

# Ontwerpen van de Nieuwe Generatie Datawarehouses

Hotel Sofitel Brussels Airport (Diegem), 12 en 13 april 2005 (14-21u)

Inclusief tools, technologieën en architecturen

Met Rick van der Lans

## I.T. Works organiseert ook:

- Reliable en Secure Web Services - 16 Maart
- Server and Storage Consolidation - 17 Maart
- Enterprise Search Solutions - 24 Maart
- ITIL - 14 April
- Logisch Databaseontwerp - 25-27 April
- Pragmatisch Modelleren met UML - 27-28 April
- Fysiek Databaseontwerp - 7-8 Juni
- Ontwerpen van Service-Oriented Architectures - 9 Juni

## Tijdens dit Seminar leert U:

- diverse vuistregels voor het ontwerpen van de logische en fysieke aspecten van een relationeel datawarehouse
- diverse technieken waarmee de effectiviteit en de efficiëntie van een datawarehouse kan worden verhoogd
- alles over Uw optimale datawarehouse-architectuur

## NIEUW in dit Seminar:

- Gespecialiseerde of klassieke database server ?
- Het Common Warehouse Model (CWM) en de relatie met XML Metadata Interchange (XMI)
- Volledig vernieuwde overzichten van tools voor metadatabeheer, business intelligence, ETL, enterprise portals, ....

# 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 **business intelligence tools, query monitor tools, copy management tools, portals en databaseservers**. Daarnaast moet er ook een architectuur gekozen worden: moeten we uitgaan van één groot centraal datawarehouse of gebruiken we datamarts ? Bovendien moet er over **integratie met andere systemen, zoals ERP, CRM en Websites**, nagedacht worden. Laatstgenoemde is belangrijk om de buitenwereld, bestaande uit klanten en leveranciers, via bv. het Internet toegang te geven tot het datawarehouse. Kortom, het ontwerpen van een warehouse vandaag is vergelijkbaar met het oplossen van een lastige puzzel. Alle stukjes moeten haarscherp passen om een volledig en passend datawarehouse te kunnen ontwikkelen.

Velen zijn inmiddels bekend met het datawarehouse concept. Bill Inmon's algemeen geaccepteerde definitie luidt als volgt: „A data warehouse organizes and stores the data needed for informational, analytical processing over a long time perspective. A data warehouse is a subject-oriented, integrated, time-variant, non-volatile collection of data in support of management's decision-making process.” Het is een database of een verzameling databases waarin zeer grote hoeveelheden gegevens worden opgeslagen. 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 conferentie is opgebouwd uit 3 hoofdonderwerpen:

1 Aan het selecteren van alle benodigde tools zal veel aandacht besteed worden. **Welke OLAP tools** dienen wanneer ingezet te worden ? Waarom is een **query monitor** noodzakelijk ? Gaan we voor een ETL- of EAI-tool (Enterprise Application Integration)? Wanneer dient een **Enterprise Information Portal** gebruikt te worden? Door het beantwoorden van dit soort vragen, krijgt de deelnemer tevens een goed beeld van de markt voor datawarehouse tools.

2 Het tweede hoofdonderwerp gaat in op het ontwerpen van de logische structuur van een datawarehouse. De spelregels zijn hier anders dan bij het ontwerpen van een productiedatabase. De 2 belangrijkste verschillen zijn **historische gegevens en aanpasbaarheid**. In tegenstelling tot datawarehouses bevatten productiedatabases weinig tot geen historische gegevens. Omwille hiervan en omdat vele analisten niet gewend zijn tijd te modelleren, zal dit onderwerp zeer uitgebreid ter sprake komen. Het tweede aspect, aanpasbaarheid, heeft betrekking op het feit dat de structuur van een datawarehouse niet statisch is, maar frequent verandert. Dit betekent dat de structuur zoda-

nig ontworpen dient te worden dat deze eenvoudig uitgebreid en gewijzigd kan worden. Er zal ook aandacht besteed worden aan CWM (Common Warehouse Model) en XMI (XML Metadata Interchange) van de OMG.

3 Vele krachtige OLAP en data mining tools zijn commercieel beschikbaar waarmee we gegevens uit het warehouse kunnen onttrekken. Maar **functionaliteit en gebruiksvriendelijkheid** zijn niet alles: queries moeten een **goede performance** hebben. Wat kunnen we en wat moeten we doen om een adequate performance te garanderen? Deze belangrijke vraag, die betrekking heeft op de fysieke specificaties van een warehouse, is het derde hoofdonderwerp tijdens dit seminar.

Opmerking:

**Dit seminar is volledig produkt-onafhankelijk**. De meeste technieken en tips zijn van toepassing op zowel Oracle, IBM DB2, MS SQL Server, Informix, Sybase, ...

## Voor Wie is dit Seminar Bestemd ?

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

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

# Ontwerpen van de Nieuwe Generatie Datawarehouses

Hotel Sofitel Brussels Airport (Diegem), 12 en 13 april 2005 (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 datawarehouses

### 2. LEVENSCYCLUS VAN HET DATAWAREHOUSEPROJECT

- Bepalen van een implementatiestrategie: top-down of bottom-up
- Analyseren van de bedrijfsdoelstellingen
- In kaart brengen van de huidige situatie
- Afbakenen van een studie
- 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 tools van o.a. Business Objects, Cognos, Information Builders, Hyperion en Microsoft
- Marktoverzicht van OLAP tools, inclusief o.a. Advantage OLAP, Applix iTM/1, BusinessObjects, Cognos PowerPlay, Hyperion BrioQuery en Essbase, Microsoft Analysis Services, MicroStrategy, Oracle9i en Oracle Discoverer
- Criteria voor het selecteren van OLAP tools
- De opkomst van analytic applications, ofwel "BI out-of-the-box"
- De rol van de Enterprise Information Portal, inclusief marktoverzicht

### 4. SELECTEREN VAN DATABASESERVERS

- Hoe geschikt zijn klassieke databaseservers, waaronder DB2, Informix, Oracle, SQL Server, Sybase en Teradata
- Wat is een in-memory databaseserver ?
- Overzicht van OLAP technologie geïmplementeerd in relationele databaseservers zoals DB2, Oracle9i en SQL Server
- Speciale database interfaces: XML for Analysis en OLE DB for OLAP
- OLAP- en datamining-functionaliteit in een RDBMS

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 aanbodgedreven analyse-aanpak
- Omgaan met feiten en dimensionele gegevens: fact-dimension diagramma's
- 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?
- Brugtabellen om hierarchieën „plat te slaan“
- Het nut van mini- en macro-dimensionale tabellen
- Historie implementeren door decomponeren van tabellen
- Nieuwe indextypes, zoals bitmap- en multi-table-index

### 7. OPSCHONEN EN KOPIËREN VAN GEGEVENS

- Hoe en waar dienen gegevens te worden opgeschoond ?
- Data-profiling tools om vervuilde gegevens te signaleren
- Overzicht van de markt van de data-profiling tools, waaronder Ascential ProfileStage, Trillium Discovery, Evoke Axio, FirstLogic IQ Insight en Business Data Quality ?
- Verschillen tussen data-cleaning en data-profiling tools
- Mogelijkheden / onmogelijkheden van data-cleaning tools

### 8. KOPIËREN VAN GEGEVENS - ETL TOOLS

- Marktoverzicht van ETL-tools, inclusief Ascential, Embarcadero, IBM, Informatica, iWay, Microsoft en SAS
- Aspecten van kopiëren: extraheren, verplaatsen, filteren, opschonen, consolideren en laden
- Synchroon houden van centraal warehouse en datamarts
- ETL versus EAI-tools: wat kiest u in welke situatie ?
- Webservices voor het binnenhalen van externe gegevens

### 9. METADATA IN HET DATAWAREHOUSE

- Het belang van metadata voor gebruikers
- Het verschil tussen technische en bedrijfsmatige metadata
- Integratie van metadata in uw warehousing architectuur
- Standaarden voor metagegevens, zoals CWM en XMI ?
- Tools voor het beheer van metagegevens, van o.a. DAG's MetaCenter, Informatica's SuperGlue en SAS

### 10. SAMENVATTING, CONCLUSIES EN TRENDS

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 1250 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.

# 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*

### 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** (12 & 13 april 2005, 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.*