

# MASTERCLASS

## Ontwerpen van de Nieuwe Generatie Datawarehouses

**Pullman Brussels Airport Hotel (Diegem), 27 en 28 mei 2009 (14-21u)**

Inclusief tools, technologieën en architecturen



### Doelstellingen van dit Seminar:

- Krijg een compleet beeld van de markt van datawarehousetools en van de selectiecriteria
- Leer technieken en vuistregels voor het ontwerpen van de logische en fysieke aspecten van een relationeel datawarehouse
- Verzamel advies over uw optimale datawarehouse-architectuur

### 3 hoofdthema's op dit Seminar:

- Selecteer de Juiste Tools
- Ontwerp Doordacht
- Denk aan de Performantie

**Gepresenteerd door Rick van der Lans (R20/Consultancy B.V.)**

# Waarom dit Seminar ?

Het ontwerpen van een datawarehouse behelst meer dan alleen maar het bepalen van een databasestructuur. Er moeten diverse soorten tools gekozen worden, waaronder OLAP-tools (On Line Analytical Processing), rapportage-tools, analytic applications, ETL-tools (Extract Transform Load), portals en databaseservers. Daarnaast moet er ook een architectuur gekozen worden: moeten we één groot centraal datawarehouse opzetten, gebruiken we datamarts of kiezen we voor een virtueel datawarehouse? Er moet ook nagedacht worden over integratie met andere systemen, zoals ERP, CRM en websites. Tevens moet er bestudeerd worden hoe er met metadata omgegaan wordt.

Kortom, het ontwerpen van een warehouse is vergelijkbaar met het oplossen van een lastige puzzel. Alle stukjes moeten haarscherp passen om een volledig en effectief datawarehouse te kunnen ontwikkelen. Helaas kan een datawarehouse niet worden gekocht. Omdat elk bedrijf anders is (verschillend qua eisen, gebruikersbehoeften, gegevensbronnen en tijdseisen), dient ieder zijn eigen warehouse te ontwerpen.

De afgelopen jaren is de aard van het datawarehouse aan het veranderen.

## Dit seminar richt zich op 5 nieuwe ontwikkelingen:

- 1 Ten eerste, jarenlang zijn datawarehouses gebouwd met een statisch karakter. Slechts periodiek (bijvoorbeeld elke week of elke maand) worden nieuwe gegevens aan het datawarehouse toegevoegd. Recentelijk beginnen gebruikers echter te vragen om gegevens die bijna 100% up-to-date zijn. We spreken dan van real-time of active datawarehouses. Deze nieuwe behoefte heeft grote invloed op welke tools ingezet kunnen en moeten worden, en op hoe data-warehouses ontworpen moeten worden.
- 2 Ten tweede, de oorspronkelijke gebruiker van het datawarehouse is de manager die beslissingen moet nemen. Tegenwoordig zien we ook dat andere gebruikersgroepen, zoals klanten, toeleveranciers en agenten, direct toegang krijgen tot het warehouse en analyses willen uitvoeren. Ook zien we niet-menselijke gebruikers, ofwel geautomatiseerde processen, die voorgeprogrammeerde beslissingen nemen. Deze processen vormen ook een nieuwe groep van gebruikers.
- 3 Ten derde, het werken met datawarehouses leidt tot data-integratie. Er bestaan ook andere vormen van integratie, waaronder applicatie-integratie. Op dit terrein is de Service Oriented Architecture (SOA) sterk in opkomst. Tijdens dit seminar wordt belicht wat de toegevoegde waarde is van het integreren van de SOA met het datawarehouse.
- 4 Ten vierde, naast het werken met gestructureerde data, vragen steeds meer gebruikers om ook ongestructureerde data te kunnen analyseren.
- 5 Ten vijfde, nieuwe technologie blijft ontwikkeld worden. Er zal onder andere ingegaan worden op datawarehouse-appliances, cubing services, business process engines, streaming database servers en enterprise service bussen.

## Voor Wie is dit Seminar Bestemd ?

- Datawarehouse-ontwerpers • Databasebeheerders
- Database-ontwerpers • Informatie- en systeemanalysten
- Iedereen die direct/indirect betrokken is bij het ontwerpen en ontwikkelen van datawarehouse-databases

*Een minimale kennis van relationele databasetechnologie en databaseontwerp wordt verondersteld.*

# Ontwerpen van de Nieuwe Generatie Datawarehouses

Pullman Brussels Airport Hotel (Diegem), 27 en 28 mei 2009 (14-21u)

## Volledig Programma - Dag 1

13.30-14.00u

Ontvangst, Registratie en Koffie/Thee

14.00u

### 1. INLEIDING

- Overzicht van de status van datawarehousing
- Introductie terminologie
- Van statische naar online ofwel actieve datawarehouses

### 2. LEVENSCYCLUS VAN HET DATAWAREHOUSEPROJECT

- Bepalen van een implementatiestrategie: top-down of bottom-up
- Analyseren van het beslissingsproces
- In kaart brengen van de huidige situatie
- Selecteren van een datawarehouse-architectuur: één groot centraal datawarehouse, verscheidene werkgroep warehouses (data marts), vele private datawarehouses, of een combinatie
- Kan Extreme Programming ingezet worden bij projecten ?

### 3. SELECTEREN VAN BUSINESS INTELLIGENCE TOOLS

- Zes groepen BI-tools: executive reporting, managed query, OLAP, data mining, BAM en spreadsheets
- De terugkeer van executive reporting met tools van o.a. Business Objects (SAP), Cognos (IBM), Information Builders, Hyperion (Oracle), Microsoft en SAS
- Marktoverzicht van OLAP tools van o.a. Cognos, Information Builders, Microsoft, MicroStrategy, Oracle en SAP
- Criteria voor het selecteren van OLAP tools
- De opkomst van analytic applications, ofwel "BI out-of-the-box"
- De rol van een Enterprise Information Portal in een datawarehouse, inclusief marktoverzicht

### 4. SELECTEREN VAN DATABASESERVERS

- Hoe geschikt zijn klassieke RDBMS-en, waaronder DB2, Informix, Netezza, Oracle, SQL Server, Sybase en Teradata
- Markt van datawarehouse-appliances, waaronder die van DATAlegro (nu Microsoft), Greenplum, HP, Netezza en Sun
- Overzicht van OLAP technologie geïmplementeerd in relationele databaseservers zoals DB2, Oracle11g en SQL Server
- Speciale database interfaces: MDX, XML for Analysis en OLE DB for OLAP
- OLAP- en datamining-functionaliteit in een RDBMS
- De opkomst van open source databaseservers

21.00 u

Einde van Dag 1

## Volledig Programma - Dag 2 (ontvangst en koffie/thee vanaf 13.30u)

14.00u

### 5. LOGISCH ONTWERP VAN DATAWAREHOUSES

- Verhogen van de flexibiliteit van een informatiemodel
- De vraag- versus de aanbod-gedreven analyse-aanpak
- Omgaan met feiten en dimensionale gegevens
- Identificeren van detailniveau's binnen dimensies
- Modelleren van historische gegevens: van statische naar dynamische databases
- Het star schema- en snowflake-ontwerp
- Ontwerpen vanuit universele gegevensmodellen
- De invloed van analytical applications op logisch ontwerp

### 6. FYSIEK ONTWERP VAN DATAWAREHOUSES

- Normalisatie of denormalisatie?
- Introduceren van kunstmatige unieke sleutels (surrogaten)
- Hoe omgaan met afgeleide (geaggregeerde) gegevens?
- Hoe werkt de Data Vault methode ?
- Historie implementeren door decomponeren van tabellen
- Nieuwe indextypes, zoals bitmap- en multi-table-index

### 7. DE KWALITEIT VAN GEGEVENS

- Hoe en waar dienen gegevens te worden opgeschoond ?
- Data-profiling tools om vervuilde gegevens te signaleren
- Overzicht van de markt van data-profiling tools
- Verschillen tussen data-cleaning en data-profiling tools
- Een speciaal produkt: AlphaBlocks van IBM
- Mogelijkheden / onmogelijkheden van data-cleaning tools

### 8. KOPIËREN VAN GEGEVENS - ETL TOOLS

- XML als taal om gegevens te kopiëren
- Marktoverzicht van ETL tools, incl. die van Ab Initio, IBM, Business Objects, Cognos, Informatica, iWay, Microsoft, Oracle en SAS
- Wat zijn datamart generatoren, zoals BI-Ready en Kalido ?
- Aspecten van kopiëren: extraheren, verplaatsen, filteren, opschonen, consolideren en laden
- Synchroon houden van centraal warehouse en datamarts
- ETL versus SOA: wat kiest u in welke situatie ?
- Webservices voor het binnenhalen van externe gegevens

### 9. METADATA EN MASTER DATA MANAGEMENT

- Het belang van metadata voor gebruikers
- Het verschil tussen technische en bedrijfsmatige metadata
- De opkomst van operationele metadata
- Wat is een Master Data Management systeem ?
- Tools voor MDM, waaronder die van IBM, Kalido, Oracle en SAP

### 10. SAMENVATTING, CONCLUSIES EN TRENDS

- Werken met ongestructureerde gegevens
- Webservices voor het importeren van externe gegevens
- Datawarehouses als informatiebron voor de buitenwereld

21.00 u

Einde van het Seminar



## Uw Lesgever:

Rick F. van der Lans is onafhankelijk adviseur, docent en auteur, en directeur van het Nederlandse R20/Consultancy B.V. in Den Haag (NL). Hij is gespecialiseerd in datawarehousing, database technologie, client/server architecturen, development tools, SQL en informatie-modellering.

Rick heeft veel grote bedrijven geadviseerd bij het opzetten van hun client/server en datawarehouse architectuur en bij het kiezen van hun database en development tools. Hij heeft de voorbije 10 jaar als spreker op conferenties in binnen- en buitenland (waaronder Europa, Midden-Oosten, de USA, Australië) een zeer goede naam opgebouwd. Hij werd tevens door diverse belangrijke software-leveranciers uitgenodigd om de keynote speech te geven. Zijn lezingen worden altijd met veel enthousiasme ontvangen en uitstekend geëvalueerd.

Rick is tevens voorzitter van het Database Systems congres, columnist bij Datanews [corporate] solutions en adviseur van de Nederlandstalige bladen Database Magazine en Software Release Magazine. Zijn populaire boeken, o.a. „Introduction to SQL” en „The SQL Guide to Oracle”, werden vertaald in verschillende talen.

## Prijs:

De prijs voor dit seminar is 1350 EUR (+ 21 % BTW), incl. deelname, handouts, diner, koffie/thee, en een bezoek aan de Datawarehousing Solutions Expo.

## Kortingen:

Bij gelijktijdige aanmelding heeft de 2de deelnemer van een bedrijf recht op 10%, de derde 20%, en elke verdere deelnemer 30% korting. Deelnemers van universiteiten en scholen hebben recht op 20% korting.

**Vroeg inschrijven wordt beloond : 20% korting bij inschrijving vóór 31 maart 2009 !!!**

# REGISTRATIEFORMULIER

*Fax dit naar (09) 241.56.56 of onze backup fax (09) 220.34.57*

*Vragen over dit seminar? Bel (09) 241.56.13 of e-mail [seminars@itworks.be](mailto:seminars@itworks.be)*

### Eerste Deelnemer

Mr. /  Mevr. Naam: \_\_\_\_\_ Functie: \_\_\_\_\_

### Tweede Deelnemer (ontvangt 10% korting)

Mr. /  Mevr. Naam: \_\_\_\_\_ Functie: \_\_\_\_\_

### Derde Deelnemer (ontvangt 20% korting)

Mr. /  Mevr. Naam: \_\_\_\_\_ Functie: \_\_\_\_\_

### Vierde Deelnemer (ontvangt 30% korting)

Mr. /  Mevr. Naam: \_\_\_\_\_ Functie: \_\_\_\_\_

### Vijfde Deelnemer (ontvangt 30% korting)

Mr. /  Mevr. Naam: \_\_\_\_\_ Functie: \_\_\_\_\_

Bedrijf: \_\_\_\_\_

Bedrijfsadres: \_\_\_\_\_

Tel: \_\_\_\_\_

Fax: \_\_\_\_\_

BTW nr.: \_\_\_\_\_

E-Mail: \_\_\_\_\_

Registreert voor het seminar:

**ONTWERPEN VAN DE NIEUWE GENERATIE DATAWAREHOUSES** (27 en 28 mei 2009, 14-21 u)

Stuur de factuur naar mijn:

Privé-adres:

Bedrijfsadres

Datum

Handtekening

*Stuur dit registratieformulier per post of fax naar: I.T. Works, I.I.C., Technologiepark 3, 9052 Gent, fax: (09) 241.56.56 of (09) 220.34.57. U ontvangt een bevestiging, factuur en routebeschrijving. Vragen i.v.m. het programma of de inhoud? Bel ons op (09) 241.56.13.*