

Mooie datawarehouse-faciliteiten zullen in de smaak vallen

Codename Viper – DB2 for LUW Version 9

Klaas Brant

Hoe IBM ooit aan de codename Viper (adder, verrader) is gekomen is mij eigenlijk een raadsel. Of het moet zijn dat men vindt dat de implementatie van XML, het hoofdthema van deze release, veel beter en geavanceerd is dan bij de concurrenten van DB2. Of dat waar is laat ik helemaal aan u over.

Feit is dat de implementatie erg bijzonder en geavanceerd is vergeleken met de XML mogelijkheden van Oracle en SQL Server. Maar er is meer dan 'alleen maar' XML. In dit artikel zetten we de highlights van versie 9 eens op een rijtje.

Native XML Store

In het vorige nummer van DB/M kwam XML al uitgebreid aan bod. DB2 is naast een SQL engine ook een XML engine geworden. Anders dan bij de concurrenten is er niet voor gekozen om XML in SQL te stoppen maar een heuse XML opslagmethode te ontwikkelen inclusief XML indexen en XQuery interface.

Het werd zaak om een nieuwe partitioning in een single node DB2 te introduceren

Een bijzonder mooie faciliteit is ook het feit dat een XML structuur door te geven is naar een stored procedure. Ook de mix van SQL en XQuery (zie DB/M nr. 6) geeft aan dat ondanks dat de engines gescheiden zijn, de query-talen toch goed in elkaar overvloeien. Ook alle ontwikkel-tools zoals Developer Workbench zijn aangepast om met XML data en structuren om te kunnen gaan. Een leuke nieuw tool, voor iedereen die nog niet zo thuis is in het Xquery-gebeuren, is XQuery Builder. Deze tool maakt op een grafische manier Xquery's zoals de diverse Query By Example tools dat doen voor SQL.

Agile development

Het is natuurlijk gevaarlijk als de leverancier zich zelf op de borst gaat kloppen dat het allemaal zo gemakkelijk gaat (agile). Wat is er dan allemaal zo gemakkelijk geworden? Om te beginnen is er

een nieuwe driver voor JDBC die niet alleen JDBC 3.0 compliant is, maar ook veel nieuwe faciliteiten heeft zoals pooling en SSL-ondersteuning. Ook is de integratie met Visual Studio 2005 verder uitgebreid en is het nu mogelijk om websites te bouwen met DB2 zonder coding te hoeven schrijven. Voor die mensen die met XML aan de slag willen is de sample database verder uitgebreid met XML data. Al met al denk ik dat de term *agile* een beetje te veel van het goede is. Het gemak waarmee applicaties gebouwd kunnen worden is met versie 9 niet echt met sprongen vooruit gegaan.

Autonomic Computing

Dit is sinds lange tijd een stokpaardje van IBM. Autonomic computing zorgt ervoor dat databases zichzelf aanpassen aan de omgeving en workload. Zoals bij mensen de hartslag omhoog gaat als zij een trap oplopen moet ook een database zich aanpassen als de workload veranderd is. Veel veranderingen op dit gebied komen uit de Stinger release (versie 8.2). In versie 9 worden kleine verbeteringen en verfijningen aangebracht. Zo kan men nu buffer pools automatisch sizen en is een aantal belangrijke parameters dynamisch geworden (prefetchers en page cleaners). Zoals wel vaker bij DB2 zijn enkele zaken die eerst optineel waren nu tot default gemaakt. Zo is de automatic statistics collections bij een CREATE DATABASE opdracht nu de default-optie geworden. Het doel van IBM met autonomic computing is om het aantal DBA-handelingen terug te brengen (lees de TCO naar beneden te brengen) en de database meer zichzelf te laten tunen. Dit is hard nodig, want zonder enige tuning is DB2 vaak een niet performante database. De optimale tuning kan pas gebeuren indien je weet waarvoor de database gebruikt wordt. Dus waarom niet de database zichzelf laten aanpassen aan de omgeving en workload? DB2 is met autonomic computing zijn tijd ver vooruit. In de volgende DB/M zal ik in een artikel verder uitleggen hoe de implementatie van de diverse autonomic computing componenten werkt.

Large database management (partitioning)

Het is altijd al mogelijk geweest om met meerdere machines een multi-node database te bouwen. In het verleden werd deze implementatie DB2/EEE genoemd. EEE stond voor Enterprise Extended Edition. Later is door IBM de naam aangepast naar Enterprise Server met Data Partitioning Feature (DPF); naar mijn smaak een ongelukkige keuze omdat er nu een nieuwe feature in DB2 is die

partitioning heet. De verwarring zal groot zijn. Door de komst van machines met enorme power, zoals meerdere processors en/of multi-core processoren, en de opkomst van zeer geavanceerde disk array's vinden veel mensen het niet wenselijk om de data op te slaan in een Multi-node shared nothing-omgeving (MPP). In een MPP-omgeving maak je van meerdere computers (nodes) één grote computer die de query's verwerkt. Een complexe en dure oplossing is het gevolg, en met de huidige power en disk-opslag zelfs onwenselijk. Maar op zich is partitioning erg bruikbaar. Veel databases kennen een vorm van partitioning die datawarehouse-bouwers graag gebruiken.

Of de mensheid zit te wachten op XML in databases is nog maar de vraag

Dus werd het zaak om een nieuwe partitioning in een single node DB2 te introduceren. Het grotere mainframe broertje van LUW heeft zoiets al zo'n 20 jaar. De ontwikkelaars in het Toronto lab introduceren in Viper de diverse opties zoals uit Informix. Daarmee is de oplossing op LUW beter dan de mainframe-oplossing. Zo zijn er, naast de diverse partitioning-mogelijkheden, mooie oplossingen met ATTACH en DETACH PARTITION die een roll-out en roll-in in een datawarehouse-omgeving gemakkelijk en flexibel maken.

Omdat de data zich nu op een enkele node bevinden komen oplossingen zoals Multi-Dimensional Clustering (MDC) ook in beeld. Dit gecombineerd met het feit dat de optimizer slim gebruik kan maken van de partitioning-structuur, maakt van versie 9 een enorm krachtige en betaalbare datawarehouse server. DPF bestaat nog steeds, maar ik denk dat IBM er niet veel meer zal verkopen. Ook bestaande DPF-klienten zullen gaan kijken of een single node server geen betere en gemakkelijkere oplossing voor ze is.

Nieuwe security opties

Na een succesvolle introductie van label based security op het mainframe is nu de beurt aan LUW om deze feature te implementeren. Bij label based security is er een hiërarchie in de security; een mooi systeem waarbij een hiërarchie bestaat in de rechten. Een voorbeeld is top-management versus middle management versus afdelingschef. Het top-management mag alles zien en de afdelingschef alleen maar wat van zijn afdeling is. Een ander mooi security feature is TRANSFER OWNERSHIP. Hiermee is een probleem, dat ook veel andere databases hebben, in DB2 verholpen. In het verleden was de creator de OWNER en kon dit niet worden overgedragen. Dat is nu anders, de owner kan de owner/rechten door middel van DDL overdragen aan een nieuwe owner.

DBA-faciliteiten

Er zijn verbeteringen aangebracht om een onderbroken database recovery gemakkelijk te kunnen herstarten en een restore (al dan niet redirected) te generen vanuit een oude backup image. Viper heeft nu de mogelijkheid om data te kunnen comprimeren. Ook deze faciliteit is reeds 20 jaar aanwezig in de mainframe DB2. Het is duidelijk dat de moderne krachtige processoren de weg openen naar veel mainframe features. DB2 LUW kan nu zeer grote databases aan (16 Terabyte of 1100 miljard rijen). Aan de operationele kant voegt IBM nu (eindelijk) de mogelijkheid toe om meerder versies van DB2 op één Windows server te hebben. Vroeger was dit niet mogelijk omdat de verschillende versies van DB2 DLL's gingen delen. Gelukkig is dit probleem verholpen.

Conclusie

DB2 LUW heeft op diverse fronten meer in zijn mars dan zijn grotere mainframe-broer. Dit is niet zo verwonderlijk, want het aantal nieuwe mainframe-gebruikers is minimaal. Hoewel er naast XML nog veel andere zaken in Viper zitten is toch XML hét thema van deze release. Viper is een mooie release, maar of de mensheid zit te wachten op XML in databases is nog maar de vraag. Ik denk dat de mooie datawarehouse-faciliteiten meer in de smaak gaan vallen.

Klaas Brant

Klaas Brant (kbrant@kbce.nl) is database-specialist en oprichter van KBCE b.v.

Online archief Database Magazine

Online archief

Online archief

Trefwoorden: zoek Zoekklas

U bent op dit moment niet ingelogd. [Inloggen]

Extra zoekcriteria:

Database Magazine

Alle magazines

Zoek in:

Alle velden

Titel

Auteur

Omschrijving

Jaar

Bladnummer

Datum:

van: Januari 1993

tot: juni 2006

Aantal artikelen per pagina: 25

[Array Publications ©](#) | [Disclaimer](#) | [Privacy statement](#)

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