The Logical Data Warehouse - Design, Architecture, and Technology

Course Outline

1. Challenges for the Classic Data Warehouse

- Integrating big data with existing data and making it available for reporting and analytics
- Supporting self-service BI and self-service data preparation
- Faster time-to-market for reports
- Polyglot persistency – processing data stored in classic SQL, Hadoop, and NoSQL systems
- Operational Business Intelligence, or analyzing zero-latency data

2. The Logical Data Warehouse Architecture

- The essence: decoupling of reporting and data sources
- From batch-integration to on-demand integration of data
- The impact on flexibility and productivity – an improved time-to-market for reports
- Examples of organizations operating a logical data warehouse
- Can a logical data warehouse really work without a physical data warehouse?

3. Implementing a Logical Data Warehouse with Data Virtualization Servers

- Why data virtualization?
- Market overview: AtScale, Cirro Data Hub, Data Virtuality, Denodo Platform, FraXses, IBM Data Virtualization Manager for z/OS, RedHat JBoss Data Virtualization, Stone Bond Enterprise Enabler, and Tibco Data Virtualization
- Importing non-relational data, such as XML and JSON documents, web services, NoSQL, and Hadoop data
- The importance of an integrated business glossary and centralization of metadata specifications

4. Improving the Query Performance of Data Virtualization Servers

- How does caching really work?
- Using caching to minimize interference on transactional systems
- Speeding up queries by caching data in analytical SQL database servers
- Which virtual tables should be cached?
- Query optimization techniques and the explain feature
- Smart drivers/connectors can help improve query performance
- How can SQL-on-Hadoop engines speed up query performance?
- Working with multiple data virtualization servers in a distributed environment to minimize network traffic

5. Migrating to a Logical Data Warehouse

- An A to Z roadmap
- Guidelines for the development of a logical data warehouse
- Three different methods for modeling: outside-in, inside-out, and middle-out
- The value of a canonical data model
- Considerations for security aspects
- Step by step dismantling of the existing architecture
- The focus on sharing of metadata specifications for integration, transformation, and cleansing

6. Self-Service BI and the Logical Data Warehouse

- Why self-service BI can lead to “report chaos”
- Centralizing and reusing metadata specifications with a logical data warehouse
- Upgrading self-service BI into managed self-service BI
- Implementing Gartner’s BI-modal environment

7. Big Data and the Logical Data Warehouse

- New data storage technologies for big data, including Hadoop, MongoDB, Cassandra
- The appearance of the polyglot persistent environment; or each application its own optimal database technology
- Design rules to integrate big data and the data warehouse seamlessly
- Big data is too “big” to copy
- Offloading cold data with a logical data warehouse

8. Physical Data Lakes or Virtual Data Lakes?

- What is a Data Lake?
- Is developing a physical Data Lake realistic when working with Big Data?
- Developing a virtual Data Lake with data virtualization servers
- Can the logical Data Warehouse and the virtual Data Lake be combined?

9. Implementing Operational BI with a Logical Data Warehouse

- Examples of operational reporting and operational analytics
- Extending a logical data warehouse with operational data for real-time analytics
- “Streaming” data in a logical data warehouse
- The coupling of data replication and data virtualization

10. Making Data Vault more Flexibile with a Logical Data Warehouse

- What exactly is Data Vault?
- Using a Logical Data Warehouse to make data in a Data Vault available for reporting and analytics
- The structured SuperNova design technique to develop virtual data marts

- SuperNova turns a Data Vault in a flexible database

11. The Logical Data Warehouse and the Environment

- Design principles to define data quality rules in a logical data warehouse
- How data preparation can be integrated with a logical data warehouse
- Shifting of tasks in the BICC
- Which new development and design skills are important?
- The impact on the entire design and development process

12. Closing Remarks