

NonStop bevat van Neoview

HP springt op de appliance-trein

Paul van der Linden

Dat de markt voor Business Intelligence (BI) en datawarehousing (DWH) volop in de belangstelling staat zal pure play vendors als Business Objects en Cognos niet zijn ontgaan. Mega vendors als Microsoft, SAP en Oracle hebben deze hotspot inmiddels ook ontdekt en oefenen vanuit hun basis van enterprise applicaties een steeds groeiende druk uit op de marktleiders.

Ook aan nieuwe aanbieders die op basis van beschikbare technologie (denk bijvoorbeeld aan 64-bit computing) intuïtieve en goedkopere oplossingen bieden is geen gebrek. QlikTech, Spotfire en Panoratio zijn hier sprekende voorbeelden van. En alsof dat nog niet genoeg druk is worden de pure play vendors nu ook vanuit een derde kant ingesloten: door de leveranciers van DWH appliances. HP Neoview past in deze derde categorie. Maar wat is een DWH appliance? En wat doet HP in datawarehousing en BI?

DWH appliances

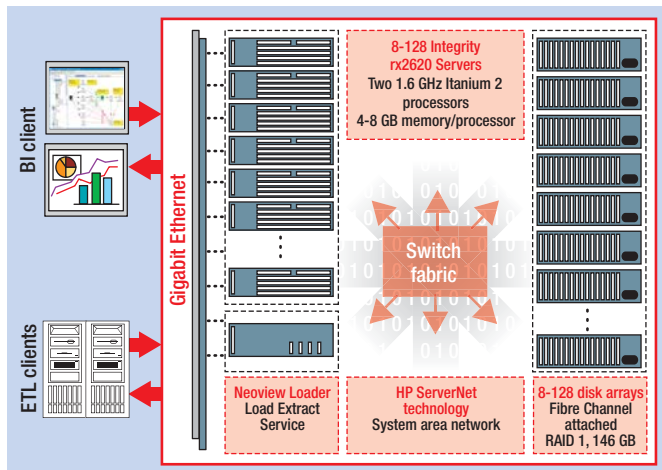
Steeds meer DWH-producten worden als DWH appliances in de markt gezet. Een appliance bestaat meestal uit de combinatie van hardware, opslag, operating systeem en database. Deze componenten zijn optimaal op elkaar afgestemd ten behoeve van BI. Netezza, DATAlegro maar ook IBM, SUN en HP zijn leveranciers die in deze categorie thuishoren. Een DWH appliance is in wezen een black box oplossing die direct uit de doos en van stroom voorzien operationeel is. Dat is in ieder geval de bedoeling. In de praktijk moeten er altijd nog wat handelingen gebeuren voordat de appliance 'up and running' is. Maar vergeleken met de reguliere datawarehouse-producten gaat het om beduidend minder voorbereidend werk en een veel snellere start. Dat is ook een van de voordelen van een DWH appliance: het is sneller operationeel en levert ook veel sneller resultaten (query's, rapporten etcetera) op.

Doordat gebruik wordt gemaakt van standaardcomponenten (processoren, schijven, software) is een appliance ook een stuk goedkoper dan een reguliere DWH-oplossing. Bedacht moet worden dat een appliance is ingericht voor een specifieke toepassing, waardoor het gebruik ook nog eens eenvoudiger wordt. Aan de andere kant is het dus wel minder generiek inzetbaar dan een reguliere oplossing. Maar de combinatie van veel sneller,

goedkoper en eenvoudiger is een garantie dat DWH appliances hun plaats zullen vinden in datawarehousing.

HP

Wie aan HP denkt, denkt aan printers en hardware. Maar er is een ander HP. HP claimt met een omzet van meer dan 91 miljard USD 's werelds grootste IT-bedrijf te zijn. De organisatie is inmiddels in meer dan 170 landen actief. Chris van Maaren, Business Manager voor HP Neoview, geeft aan dat de organisatie volgens marktresearcher IDC ook in Nederland veruit het grootste IT-bedrijf is. In Nederland zijn circa 2400 mensen werkzaam. Een belangrijk onderdeel van HP vormen de research-laboratoria (HP Labs), die zich bevinden in locaties zoals Palo Alto (US), Bristol (UK), Haifa (Israel), Bangalore (India) en Tokyo (Japan). Wie denkt nooit met HP in contact te komen vergist zich. Zo is waarschijnlijk iedere Nederlander HP-gebruiker. Geld pinnen uit een betaalautomaat (98 procent betaalautomatenverkeer), een telefoontje plegen op de vaste telefoon (40 miljoen telefoongesprekken per dag), een SMS-berichtje versturen, handelen op de Amsterdamse aandelenbeurs. Al deze transacties worden 24 uur per dag, 7 dagen in de week verwerkt door computersystemen van HP. Daarnaast gebruiken zeker zes van de tien Nederlandse technologiegebruikers een HP printer of HP PC (marktaandeel van 17 tot 63 procent). Veruit de meeste websites in ons land draaien op HP-systemen. Het merendeel van de Windows- en Linux-gebruikers vertrouwt op HP-systemen. Veel van wat hier wordt gepresteerd maakt gebruik van NonStopSQL, een stuk technologie dat in transactie-processing bewezen heeft zowel schaalbaar als betrouwbaar te zijn, ook bij enorme hoeveelheden data en gebruikers en bij complexe en gemengde workloads. Neoview, dat gericht is op datawarehousing, is gebaseerd op de uitgekristalliseerde en in de praktijk bewezen NonStop-technologie. Om tot HP Neoview te komen is rijk geleend aan



Afbeelding 1: Standards-based commodity components.

reeds beschikbare technologie die in de verschillende laboratoria is ontwikkeld. Goed beschouwd gaat het daarbij om vier bouwstenen. De eerste bouwsteen bestaat uit de servers en de disk array's. Hiervoor is gebruik gemaakt van de bestaande Integrity servers en van de StorageWorks Fibre Channel disk storage solutions. Bestaande elementen derhalve, die hun waarde reeds hebben bewezen en gebaseerd zijn op industriestandaarden. De tweede bouwsteen betreft het Enterprise Data Warehouse (EDW). Ook hier kon HP gebruik maken van bestaande technologie en wel van de clustering- en database-technologie van hun Integrity NonStop servers. Vervolgens zijn platform en software helemaal ingericht voor Business Intelligence (derde bouwsteen). Dit moet ervoor zorgen dat zowel de interfacing met de ETL-component als met de front-end tools subliem is. De vierde bouwsteen betreft het appliance-karakter van HP Neoview. Hiermee wordt bedoeld dat Neoview een black box is die voor verscheping naar de klant wordt geconfigureerd, geïntegreerd en getest. Installatie, setup, initialisatie en administratie moeten vooral eenvoudig zijn. HP installeert en initialiseert het product dan ook ter plekke en zorgt voor remote monitoring, remote support en analyse van incidenten. Routine-administratieve jobs worden waar mogelijk *gescheduled*. Daarnaast is er eenvoudig te bedienen software om het platform te monitoren en te beheren.

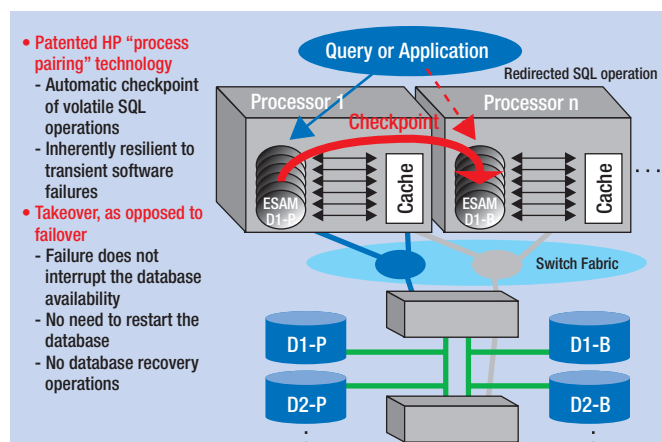
Wat zit er in de doos?

HP Neoview data warehousing platform, zoals het volledig heet, bestaat uit hardware, opslag, relationele database software, data services software (high speed load/unload, transaction management, backup & restore etcetera) en professional services (remote monitoring, remote support, incident analyses etcetera). In termen van hardware gaat het om een MPP-oplossing. De HP Integrity servers worden gebruikt voor de afhandeling van query's en database-processen. Afhankelijk van de opstelling is er sprake van tussen de 16 en 128 HP Integrity rx2620 servers. Elke server beschikt over twee 64-bit Intel Itanium 2 processors van 1,6 GHz. Elke processor heeft 4 of 8 GB aan main memory en 3 GB aan secondary cache. Wie moet groeien

kan dit doen in stappen van 16 processoren (en bijbehorende schijven). Voor het laden en extraheren van data maakt de Neoview Loader gebruik van HP ProLiant servers. Deze servers draaien onder Red Hat Linux Advanced Server. Voor de opslag van gegevens wordt een beroep gedaan op HP StorageWorks Fibre Channel Disk Array Modules (FCDM) met EVA JBOD-schijven. De hoeveelheid data die kan worden opgeslagen ligt tussen 3 TB en 55 TB. Alle data (zowel gebruikersdata als systeemdata) worden automatisch volgens RAID-1 opgeslagen. Tussen de ProLiant SMP server en de Integrity server in staat de HP switch fabric. Deze switch fabric is gebaseerd op HP ServerNet technologie en verzorgt de processor-naar-processor communicatie en de disk IO-bandbreedte. Het geheel wordt gecompleteerd met een virtual tape server (database- en log-backups ten behoeve van database recovery) en Gigabit Ethernet voor snelle ODBC, JDBC en Neoview Loader-connecties.

Consequenties

HP is ervan overtuigd dat HP Neoview een belangrijke plaats verdient in datawarehousing. Ze zijn dan ook bereid om garanties te geven op de prestaties ervan. Dat is mooi. Organisaties moeten er wel rekening mee houden wat inzet van HP Neoview in hun datawarehousing-landschap betekent. HP Neoview verzorgt niet de ETL-component. Daarvoor is een apart ETL-product nodig zoals bijvoorbeeld IBM DataStage, BO DataIntegrator of Informatica PowerCenter. De resultaten van de ETL-slag kunnen wel worden doorgegeven aan Neoview. HP Neoview levert ook niet de front-end tooling. Dit wordt wijselijk overgelaten aan partijen als Business Objects, Cognos, SAS en anderen. Deze tools kunnen hun query's aan Neoview aanbieden, die dat dan weer in razend tempo uitvoert en de resultaten terugsluist. Een opzet die in praktijk zeker werkt en goede resultaten kan opleveren. Twee opmerkingen kunnen hierbij worden gemaakt. In datawarehousing zit de grootste inspanning in het ETL-stuk. Dit is ook het meest weerbarstig stuk waarin vele datastromen bij elkaar worden gebracht en gefilterd, getransformeerd en verrijkt moeten worden. Hier gaat veel werk en tijd inzitten. Wie nu beschikt over een supersnelle jongen als Neoview komt dan ook



Afbeelding 2: Fault-tolerant process pairing.

spoedig op de gedachte om deze ETL-jobs op de Neoview server te draaien om zo mee te kunnen profiteren van de daar beschikbare spierkracht en intelligentie. Dat lijkt een goed idee, maar wie dit implementeert krijgt van HP geen garanties over de prestaties die hij mag verwachten. Dat is ook logisch omdat HP Neoview geoptimaliseerd is om BI-query's af te handelen en niet om ETL-jobs uit te voeren.

De tweede opmerking die gemaakt kan worden is dat HP Neoview ook de database-component levert. Wie beschikt over een (laten we zeggen) Oracle database exporteert de gegevens naar Neoview en kan afscheid nemen van de Oracle DWH database en de bijbehorende DBA. Ook dit is een resultaat van de appliance (black box) gedachte achter HP Neoview. Waar HP Neoview uitstekend past is in DWH-omgevingen waar betrouwbaarheid, schaalbaarheid, snelheid, gemak en kosten een rol spelen. Door parallelle query-afhandeling, data partitioneringstechnologie, caching en een 'take-over'-technologie in plaats van de gebruikelijke 'fail-over'-technologie, kan Neoview al deze vereisten invullen. Op basis van de bewezen technologie van NonStop en aangevuld met vele DWH-specifieke oplossingen biedt Neoview in datawarehousing een soortgelijke oplossing als NonStop dat voor transactie-processing is.

Conclusie

HP Neoview is een DWH appliance die met name geschikt is voor omgevingen met veel data, veel gebruikers, complexe en gemengde workloads. Waar beschikbaarheid, betrouwbaarheid, schaalbaarheid, snelheid, gemak en kosten een rol spelen is Neoview een product dat in overweging genomen dient te worden. HP Neoview heeft veel te danken aan de bewezen technologie van NonStop, bekend uit de wereld van transactie-processing. Maar omdat Neoview voor datawarehousing en Business Intelligence is bedoeld zijn er ook veel specifieke oplossingen in meegenomen afkomstig uit HP's research-laboratoria. Als een echte DWH appliance levert Neoview hardware, opslag, besturingssysteem en database. ETL-tool en ontsluitings-tools worden niet meegeleverd, maar kunnen op Neoview worden aangesloten. Neoview komt uit een goed nest en heeft slechts twee uitdagingen. De eerste uitdaging is dat HP nu niet meteen geassocieerd wordt met datawarehousing en BI. Met de komst van Neoview is dat niet meer terecht. De tweede uitdaging bestaat eruit dat het product nog relatief nieuw is en derhalve op bepaalde punten nog moet groeien. De Winter Corporation heeft het dan bijvoorbeeld over de volwassenheid van algoritmes waarmee complexe en grote datavolumes afgehandeld worden. Beide uitdagingen liggen in elkaars verlengde. Als de eenvoud en snelheid van Neoview enige indicatie zijn voor de marktacceptatie, kan HP met Neoview binnen *no time* non-stop aan de bak.

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).

Hier geen nummer

0800-5432101

Werken bij Valid is werken voor een ICT dienstverlener waar persoonlijke aandacht nog de normaalste zaak van de wereld is. Voor onze collega's én voor onze klanten. Bij Valid krijg je wat je verdient: uitdagende projecten bij toonaangevende klanten, een uitmuntend salaris, een uitdagend bonussysteem en een individueel budget voor opleidingen en trainingen.

Ben je een ervaren BI Consultant, Oracle DBA of Software Engineer en toe aan een op het lijf geschreven uitdaging in Utrecht, Eindhoven of Maastricht? Neem dan contact op met Bart Meex via bovenstaand telefoonnummer of mail je CV naar work@valid.nl.

www.valid.nl

