

Teradata blijft high-end oplossing

# Kampioen ondanks DWH Appliances

Paul van der Linden

**Tussen 19 en 21 juni vond in een 'tropisch' Budapest de zevende Technical Conference plaats van Teradata. De datawarehouse-leverancier die door Gartner al langere tijd wordt gezien als een duidelijke leider heeft inmiddels te maken met het verschijnen van datawarehouse appliances.**

Hoe denkt Teradata hierover en over leveranciers als Netezza en Datallegro? Is het een bedreiging of een hype? En is Teradata zelf misschien 'de vader van alle datawarehouse appliances'? Tijdens de opening van de Technical Conference wist Hermann Wimmer, vice president voor EMEA, in zijn beste Engels te vertellen dat dit de laatste editie van het event is. Dit jaar staat er nog netjes onder de naam Teradata dat het een onderdeel van NCR is. Maar begin van het jaar werd duidelijk dat de banden tussen NCR en Teradata worden verbroken. Dat opent dus mogelijkheden voor een andere opzet.

## Teradata Beginner

Zover was het dit jaar echter nog niet. Verdeeld over zes tracks (Teradata Beginner, SQL Expert, Enterprise Architect, Analytical Expert, 'Active' Specialist en Warehouse Manager) werd gedurende drie dagen een ware stortvloed aan informatie over de genodigden uitgestort. De indeling in verschillende tracks zorgt er daarbij voor dat iedereen aan zijn trekken kan komen. Wie nog weinig van Teradata of de relationele database-theorie weet is bij 'Teradata Beginner' aan het goede adres. Het is een soort opleiding waar met behulp van korte lessen (en vragen en antwoorden!) de basis van Teradata wordt onderwezen. Gewapend met een heuse Student Guide kom je dan ook weer uit zo'n sessie. Deelname aan een beginnerssessie is een bijzonder goed idee.

## Fysiek afgesplitste datamarts zijn volgens Teradata overbodig

Teradata heeft namelijk een heel duidelijke visie op datawarehousing. En die visie vormt het uitgangspunt voor de ontwikkelingen die terug zijn te vinden in haar producten. Eén van die

uitgangspunten is gebruik van een genormaliseerd relationeel model. Dat is dus de opzet zoals die door Date, Codd en Inmon wordt gepropageerd. Natuurlijk kan ook met een stermodel á la Kimball worden gewerkt, maar dat is niet nodig. De hardware die Teradata gebruikt is namelijk snel genoeg om met grotere aantallen tabellen en joins geen problemen te hebben; waardoor de facto het betere beheer van een genormaliseerd model resteert.

Wie tegenstribbelt dat een stermodel een meer herkenbaar model oplevert moet bedenken dat dat betrekking heeft op de presentatie van de data. En zoals chief technical officer Stephen Brobst aangeeft; "We don't want to own the GUI".

Active datawarehousing is ook zo'n uitgangspunt. Met 'active' wordt hier bedoeld dat niet alleen de strategische en tactische, maar ook de operationele niveaus in de organisatie ondersteund moeten kunnen worden. Hetgeen ook betekent dat detailgegevens in het datawarehouse aanwezig moeten zijn om vragen direct te kunnen beantwoorden. Wie het nog niet vermoedde: fysiek afgesplitste datamarts zijn volgens Teradata overbodig.

Het enterprise datawarehouse kan het allemaal aan. En last but not least: het multi-temperature datawarehouse. Al naar het gelang het gebruik van data wordt voor een andere opslagmethode gekozen. Wie meer wil weten over het multi-temperature datawarehouse boft: Teradata kondigde tijdens de conferentie aan hier nu ook consultancy services voor te hebben.

## DWH Appliances

Vanaf het begin heeft Teradata ervoor gekozen om een combinatie van hardware en software te leveren. Het grote voordeel daarvan is een afgestemde samenwerking tussen beide componenten, waardoor een optimale prestatie wordt gerealiseerd. Eén van de argumenten waarmee Teradata dan ook schermt is een uitstekende schaalbaarheid. Doorgroeien naar grotere hoeveelheden gebruikers en grotere datavolumes betekent dat de prestaties van het systeem niet veranderen. Om dit voor elkaar te krijgen maakt Teradata gebruik van een zogenaamde massively parallel processing (MPP) architectuur. Deze maakt gebruik van standaard Intel XEON 32/64-bit processors. Opslag van data (storage) is eveneens onderdeel van de geboden oplossing. Teradata kan onder Unix, Linux en MS Windows draaien.

De combinatie van hardware, database en dataopslag is ook precies hetgeen DWH Appliance leveranciers als Datallegro, Netezza en Plumtree bieden. DWH Appliances staan momenteel sterk in

de belangstelling omdat ze hogere prestaties leveren tegen lagere kosten. Zo maken onder andere Wolters Kluwer, Orange, T-Mobile en Flora Holland gebruik van Netezza. Betekent deze ontwikkeling nu dat Teradata het vanaf het begin aan goed heeft gehad door te kiezen voor een combinatie van hardware, software en opslag, waar anderen zich alleen op software hebben toegelegd? En kan Teradata in die zin dan worden gezien als de 'vader van de DWH Appliances'? CTO Stephen Brobst is niet helemaal gelukkig met deze ongewenste felicitatie. "Nee, we zijn geen DWH Appliance. Ja, we hebben wel de kenmerken ervan maar de DWH Appliance vendors, Netezza, Datallegro, Plumtree, zijn allemaal datamart-oplossingen." Dit in tegenstelling tot Teradata, dat zich presenteert als een enterprise datawarehouse-oplossing. Dat betekent onder andere schaalbaarheid en dat is ook precies hetgeen de DWH Appliance volgens Brobst missen. "DWH Appliances kunnen niet schalen vanwege de high density disk drives. Het zijn gigantische scanning machines – ze indexeren de data ook niet". Moet de adoptie van DWH Appliances en de aandacht die ze nu in de markt krijgen dan gezien worden als de zoveelste hype? Dat blijkt ook niet zo te zijn. Brobst: "Natuurlijk is er wel een plaats voor DWH Appliances. Als datamart kunnen ze prima voldoen. Maar als Enterprise datawarehouse zijn ze niet geschikt."

### Teradata en ... Microsoft?

Eén van de presentaties had betrekking op de combinatie van Teradata en Microsoft (Enterprise Architect-track). Het is nu mogelijk om Analysis Services (Microsoft SQL Server) te koppelen met Teradata. Het gaat hierbij om de ROLAP-optie. Deze combinatie wordt voornamelijk ingezet bij nieuwe BI-applicaties bij gezamenlijke klanten. Een opmerkelijk partnership waarbij Teradata, een echte high-end oplossing, hand in hand gaat met Microsoft, dat niet te boek staat als technologische hoogvlieger. CTO Brobst is hier kort over: "Bij 2-tier klanten waar een lage prijs per seat een issue is komen we deze vraag tegen." Aan de andere kant van het spectrum wordt Bank of America door Scott Gnuau (Vice President en General Manager R&D) opgevoerd als hooggewaardeerde klant. Het datawarehouse van de Bank of

America krijgt maandelijks 4 Terabyte aan informatie erbij en heeft momenteel een omvang van 125 Terabyte. Tijdens de conferentie werd ook versie 5.1 van Teradata Warehouse Miner geïntroduceerd. Met behulp van Warehouse Miner kan datamining in de Teradata database zelf plaatsvinden. Ook is het mogelijk om de datamining software van derde partijen hiervoor te gebruiken. Op datamining-gebied werkt Teradata al langer samen met KXEN, SAS, SPSS en Fair Isaac. Al deze partijen gebruiken het op XML gebaseerde PMML (Predictive Model Markup Language) waardoor de analytische modellen makkelijk uit te wisselen zijn.

Waar het gaat om business rules engines wordt Fair Isaac door Stephen Brobst genoemd als 'best in class'-oplossing. Business rules engines ziet Brobst als een onderwerp dat pas aan de orde komt als je in de vijfde fase bent van het Teradata volwassenheidsmodel. Brobst: "Een business rules engine gaan we zelf niet bouwen."

### Conclusie

Teradata is een high-end datawarehouse-oplossing die door Gartner al enige tijd als leider (in DWH-servers en DWH-databases) wordt gezien. De combinatie van hardware, software en dataopslag zou tot de conclusie kunnen leiden dat Teradata de DWH Appliance avant la lettre is. Dat is echter tegen het zere been van CTO Stephen Brobst die DWH Appliances ziet als beperkt schaalbare datamarts. Teradata is daarentegen een enterprise datawarehouse dat met name bedoeld is voor organisaties met heel veel gebruikers en heel veel data. Dat steeds meer organisaties deze kenmerken vertonen zal Teradata zeker niet erg vinden. Maar haar focus en ontwikkelingen blijven liggen bij de allergrootste datawarehouses en dat betekent dus blijven meegroeien. Daar verandert de scheiding van NCR en de door Wimmer aangekondigde 'andere opzet' verder niets aan.

**Paul van der Linden** (Paul.PFH.vanderLinden@AtosOrigin.com) is senior consultant Data Warehousing/BI bij Atos Origin en geeft leiding aan Data Warehousing Cost & Lifecycle Management (CLM).

## Bouw mee aan de In Summa oplossing

**SSAS, SSIS, SSMS, SSRS....**

**Zie jij kansen i.p.v afkortingen? Wil jij onze juniors & mediors al jouw SQL 2005 kennis bijbrengen? Bekijk onze vacature voor BI&DWH Specialist op [www.insumma.nl/specialist](http://www.insumma.nl/specialist)**

