

# HADOOP EN NoSQL: dé Technologie voor Big Data

Leveranciersafhankelijk overzicht van de bouwstenen  
voor uw innovatieve databasetoepassingen van morgen

*12 Juni 2014 (14-21u), Golden Tulip Brussels Airport (Diegem)*

## Waarom mag u dit seminar niet missen ?

Big Data, Hadoop, NoSQL, analytical SQL databaseservers, NewSQL, MapReduce, appliances zijn de populaire termen van vandaag. Is dit technologie voor alleen de grote, bekende websites, zoals Bol.com, Amazon, LinkedIn, Twitter en Facebook, of heeft deze technologie wel degelijk traditionele organisaties iets te bieden?

**In dit seminar worden deze nieuwe ontwikkelingen helder toegelicht en wordt duidelijk gemaakt hoe, waar en waarom ze ook zeer relevant kunnen zijn voor meer traditionele organisaties.**

Tijdens het seminar zullen voorbeelden van toepassingsgebieden van big data gegeven worden. Een belangrijk toepassingsgebied is uiteraard analytics. Big data applicaties maken vormen van analytics mogelijk waarvan veel organisaties momenteel alleen kunnen dromen. **Big data applicaties vereisen de opslag en verwerking van (soms extreem) grote hoeveelheden gegevens.**

Maar **welke gegevensopslagtechnologie dient ingezet te worden?** Bestaande, meer traditionele technologie haalt misschien niet de benodigde schaalbaarheid of alleen voor een onacceptabel hoge prijs. Organisaties kunnen ook kiezen voor een van de volgende oplossingen:

- Hadoop: een platform speciaal ontwikkeld voor big data systemen
- NoSQL databaseserver, zoals CouchDB, Cassandra en MongoDB
- Een analytical SQL databaseserver, zoals Aster Database, Oracle Exalytics, SAP HANA en Teradata
- Een NewSQL databaseserver, zoals Clustrix, GenieDB, NuoDB, Translattice of VoltDB

Tijdens dit seminar worden al deze nieuwe technologieën kritisch besproken. Voor- en nadelen worden toegelicht en mogelijkheden en technische eigenschappen worden besproken. Na het seminar heeft u een gedegen en compleet overzicht van deze snel veranderende wereld van gegevensopslagtechnologie om big data systemen te ontwikkelen.

## Concreet krijgt u antwoord op volgende vragen tijdens dit seminar:

- Wat wordt precies bedoeld met Big Data en NoSQL, en waarom kan dit nuttig zijn voor uw bedrijf ?
- Welke “flavours” van NoSQL bestaan er en wat zijn de bekendste producten ?
- Welke rol speelt Hadoop ?
- Wat zijn de belangrijkste eigenschappen, voordelen en nadelen van analytische database servers ?
- Hoe ontsluiten we de kracht van big data databases, Hadoop en NoSQL voor rapportering en data/tekst-analyse, en wat wordt bedoeld met de “SQL-ficatie van NoSQL” ?
- Waar gaat dit naar toe, m.a.w. wat mogen we nog verwachten van de Big Data, NoSQL, NewSQL en analytical database revolutie ?
- Wat is NewSQL en waarom/wanneer is dit nuttig voor uw bedrijf ?

## Voor wie is dit seminar ?

Dit seminar is bedoeld voor iedereen die alles wil leren over de nieuwste ontwikkelingen in de databasewereld, o.a.:

- database specialisten,
- database administrators,
- datawarehouse- en database designers,
- database- en business intelligence experts,
- consultants,
- technology planners,
- project managers,
- systeem- en infrastructuur analisten,
- ...

## Spreker



### RICK VAN DER LANS

is onafhankelijk adviseur, docent en auteur op de terreinen datawarehousing, business intelligence, applicatie-integratie en databasetechnologie. Als consultant heeft hij door de jaren heen veel grote bedrijven (nationaal en internationaal) geadviseerd bij het ontwerpen van hun datawarehouse-architectuur en bij het maken van een toolkeuze.

Rick heeft als spreker op conferenties een zeer goede naam verworven zowel in binnen- als

buitenland. Hij weet als geen ander een goede balans te vinden tussen op de praktijk toegesneden technologische en strategische zaken.

Hij schrijft voor het bekende BeyeNetwork.com. Tevens verschenen van zijn hand diverse artikelen in toonaangevende vakbladen.

Verschillende van zijn boeken, waaronder het populaire SQL Leerboek, zijn in vele talen gepubliceerd. Recent is van zijn hand verschenen Data Virtualization for Business Intelligence Systems.

Meer info over R20/Consultancy BV vindt u op [www.r20.nl](http://www.r20.nl)

## Registratievoorwaarden:

Dit seminar wordt georganiseerd door I.T. Works.  
U kan on-line inschrijven op [www.itworks.be](http://www.itworks.be)  
De deelnameprijs bedraagt 720 EUR (+21% BTW).

# VOLLEDIG PROGRAMMA: HADOOP EN NOSQL

## 13u30 REGISTRATIE, KOFFIE/THEE EN CROISSANTS

## 14u00 INLEIDING

- De beperkingen van klassieke SQL databaseservers
- Hoe belangrijk zijn de concepten declaratieve databasetaal en opslagonafhankelijkheid?
- Trends in de databasemarkt
- Het belang van meer data scalability en query performance

### BIG DATA: HYPE OF REALITEIT?

- Het inzetten van Big Data om een concurrentieel voordeel te krijgen
- De vier V's van Big Data: volume, variety, velocity, en variability
- Big Data is niet alleen voor Google en Yahoo, maar ook voor traditionele organisaties
- Verschillen tussen semi-structured, poly-structured, multi-structured en unstructured data
- Voorbeelden van Big Data: sensordata, (micro-)eventdata en clickstream data.

### DE WERELD VAN HADOOP TOEGELICHT

- De Hadoop stack: HDFS, MapReduce, Hive, HBase, YARN, ZooKeeper, Pig, HCatalog, enzovoorts
- Eigenschappen en consequenties van HDFS en bestandsformaten
- Alternatieve implementaties onder andere van MapR, Amazon (Hadoop as a service) en ScaleOut (Hadoop in-memory)
- Gebruik van MapReduce voor analytics en reporting
- Storm voor streaming data
- De rol van Cloudera, HortonWorks en MapR

## 15u45 KORTE PAUZE MET KOFFIE/THEE

## 16u00 NoSQL DATABASE SERVERS

- Classificatie van NoSQL databaseservers: document stores, column family stores, key-value stores en graph data stores
- Zijn NoSQL producten werkelijk databaseservers?
- NoSQL en Clusters, data replication en sharding
- Nieuwe transactiebegrippen, waaronder CAP theorie, eventual consistency, replication factor en partition tolerance
- Aggregate data model versus relationeel model
- Marktoverzicht, inclusief Cassandra, CouchDB, Cloudera, DataStax, InfiniteGraph, MongoDB, Neo4j en Riak

- Wat is de rol van in-memory datagrids bij big data?
- Enterprise NoSQL systemen, zoals InterSystems Caché en MarkLogic, combineren de voordelen van NoSQL met die van SQL

## 17u00 ANALYTICAL SQL DATABASESERVERS VOOR BIG DATA SYSTEMENS

- Classificatie van analytical SQL databaseservers
- De voor- en nadelen van kolomgebaseerde opslagstructuren
- Wat is in-database analytics en wat is de relatie met Google's MapReduce?
- Marktoverzicht, inclusief Actian Matrix en Vector, EMC/Greenplum, Exasol, HP Vertica, IBM/Netezza, InfoBright, Kognitio WX2, Microsoft PDW, Oracle Exalytics, SAP HANA en Sybase IQ, Teradata Appliances en Teradata Aster Database

## 17u45 DINER BUFFET

## 19u00 SQL-ON-HADOOP: BENADEREN VAN NOSQL DATA MET SQL

- Kopiëren van gegevens tussen Hadoop en SQL databases via ETL oplossingen, zoals die van Informatica, Sqoop en Flume
- Overzicht van SQL-on-Hadoop engines: Apache Hive, CitusDB, Cloudera Impala, Concurrent Lingual, Hadapt, InfiniDB, JethroData, MammothDB, Apache Drill, MemSQL, Pivotal HawQ, ScleraDB en Splice Machine
- Datavirtualisatie voor het combineren van NoSQL en SQL systemen, met producten van Cirro, Cisco/Composite, Denodo, Informatica, Red Hat en StoneBond

## 20u00 NEWSQL DATABASESERVERS

- NewSQL betekent high-performance transactie-georiënteerde SQL systemen
- Simpelere transactie mechanismen om scale-out te realiseren
- Het voorkomen van multi-table joins die niet geparalleliseerd kunnen worden
- Wat betekent de term geo-compliance ?
- Marktoverzicht inclusief Akiban, Clustrix, GenieDB, NuoDB, TransLattice, Pivotal SQLFire en VoltDB

## 20.45u SAMENVATTING EN AFSLUITING

## 21u00 EINDE VAN DIT SEMINAR.

# REGISTRATIEFORMULIER

Fax naar: (09) 241.56.56 Vragen ? Bel (09) 241.56.13 of Email seminars@itworks.be

Dhr. /  Mevr.

Naam : ..... Functie : .....

Bedrijf : .....

Bedrijfsadres : .....

Tel : ..... Fax : ..... BTW nr. : .....

E-mail : .....

Gaat akkoord met de voorwaarden, en neemt deel aan:

## Hadoop en NoSQL: de Technologie voor Big Data

12 Juni 2014 (14-21u)

Stuur de factuur voor 720 EUR (+21% BTW) naar mijn bedrijfsadres, tenzij hierna anders vermeld:

Extra info voor factuur (bv. bestelbonnr.): .....

Datum : ..... Handtekening : .....

Stuur dit registratieformulier per post of fax naar: I.T. Works, Innovatiecentrum, Technologiepark 3, 9052 Gent, fax: (09) 241.56.56.

U ontvangt per kerende een bevestiging, factuur en routebeschrijving. Uw gegevens worden door I.T. Works bvba in een bestand opgenomen zodat wij u op de hoogte kunnen houden van onze activiteiten. Overeenkomstig de wet van 8/12/1992, heeft u recht op inzage, correctie of schrapping ervan.

U kan ook on-line inschrijven op [www.itworks.be](http://www.itworks.be)