

RDC ziet zonnige toekomst voor Tamino

# XML spin in SOA-web

Hans Lamboo

**Op welke manier wordt een native XML database in de praktijk ingezet? Het RDC Datacentrum in Amsterdam werkt nu een half jaar met Tamino van Software AG. Nu nog slechts 500 MB, maar de verwachtingen voor de toekomst zijn groot.**

Sprekend over XML met Richard van der Nat en Danny Kruidbosch, bij RDC verantwoordelijk voor respectievelijk IT-infrastructuur, en IT-architectuur en Netwerkontwikkeling, komt een aantal opmerkelijke feiten aan het licht. Zo blijkt men al 2,5 jaar geleden begonnen te zijn met de inrichting van een Service Oriented Architecture (SOA).

"We hebben zo'n 15 Terabyte aan data in voornamelijk DB2 en een mainframe-omgeving die draait onder VM-VSE, daaronder draait IDMS", zegt Richard van der Nat. "Het werd een jaar of drie geleden echt tijd om die gehele infrastructuur eens onder de loep te nemen." Er werd weinig toekomst gezien voor het mainframe, wat bevestigd werd door een onafhankelijk extern advies. "Vaak hoor je over mainframes dat ze fantastisch zijn, zoveel beschikbaarheid, zo hoogwaardig. Maar daar gaat het niet om, het gaat om de software die erop draait en die het vaak laat afweten. Onze Klanten wensen meer flexibiliteit en snellere 'time to market' werd noodzakelijk. Bovendien is een mainframe niet goedkoop en gaat de laatste mainframebeheerder die ik ken volgend jaar met pensioen." Dus werd besloten de gehele infrastructuur te herontwerpen. Van der Nat koos intuïtief voor een service-georiënteerde aanpak en architectuur-denken. "We liepen daarmee voor de muziek uit, ja. We hebben ook mensen van buiten RDC en onze huisleverancier IBM mee laten denken. Resultaat: een SOA moest het worden", vertelt Van der Nat.

## Spannend

Er moest een plan gemaakt worden hoe de oude wijn in nieuwe zakken kon worden gedaan, zoals Van der Nat dat noemt, en hoe die nieuwe zakken er dan precies uit moesten zien. Van der Nat besloot om met zijn eigen mensen de inrichting van en de migratie naar de SOA te gaan doen. "Niks met z'n allen naar de hei en dan een half jaar overleggen, maar gewoon: handen uit de mouwen, aan de slag. Ik wilde de mensen laten meegroeien in het nieuwe concept", legt Van der Nat uit. Tegelijkertijd moest

worden voldaan aan hoge beschikbaarheid van de operationele systemen. De inrichting van de SOA ging dus met horten en stoten – maar er werd wel een op Websphere J2EE gebaseerde SOA gerealiseerd. "Het valt niet altijd mee om de mensen mee te krijgen naar de wereld van het infrastructuurdenken. Maar er staat nu wel wat. Ik denk dat we redelijk vooroplopen in de wereld van SOA", stelt hij. "We zijn heel dapper geweest. Nu terugkijkend kan ik heel arrogant, als de Amsterdammer die ik ben, zeggen dat het goed gaat. We hebben nu een stevige basis voor de toekomst, een fundament waarvan we weten dat het goed gaat werken."

## Tamino is straks gewoon een SOA-component

Alle functionaliteit wordt uit het mainframe gehaald, dat zo wordt uitgefaseerd. Nadat alle toepassingen zijn 're-designed' en 'reployed' in de nieuwe omgeving, gaat de knop van het mainframe om. Van der Nat verwacht dat dat aan het eind van dit jaar is, als de SOA geheel operationeel is. "En dat is pas het begin, omdat we dan niet meer gedaan hebben dan één op één onze toepassingen overbrengen naar de nieuwe SOA-omgeving."

Op dit moment zit de SOA van RDC in de acceptatiefase. Er zijn enkele klanten gemigreerd die nu parallel draaien aan de oude omgeving. Het blijft voorzichtig handelen, want tijdens de verbouwing blijft de winkel gewoon open; op zich al een zware

opgave in een omgeving met bijzonder grote transactie-volumes: zo vinden er op piekmomenten zo'n 70 APK-meldingen per minuut plaats vanuit de 15.000 automotive-klanten.

"We vinden ook dat onze SOA relatief weinig heeft gekost ten opzichte van de gehoorde miljoenen die andere, grote bedrijven hebben moeten investeren. We hebben onze SOA-omgeving gerealiseerd tegen een redelijke prijs", zo meent Van der Nat. Er breekt nog een spannende tijd aan voor hem en zijn mensen, want de komende maanden moeten alle klanten gemigreerd worden naar de SOA.

## Native XML

In een grote relationele omgeving lijkt XML database Tamino een vreemde eend in de bijt. Danny Kruitbosch vindt van niet. "Het idee voor een native XML database kwam bij ons op, door de informatie die van de Rijksdienst voor het Wegverkeer wordt afgenomen. Het gaat daarbij om de kenteken- en voertuiggegevens en Europese typekeuringen. Die laatste groep bevat erg veel technische informatie. Van beide krijgen wij de bestanden in XML formaat en die willen we het liefst in het originele formaat opslaan. Want je wilt al die grote bestanden toch niet in een relationele database opslaan, om het vervolgens weer terug te converteren naar XML."

Tamino is nu een half jaar in productie. De RDW voorzagt RDC van de complete historie die werd opgeslagen in de XML database. "Eigenlijk is dat nu nog het enige dat erop draait. En in onze oude omgeving trekken we informatie uit de XML database, vertalen dat naar een relationeel model en slaan dat op in een relationele database. Nóg wel", benadrukt Kruitbosch, "het is een tussenoplossing omdat we nu middenin het redesign-project zitten. We moeten eerst het mainframe uitfaseren en dan gaan we verder met de uitbouw van de XML database. Omdat we straks een SOA gebruiken hebben we veel interfaces in XML geschreven. Webservices, maar ook bestanden die in een XML formaat worden uitgewisseld. Die gaan we straks ook allemaal in het native formaat opslaan."

Query's op de XML data vinden nu dus plaats in de relationele database, maar het ligt wel degelijk in de bedoeling dat uitsluitend nog in de XML database te gaan doen. Volgens Kruitbosch komt dat vooral door de SOA: "Je hebt een interface naar de data, maar hoe je die data daar uithaalt maakt niet uit. Een reden temeer om XML data in het native formaat op te slaan en te bevragen. Tamino is straks gewoon een SOA-component." De hele historie van de Europese typekeuringen omvat nu zo'n 500 MB. Het zijn geen grote bestanden en er komt nu nog niet dagelijks nieuwe informatie bij. Maar er zijn inmiddels veel meer XML modellen die binnenkort ook in de XML database zullen worden opgeslagen.

"De SOA hangt om het zo te zeggen van XML aan elkaar", zegt Kruitbosch. "Als we informatie binnen de SOA verplaat-

sen gaat dat allemaal via XML. Soms zijn het webservices, soms JMS, soms file-uitwisseling. Met Tamino zijn we klaar voor de toekomst. De uitwisseling van berichten met de RDW was de aanleiding om een native XML database aan te

## RAI Documentatie Centrum

Het RDC Datacentrum werd 40 jaar geleden in het leven geroepen door de Nederlandse vereniging van Automobiel Importeurs, de RAI. Behalve dat de RAI een eigen tentoonstellingshal voor de autobranche liet bouwen – de fameuze Amsterdam RAI – was het ook nodig de steeds groeiende hoeveelheid documentatie over de automarkt in het geheel en de automobielen in het bijzonder overzichtelijk te bewaren. Hiertoe werd het RAI Documentatie Centrum opgericht, kortweg RDC. Deze regelde tevens centraal de afgifte van kentekenbewijzen voor nieuwe auto's via de Rijksdienst voor het Wegverkeer (RDW).

Toen de APK-keuring zijn intrede deed, werd het nodig een centrale service-organisatie in te richten die onder meer de elektronische aan- en afmelding – in aanleg met Videotext – van APK-keuringen bij de RDW ging regelen. Het RDC was op dat moment klaar met het digitaliseren van alle informatie en de opslag daarvan in databases. Samen met de andere belanghebbende bij de APK-keuring, de BOVAG, werd het RDC Datacentrum opgericht, waarvan de RAI Vereniging en de BOVAG allebei ongeveer voor de helft eigenaar zijn. De communicatie tussen de automotive-branche en de RDW is lange tijd voorbehouden geweest aan het RDC Datacentrum, waar men kon beschikken over de volledige kentekenadministratie inclusief de persoonsgegevens van de eigenaar.

Kernactiviteit van het RDC Datacentrum (RDC) is het verzamelen van data uit de automotive-branche en het verschaffen van statistische gegevens aan importeurs, dealers, fabrikanten, schadeherstellers en garagebedrijven.

In het jaar 2000 raakte het RDC zijn monopoliepositie kwijt, doordat in het kader van de wet op de privacy de NAW-gegevens van kentekenhouders niet langer aan derden mogen worden verstrekt – deze informatie probeert het RDC thans via allerlei andere kanalen te verzamelen.

Daarnaast exploiteert RDC een ASP-farm voor zo'n 3.000 gebruikers, waar echter vooral MKB'ers uit de financiële sector gebruik van maken, en niet de beoogde en verwachte MKB'ers uit de automotive branche. Er worden zo'n 100 producten aangeboden, waaronder de MS Office Suite. Als derde activiteit biedt RDC hosting en housing aan, ook voornamelijk aan bedrijven van buiten de automotive-branche.

RDC heeft ongeveer 17.000 klanten, waarvan 15.000 uit de automotive-branche.

schaffen; we hebben daardoor de gelegenheid gekregen om de technologie, het concept te leren kennen."

## Meer redenen

Ook het persoonlijke gedeelte van het kenteken, dat vroeger dus deel uitmaakte van de RDW-informatie, moet via allerlei kanalen aangeleverd gaan worden in XML formaat. Ook deze data zullen worden opgeslagen in Tamino, om er vervolgens verschillende validaties en kwaliteitscontroles op te doen. Dat gebeurt onder meer door de aangeleverde adressen te vergelijken met bestanden van partijen waar RDC een partnership mee heeft.

## Uiteindelijk wordt een waardeoordeel over het geleverde adres gegeven

Uiteindelijk wordt een waardeoordeel over het geleverde adres gegeven. Daar komt nog een moeilijkheid bij: de aanleverende partij kan bedingen of en zo ja aan wie het door hem verstrekte adres mag worden vrijgegeven. Ook hier is gekozen voor aanlevering in XML formaat, hoewel ook 'good old' CSV wordt ondersteund. Ook deze informatie zal in Tamino worden opgeslagen.

"We geven eigenlijk invulling aan de ideeën van iemand als Frank Cohen: als je binnen een SOA alleen maar XML verschuift, gebruik dan in je middle-tier XQuery en XML databases. Dat is waar we naartoe groeien", stelt Kruitbosch. "Je kunt op de middle-tier bijvoorbeeld query's doen als je de XML database gebruikt als cache, waarin je eerder opgevraagde informatie nog eens *cached*. Die kun je er met XQuery heel snel uithalen. Ook een SOAP-bericht kun je heel makkelijk parsen met XQuery."

## Standaard stellen

Wellicht is de positie van RDC zo sterk, dat ze een standaard kunnen instellen voor de automotive-branch. Van der Nat is tamelijk sceptisch. "In Amerika heb je de standaard STAR. Maar Amerika is toch iets heel anders dan Nederland, waarin we het – ondanks de vele verenigingen – toch nooit met elkaar eens zijn. Als je kijkt naar de BOVAG, daarvan is ruwweg tweederde van de automotive-bedrijven lid, eenderde dus niet. Van de RAI-vereniging zijn ook niet meer alle importeurs lid. Redenen waarom je een standaard moeilijk van de grond krijgt. Wij spreken bij RDC al jarenlang over een standaard datamodel voor een auto. De invoer van zo'n uniform datamodel kan erg lucratief zijn, want het verhoogt de efficiëntie enorm. Omdat er zoveel tegenstrijdige belangen zijn kunnen de marktpartijen het

niet met elkaar eens worden – voor een leasemaatschappij is een auto iets anders dan voor een garagebedrijf. We kennen in Nederland 3.000 dealers en 39 importeurs, er zijn 30 software-pakketten in omloop voor de automotive-branch. Iedereen doet het op z'n eigen manier. Het is aan ons om overall connectors voor te bouwen. En XML is daarvoor buitengewoon geschikt."

## Spin

RDC kent wel een zekere regulering van bewaarplicht, maar dat is veel minder complex dan in andere branches. Door de toename van de hoeveelheid persoonsgegevens verwacht Van der Nat wel steeds meer met regelgeving die opslag van bijvoorbeeld e-mail verkeer in native formats voorschrijft, geconfronteerd te worden.

"Als je kijkt naar onze markt en onze doelgroepen, is XML in de toekomst de spin in het web. Door de gevarieerdheid in communicatiestromen met klanten en toeleveranciers is dat onontkoombaar. De ervaring met XML die we inmiddels hebben en nog gaan opbouwen, levert ongetwijfeld nog veel meer nieuwe ideeën voor de toepassing van Tamino", besluit Van der Nat.

**Hans Lamboo** is hoofdredacteur van Database Magazine.

## Online archief Database Magazine

Database Magazine-lezer opgelet! Artikelen over onderwerpen als Datawarehousing, SQL, ETL, Business Intelligence, Relationale databases, modellering en nog veel meer vindt u in het Online Archief van Array Publications. Vaktijdschriften als Storage Magazine, Database Magazine, IT Service Magazine, Java Magazine en ons Oracle vakblad Optimize hebben hun artikelenarchief online gezet. Met een Google-achtige zoekstructuur vindt u snel wat u zoekt op [www.dbm.nl](http://www.dbm.nl)