Training - Big Data and Cloud for Data Warehouse Professionals

Sonra Intelligence Limited

6-9 Trinity Street
Dublin 2
Ireland
hello@sonra.io
www.sonra.io

Big Data and Cloud for Data Warehouse Professionals.

How to solve common data warehouse headaches with big data concepts and technologies.

Audience: This course shows data warehouse professionals how they can benefit from big data technologies, tools, and concepts to solve common problems in their Enterprise Data Warehouse. It caters for data warehouse managers, architects, and developers. Anyone who is involved in data warehousing and business intelligence.

What you will learn: The traditional approach to data warehousing has served us well over the last 25 years. However, various cracks have shown up over the last couple of years and we as data warehouse managers, architects, and developers are faced with a multitude of problems.

- The data warehouse is too slow to turn around questions by the business. Anything between 3-9 months is spent to get a new subject area into the data warehouse. 80-90% of enterprise data never makes it into the data warehouse. Important business decisions are made without seeing the full picture.
- The digitisation of all aspects of life has led to an explosion of data volumes. Machine generated data fuels the flames of data growth. Our data warehouse is bursting at the seams. License costs are soaring.
- Traditional data warehouse architecture and technology does not cater well for certain workloads, e.g. unstructured data, streaming data, and graph data.
- Advanced analytics has always been a secondary concern in the EDW. This is somewhat strange as advanced analytics applications deliver the greatest ROI.

In this course, we will describe these headaches in detail and prescribe various remedies.

The first part lays the foundation. We will clarify important big data and distributed compute concepts that are crucial for understanding the solutions we prescribe.

In the second part we cover today's data warehouse issues in some detail.

In the third part we will prescribe the remedies.

In part four we summarise what we have learned and close with an overview on the next generation data warehouse architecture.

Benefits: The course will help you to achieve the following business objectives

- Turn business questions around faster. The concepts and tools presented in this course will help you reduce the time it takes to deliver data to decision makers.
- Democratize data and become data driven. We will show how to make more data available to more employees in your organisation.
- Save license and hardware costs by offloading workloads to cheaper storage technologies

- Use advanced analytics to predict outcomes, get to the root cause of problems, and generate new revenue streams
- Some decisions need to be taken within minutes or even seconds of an event happening. Credit card fraud detection is a good example. React to events as they happen.
- You will learn when to use Hadoop and when to use concepts or tools from traditional data warehousing.
- Become a data evangelist. Propel your data analytics career. Save your organisation millions.

Coursework: The training is based on slides. Some demonstrations will be delivered during the training.

Duration: 1 day

Pre-requisites: Understanding of data warehouse concepts and architecture.

Price: 700 Euros per participant (exclusive of VAT)

The Trainer: Uli Bethke

Uli has 17 years' hands on experience as a consultant, architect, manager, and developer in the data industry. He frequently speaks at conferences and is globally recognised as an industry thought leader. Uli has architected and delivered Big Data projects in Europe, North America, and South East Asia. He understands both the traditional world of data warehousing AND the *brave* new world of Big Data.

Uli is a regular contributor to blogs and books, holds an Oracle ACE award, and chairs the Hadoop User Group Ireland (~1,100 members, www.hugireland.org). He is also a co-founder and VP of the Irish chapter of DAMA, a non for profit global data management organization. He is also a co-founder of the Irish Oracle Big Data User Group.

He holds degrees from Freie Universität Berlin, Albrecht Ludwigs Universität Freiburg, and the University of Ulster, Coleraine.

Part I - Data Warehousing in the Age of Big Data. The end of an era?

Big Data Concepts

What is Big Data?
The evolution of Big Data processing paradigms and frameworks

Distributed Compute Concepts

Storage Types and Data Access Patterns
What is a distributed system? When should we use one?
Distributed Database Systems (MPP)
Distributed MapReduce Systems (Hadoop)
Distributed Dataflow Systems (Spark)

Part II - Data Warehouse Headaches

The trouble with the Enterprise Data Warehouse

Performance & License Costs
Project Timelines - Excessive Turnaround Times
Streaming Analytics
Advanced Analytics
Is the Enterprise Data Warehouse obsolete?

Part III - Data Warehouse Remedies

Data Warehouse Optimization

Staging Area Offload
ETL Offload
Smart Archiving
Query Offload
Approximate Query Engines
Federated Query Engines
Unstructured Data Offload
Reference Architecture, Tools, and Technologies

Streaming for Real-time analytics

Why real-time analytics?
Streaming Architectures
What is a queuing system? Why do we need it?
Apache Kafka – A Distributed Queue
Streaming Concepts
- Batch vs Streaming
- Stream Processing and State
- Processing Windows
- Processing Semantics
- Lambda and Kappa Architecture Demystified
- Types of Streaming Engines

Reference Architecture, Tools, and Technologies

Self-Service Analytics, Data Preparation, and the Enterprise Data Hub

The Concept of the Enterprise Data Hub
Data Hub vs Data Warehouse
The Concept of Data Preparation
Data Preparation Tools
Self-Service Analytics
Governance: Data Hub vs Data Swamp
Demo: Data Preparation in Action
Reference Architecture, Tools, and Technologies

Advanced Analytics & Data Science

Traditional Advanced Analytics Architecture
Data Science Platforms
From Big to Smart Data
Types of advanced analytics

Part IV – The Next Generation Data Warehouse

Skills, Roles
Data Warehouse in the Cloud
Reference Architecture, Tools, and Technologies