Big Data
Multi-Platform Analytics

(Hadoop, NoSQL, Graph, Analytical Database)

Presented By:

Mike Ferguson
Intelligent Business Strategies Limited

2 Day Workshop

London: 25-26 September 2014

Amsterdam: 29-30 September 2014

www.unicom.co.uk/bigdata

Organised By:

UNICOM

www.facebook.com/unicom.seminars
@UNICOMSeminars
www.youtube.com/unicomseminars

UNICOM Seminars
unicomseminarsltd.wordpress.com

UNICOM

Big Data

London
Amsterdam

OVERVIEW

This new two day workshop is aimed at getting Data Scientists, Data Warehousing and BI professionals up to scratch on Big Data, Hadoop, other NoSQL DBMSs and Multi-Platform Analytics. What is Big Data? How can you make use of it? How does it fit within a traditional analytical environment? What skills do you need to develop for Big Data Analytics? All of these questions are addressed in this new knowledge packed workshop.

AUDIENCE

IT directors, CIO's, IT Managers, BI Managers, data warehousing professionals, data scientists, enterprise architects, data architects

LEARNING OBJECTIVES

Attendees to this seminar will learn:

- What Big Data is
- How Big Data creates several new types of analytical workload
- Big Data technology platforms beyond the data warehouse
- Big Data analytical techniques and front-end tools
- How to analyse un-modelled, multi-structured data using Hadoop, MapReduce & Spark
- How to integrate Big Data with traditional data warehouses and BI systems
- How to clearly understand business use cases for different Big Data technologies
- How to set up and organise Big Data projects including skills
- How to make use of Big Data to deliver business value

MODULE 1: AN INTRODUCTION TO BIG DATA

This session defines big data and looks at business reasons for wanting to make use of this new area of technology. It looks at Big Data use cases and what the difference is between traditional BI and Data Warehousing versus Big Data

- What is Big Data?
- Types of big data
- Why analyse Big Data?
- The need to analyse new more complex data sources
- Industry use cases - Popular big data analytic applications
- What is Data Science?
- Data Warehousing and BI Versus Big Data
- Popular patterns for Big Data technologies

MODULE 2: AN INTRODUCTION TO BIG DATA ANALYTICS

This session looks at Big Data Analytical workloads, the technology components involved and how you can integrate these with existing DW/BI systems in a new architecture for end-to-end analytics and to enrich business insight. It also looks at how to preserve existing investment in data management and BI tools across DW and Big Data platforms

- Types of Big Data analytical workloads
- Streaming data analytics at high velocity
- Exploratory analysis of multi-structured data
- Complex analysis of structured data
- Graph analytics
- Challenges when managing and analysing big data
- Key components in a Big Data Analytics environment
- Preserving existing BI/DW investments
- The Big Data Extended Analytical Ecosystem

MODULE 3: BIG DATA PLATFORMS AND STORAGE OPTIONS

This session looks at platforms and data storage options for big data analytics

- The new multi-platform analytical ecosystem
- Beyond the data warehouse - Hadoop NoSQL and analytical RDBMSs, NewSQL DBMSs
- NoSQL DBMSs
- Key Value stores, Document DBMSs, Column Family DBMSs and Graph databases
- An introduction to Hadoop and the Hadoop Stack
- HDFS, MapReduce, Pig & Hive
- Hadoop 2.0 Spark Framework
- SQL on Hadoop options
- The Big Data Marketplace
- Hadoop distributions – Cloudera, Hortonworks, MapR, IBM BigInsights, Microsoft HD Insight, PivotalHD
- Big Data Appliances – Oracle Big Data Appliance, IBM PureData System for Hadoop, HP HA-VeN, Teradata Aster Discovery Server,
- NoSQL databases, e.g. Datastax, Neo4J, Yarcdata, MongoDB, Riak
- Analytical databases and DW appliances, e.g. Teradata, Exasol, IBM PureData, Oracle Ex-adata, SAP HANA, Kognitio, Actian ParAccel
- Analytical appliances – SAS LASR, MicroStrategy PRIME
- The Cloud deployment option – Microsoft Windows Azure, IBM, Amazon Elastic MapReduce, Altiscale Data Cloud
- Creating a multi-platform analytical ecosystem

MODULE 4: BIG DATA INTEGRATION AND GOVERNANCE IN A MULTI-PLATFORM ANALYTICAL ENVIRONMENT

This session will look at the challenge of integrating and governing Big Data and the unique issues it raises. How do you deal with very large data volumes and different varieties of data? How does loading data into Hadoop differ from loading data into analytical relational databases? What about NoSQL databases? How should low-latency data be handled? Topics that will be covered include:

- Types of Big Data
- Connecting to Big Data sources, e.g. web logs, clickstream, sensor data, and multi-structured content
- Supplying consistent data to multiple analytical platforms
- Loading Big Data – what's different about loading HDFS, Hive & NoSQL Vs analytical relational databases
- Change data capture – what's possible
- Data warehouse offload
- Tools for ELT processing on Hadoop – The Enterprise Data Refinery
- ETL tools Vs Pig Vs self-service DI/DQ
- Dealing with data quality in a Big Data environment
- Parsing unstructured data
- Governing data in a Data Science environment
- Joined up analytical processing from ETL to analytical workflows
- The impact of data scientist and end user self-service DQ/DI – Paxata, Trifacta, MS Excel, MicroStrategy
- Mapping discovered data of value into your DW and business vocabulary
- Big data audit, protection and security – Dataguise, IBM Guardium, Protegrity

MODULE 5: TOOLS AND TECHNIQUES FOR ANALYSING BIG DATA

This session looks at tools and techniques available to data scientists, business analysts and traditional DW/BI professionals to analyse big data. It looks how different types of developers and users can exploit Big Data platforms such as Hadoop and NoSQL databases using programming techniques, text analytics, search, self-service BI tools as well as how vendors are making it easier to gain access both the NoSQL/Hadoop world and the Analytical RDBMS world by using data virtualisation.

- Data Science projects
- Creating Sandboxes for Data Science projects
- Options for analysing unstructured content – Text analytics, c ustom MapReduce code and MapReduce developer tools
- Using R as an analytical language for Big Data
- Text analysis and visualisation, Sentiment analysis and visualisation
- Clickstream analysis and visualisation
- Analysing big data using MapReduce BI Tools and applications for Hadoop, e.g. Datameer, Karmasphere, Platfora, IBM Customer Insight
- Exploratory graph analysis and visualisations
- Using search to analyse multi-structured data
- Creating search indexes on multi-structured data
- Building dashboards and reports on top of search engine indexed content
- The integration of search with traditional BI platforms
- Guided analysis using multi-faceted search
- The marketplace: Apache Solr, Attivio, Cloudera Search, Connexica, DataRPM, HP IDOL, IBI WebFocus Magnify, IBM Watson Explorer, LucidWorks, Microsoft, Oracle Endeca Quid, Splunk
- Analysing Big Data using Self-Service BI Tools, e.g. Tableau, QlikView, Spotfire, SAS Visual Analytics, MicroStrategy, SAP Lumira
- SQL connectivity initiatives to Big Data – e.g. Impala, Hive, Stinger, Shark on Spark, HawQ, IBM BigSQL, CitusDB JethroData, Splice Machine
- Big data analytics – query performance enablers
- Managing stream computing in a Big Data environment
- Tools and techniques for streaming analytics

MODULE 6: INTEGRATING BIG DATA ANALYTICS INTO THE ENTERPRISE

This session looks at how new Big Data platforms can be integrated with traditional Data Warehouses and Data Marts. It looks at stream processing, Hadoop, NoSQL databases, Data Warehouse appliances and shows how to put them together in an end-to-end architecture to maximise business value from Big Data

- Integrating Big Data platforms with traditional DW/BI environments – what's involved
- Integrating stream processing with Hadoop and Analytical DW Appliances
- Integrating Hadoop with DW Appliances and Enterprise Data Warehouses
- Tying together front end tools
- Options for implementing multi-platform analytics
- Cross-platform analytical workflows
- The role of Data Virtualisation in a Big Data environment
- Multi-platform optimisation

PRESENTER

Mike Ferguson is Managing Director of Intelligent Business Strategies Limited. As an analyst and consultant he specialises in BI/Analytics, Big Data and Data Management. With over 32 years of IT experience, Mike has consulted for dozens of companies on BI, technology selection, Big Data, enterprise architecture, and data management. He has spoken at events all over the world and written numerous articles. Mike provides articles, blogs and his insights on the industry. Formerly he was a principal and co-founder of Codd and Date Europe Limited – the inventors of the Relational Model, a Chief Architect at Teradata on the Teradata DBMS and European Managing Director of Database Associates. He teaches popular master classes in BI, Big Data Analytics, Data Governance & Master Data Management

Big Data

London
Amsterdam

Registration

2 Day Workshop Fee:
£1095+VAT

1st Delegate
Dr/Mr/Ms/Mrs.....First Name.....
Surname.....
Position.....
Head of Department.....

2nd Delegate
Dr/Mr/Ms/Mrs.....First Name.....
Surname.....
Position.....
Head of Department.....

Contact Details
Organisation.....
Address.....
Post Code.....
Tel.....Mobile.....
Fax.....
Email.....

INVOICE TOTAL

☐ I enclose a cheque made payable to UNICOM Seminars

☐ Please charge my:

☐ VISA ☐ MasterCard ☐ Switch ☐ American Express

Card No.

☐☐☐☐ ☐☐☐☐ ☐☐☐☐ ☐☐☐☐

Expiry Date: ☐☐ / ☐☐

Purchase Order

☐ Please invoice (include invoice address)

F.A.O.....

Organisation.....

Address.....

.....

.....

Post Code.....

Tel.....Mobile.....

WHAT THE REGISTRATON FEE INCLUDES:
The registration fee for the training course or the event covers the following: attendance, copy of the documentation and materials, examinations where applicable and light refreshments. Accommodation is not included unless otherwise specified.
Joining instructions will be sent to you approximately one week before the event (if for any reason these are not received, please contact UNICOM).

PAYMENT TERMS:
Payment is required in advance of the event or at the latest, paid at the event.
All invoices carry a 10% surcharge, which is payable if the fee remains unpaid on the day of the event and 30 days thereafter; should the invoice remain unpaid beyond 30 days and up to 45 days the surcharge increases to 15% and if unpaid after 45 days the surcharge increases to 20%. For credit card payments a 2.5% fee amount is charged or for American Express cards the fee is 3% of the total amount.

CANCELLATION AND SUBSTITUTION TERMS:
What happens if I have to cancel? If you confirm your **CANCELLATION** in writing up to fifteen (15) working days before the event or training start date and if the invoice has already been paid you will receive a refund less a 10% + VAT service charge; if the invoice has not been paid at that point you the a credit note for the existing invoice will be raised and a new invoice raised for the 10% +VAT service charge – the service charge invoice is due for payment by the original event / training start date. Regrettably, no refunds can be made for cancellations received less than 15 working days prior to the event and the invoice will remain due. **SUBSTITUTIONS** are welcome at any time – there is no fee for sending a substitute delegate on any event or training. If it is more than 15 working days but less than 5 working days before the course or training start date, you may TRANSFER your registration to a future date within a 6 month period. If it is less than 15 working days to the event / training start date you can still TRANSFER your booking to a future event future date within 6 months but an additional transfer fee of £125+VAT per person per event day will be charged (e.g. the transfer fee for a 2 day training is £250+VAT), invoices for transfer fees are due for payment within 7 days of the invoice date invoice.
As we cannot guarantee that exactly the same event or training will be available, the transfer will be open to any other UNICOM event taking place within six months from the date of the original event. **TRANSFERS** are not accepted less than five (5) working days before the event or training unless there are exceptional circumstances and the acceptance of the transfer is at the discretion of UNICOM.
Where a transfer has been made and a future date selected, the standard cancellation terms and conditions apply to the transferred booking just as if it were a new booking.
UNICOM reserves the right to amend the event / training content programme if necessary and cannot guarantee repeats of the same event or training. All transfers and cancellations must be made in writing either by email or letter and are only valid when confirmed by email or on writing by UNICOM. Transfers and cancellations are not accepted by telephone.

INDEMNITY:
Should for any reason outside the control of UNICOM Seminars Ltd, the venue or the presenters change, or the event be cancelled due to but not exclusively to industrial action, adverse weather conditions, an act of terrorism, presenter illness or other reasons beyond its control UNICOM Seminars Ltd will make reasonable endeavour to reschedule, but the client hereby indemnifies and holds UNICOM Seminars Ltd harmless from and against any and all costs, damages and expenses, including attorneys fees, which are incurred by the client as a consequence beyond the attendance fee due to UNICOM. The construction validity and performance of this Agreement shall be governed by all aspects by the laws of England to the exclusive jurisdiction of whose court the Parties hereby agree to submit.

UNICOM