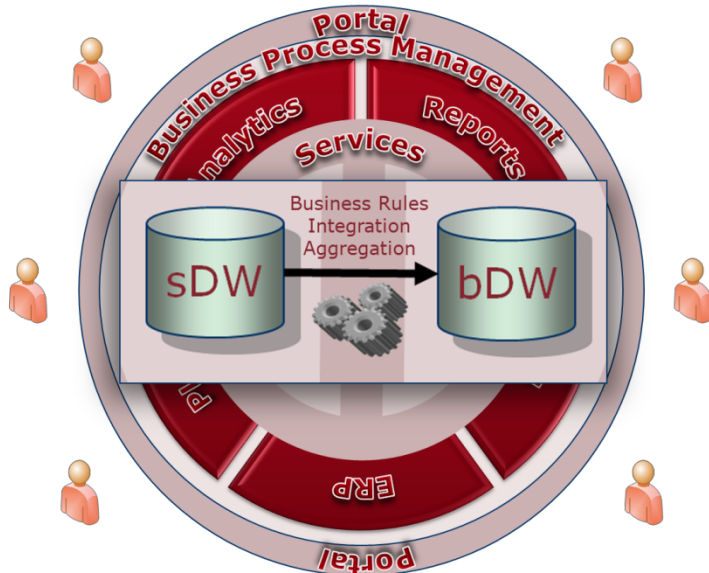
Dit integratie platform onderscheidt verschillende lagen waarvan de belangrijkste, de bron integratielaag, en de business integratielaag zijn.



Figuur 1. Integratie platform dat vele doeleinden kan bedienen

## Bron integratielaag

Deze laag is in de figuur te herkennen als sDW of **source** datawarehouse laag. Deze laag verzamelt de data van de bronsystemen en voegt attributen toe die het opbouwen van historie mogelijk maken. Op dit punt kan alleen sprake zijn van semantische integratie zodra business keys in verschillende bronnen hetzelfde zijn. Denk bijvoorbeeld aan klantcodes die in twee verschillende bronnen (gelukkig!) hetzelfde zijn. Belangrijk is dat de data nog niet is gewijzigd in deze verzamelstap, zodat dit de traceerbaarheid van data ten goede komt. Een belangrijke eis in het licht van nieuwe wetgeving (Basel II, Solvency II, USGAAP, IFRS etc).



Figuur 2. Bron en business integratielagen

De bronlaag bevat:

1. Een minimum data model, dat nodig is voor informatie uitwisseling met alle partners. We noemen dit het communicatie model. In ons voorbeeld, een data model dat alle attributen bevat nodig voor het verschaffen van de RSGB data;
2. Een meer uitgebreid data model, dat de andere attributen bevat om operationele, tactische en strategische (management) rapportage te kunnen ondersteunen;

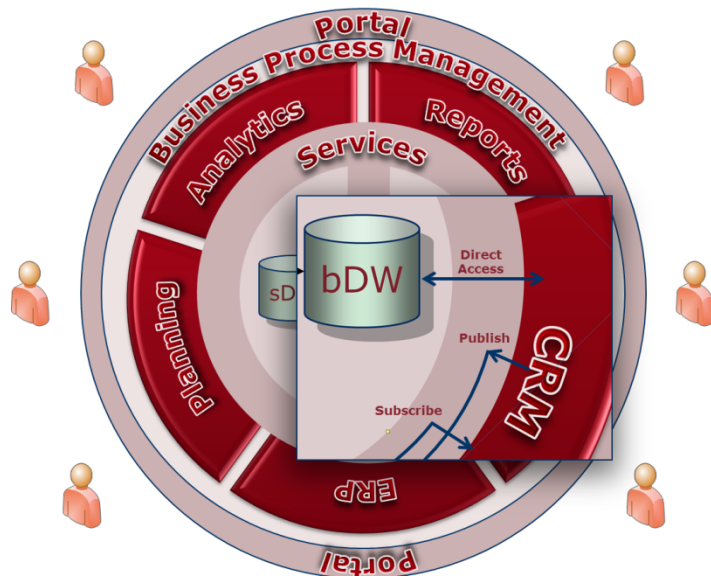
## Business integratie laag

Deze laag is in de figuur te herkennen als bDW of **business** data warehouse laag. Een deel van de data in de bronlaag, heeft behoefte aan wijziging/verrijking, om zodoende aan bepaalde uitwisselings- of rapportage behoeften te kunnen voldoen. Daarom wordt (een deel van) deze data getransformeerd en opgeslagen in de business laag. Datamarts worden dan typisch afgeleid van zowel de bron- als de businesslaag. Het aldus kunnen beschikken over voorbereide, afgeleide data in de businesslaag, voorkomt dat dezelfde afleiding keer op keer moet worden uitgevoerd richting de datamarts. Het onderscheid tussen de bron integratielaag en de business integratielaag is louter een logische. Beiden kunnen uitstekend bestaan in één technisch database schema.

## Services

Beide lagen worden omgeven door wat wij 'services' noemen in figuur 1. Alle data die erin en eruit gaan worden afgehandeld door deze services. Dat kan zowel Enterprise application integration (EAI) betreffen, als extractie, laad en transformatie (ETL, ELT), afhankelijk van de beschikbaarheidseisen van de gebruiker. Doordat deze laag als een soort van gatekeeper functioneert kan data gecontroleerd stromen, en is het te allen tijde mogelijk uitspraken te doen over de herkomst en

bestemming van de data alsook eventuele wijzigingen die de data heeft ondergaan. Het spreekt voor zich, dat ook tussen de bron integratielaag en de business integratielaag deze services van toepassing zijn.



Figuur 3. De rol van services in de architectuur

Via publish en subscribe mechanismen kan tussen applicaties informatie worden uitgewisseld. Daarnaast kan door applicaties data uit de bron- en/of de business integratielaag gehaald worden.

### Geen onderscheid tussen applicaties

Wat duidelijk wordt in deze figuur, is dat alle functies (of applicaties) gegroepeerd zijn rond het integratieplatform. Er is geen 'stroom van bron naar target' zoals we gewend zijn in traditionele data warehousing presentaties. Data kan vrij stromen tussen de verschillende functies.

Er zijn BI applicaties (planning, consolidatie) die nieuwe data creëren (het plan, journaalposten) die weer vereist zijn in ERP systemen voor verdere verwerking. Of wat gedacht van het CRM systeem dat verrijkte data nodig heeft die in de business laag verblijft en die nodig is om over een totaal klantbeeld te kunnen beschikken. Of wat gedacht van een klant die aan de telefoon een bestelling probeert te maken en waarvoor een controle op kredietwaardigheid in het ERP systeem moet worden uitgevoerd via EAI technologieën (omdat de data nog niet aanwezig is in het sDW of het bDW)?

Terugkomende op de vraag, waar de data blijft als applicaties meer en meer worden ontmanteld in services; Het mag nu evident zijn dat er in dit artikel voor wordt gepleit de data centraal op te slaan in de bron- en de business integratielaag ter ondersteuning van de functies als genoemd. Deze benadering maakt een organisatie minder afhankelijk van haar bronsystemen. Het is zelfs denkbaar dat een bron eenvoudiger kan worden vervangen doordat de initiële vulling vanuit het integratieplatform kan worden geregeld.

### Governance

De vraag wordt wel eens gesteld hoe het nu zit met de verantwoordelijkheid voor deze data? Wie is de eigenaar van de data? Het antwoord op deze vraag staat redelijk los van deze architectuur overwegingen. Tenslotte is het zo, dat bepaalde rollen in een organisatie verantwoordelijk zijn voor bepaalde attributen van een bedrijfsobject als klant, medewerker, product et cetera. Zelden is iemand verantwoordelijk voor alle attributen betreffende een bedrijfsobject. Neem bijvoorbeeld het bedrijfsobject 'product'. Typisch is sprake van attributen die ontstaan in het R&D proces, het is dus

logisch dat er een verantwoordelijke is binnen dit domein. Evenzo zijn er productattributen nodig ter ondersteuning van het productieproces, de voorraadvorming, of de sales en marketing ervan. Zo kan het gebeuren dat er honderden product attributen zijn met meerdere eigenaren. Qua governance ligt het dan ook voor de hand, dat alle eigenaren verantwoordelijk blijven, en zeggenschap houden over deze data. Ook als deze verhuist naar een integratieplatform, of als deze wordt gebruikt om nieuwe data af te leiden. Dit dwingt overleg af over het gebruik van deze attributen, in bronnen en rapportagesystemen.

### **Data vault**

Terug naar de gemeenten, ontdekten we al snel dat zij een data model zochten dat geschikt is voor verzameling en distributie van data, als nodig voor de doeleinden hierboven omschreven. Na het data vault modelleer concept te hebben toegelicht, waren ze blij verrast te horen over deze elegante manier van data modelleren, die zo geschikt lijkt voor hun specifieke behoeften. Vooral het feit dat data vault geleidelijke uitbreiding van het model toestaat, zonder de noodzaak bepaalde zaken opnieuw te moeten bouwen als nieuwe data elementen worden toegevoegd.

Een ander belangrijk aspect waar het gemeenten betreft, is dat zij goed in staat zijn de juiste business keys te identificeren, in relatie tot de zaken waar zij verantwoordelijk voor zijn (bv burgers, gebouwen, locaties). Dit verhoogt de kans van succes, waar het toepassen van data vault betreft.

Een derde, zeer interessant aspect, is gelegen in het feit dat data vault zo generiek is qua data modellering, dat het schreeuwt om automatisering. Dit wordt nog eens versterkt, door het feit dat toepassing ervan het aantal tabellen enorm doet groeien.

### **Quipu**

Vandaag de dag mag je stellen, dat er een aantal zeer interessante innovaties op de markt zijn om data vault te ondersteunen. Er is zelfs een zeer succesvol open source hulpmiddel genaamd Quipu ([www.datawarehousemanagement.org](http://www.datawarehousemanagement.org)) dat al door een aanzienlijk aantal, toonaangevende, grote ondernemingen wordt gebruikt. Quipu genereert de staging laag, de data vault lagen en de data mart laag voor het beschreven informatie of data warehouse platform. Dit helpt enorm in het verkorten van project doorlooptijden.

Natuurlijk moet de nodige aandacht worden besteed aan analyse en ontwerp. Daar staat tegenover dat dan de verschillende data modellen en bijbehorende laadcode kan worden gegenereerd in plaats van dat het moet worden geprogrammeerd door een consultant. De afhankelijkheid van vele externe consultants wordt hierdoor significant teruggebracht. Het gros van de inspanning kan zich concentreren op het ontwikkelen van het meest gecompliceerde stuk te weten de meer complexe transformaties tussen de bron datawarehouse laag, de business datawarehouse laag en de datamarts.

### **Over de auteur**

drs Pieter Rambags is managing partner van Nippur([www.nippur.nl](http://www.nippur.nl)). Hij is in 1989 afgestudeerd aan de Universiteit van Tilburg aan de vakgroep Bestuurlijke Informatiekunde. In 2002 heeft hij samen met partner Peter Kurstjens Nippur opgericht.

Nippur consultants nemen leidende rollen in BI projecten bij multi-nationale en grootzakelijke ondernemingen. Deze rollen zijn enterprise BI architect, BI projectmanager en BI informatie analist. Nippur consultants doceren de Adapt methodologie. Een techniek onafhankelijke methode die organisaties helpt hun BI informatie behoeften in kaart te brengen ([www.olapmodeling.org](http://www.olapmodeling.org)). Nippur helpt bedrijven bij het definiëren van de BI strategie. Een bewezen framework en aanpak die helpt bij het succesvol kunnen uitvoeren daarna van BI projecten.