EDW Positioning
Based on the SAP Real-Time Data Platform

July, 2013
Typical needs in a Enterprise world

Introduction
Different Analytical needs and the consequences in IT architectures

- High variety of information sources
- Extensive information needs

Complex architecture landscapes including different kind of Data Marts, Data Warehouses and EDWs
A data mart is a collection of data coming from subject areas organized for decision support based on the needs of a given domain / group of users.

**Characteristics of Data Marts:**
- Maximum flexibility for LOBs in data modeling and integration of LOB specific data (Agile)
- Supporting local business execution
- Support strategic decision making in LOBs
- Independently of, or sourced from the centralized corporate EDW layers
- Reporting on large volumes of granular, transactional data
- Data latency determined by business requirements – batch to real time
- High Volume Unevenly structured data (Big Data)
- Predictive Analysis
- Event and Streaming sources

Different Needs …
Different Types of Data Marts for Reporting
Examples of Different Types of Data Marts

**Operational Data Marts:**
- Real Time Data
- Reporting on large volumes of granular, transactional data
- Supporting local business execution

**Agile Data Marts:**
- Independently of the highly governed centralized corporate EDW layers
- Maximum flexibility for LOBS in data modeling and integration of LOB specific data
- Support strategic decision making in LOBs
Unlock The Power of Your Data Across The Enterprise
Enterprise Data Warehousing – the single point of truth

EDW * – characteristics:
- Consolidate the data across the enterprise to get a consistent and agreed view on your data
- Combine SAP and other sources together
- Standardized data model on corporate information
- Supporting decision making on all organizational levels

EDWs require a database plus EDW orchestration tools:
- EDW orchestration tools are used to manage, model and run the Enterprise Data Warehouse
- EDW orchestration tools can be shipped as an integrated EDW application
- Custom built EDWs are based on loosely coupled orchestration tools e.g. ETL tools, Data Modeling tools, Monitoring tools etc.

* EDW and DW are often used interchangeably. For the remainder of this presentation only the term EDW will be used.
Example of a Customer Enterprise Data Warehousing landscape

Example for global EDW implementation:
Architecture to provide the base for entire commercial business intelligence solutions:

- Local, regional, global Data Marts in addition to global Data Warehouse necessary to meet business users need
- Data Warehouse as single point of truth with harmonized master data to feed departmental Data Marts and external systems with cleansed data
- Depending on users requirements reporting on departmental Data Mart or Data Warehouse is possible
Different Approaches
What approach is right for you?

Different approaches to implement an analytical data foundation

• Using an Integrated EDW application
  • Model-driven pre-packaged Data Warehouse application as a central component
  • Prebuilt Information – and process content
  • Out of the box tool set for modeling, managing, operating, and governing an EDW and its data marts

• Custom Built Enterprise Data Warehouse
  • Loosely coupled orchestration tools
  • Higher efforts for development and maintenance
  • High flexibility to build custom data models and processes with little enforced governance
  • Open environment to easily import industry models
Different Approaches
What approach is right for you?

• Custom built data marts without EDW integration layer
  • Maximum flexibility for custom data marts specifically built for isolated use cases
  • Easy to build and fast realization time
  • Low level of integration among different data marts

Combination of approaches
• Grow from custom built DMs to an EDW
• Combine custom built DMs or EDW with integrated EDW application
Data Volume and Complexity considerations
Enterprise Data Warehouse, Data Marts

Very large Data Marts
- Internet scale business process (e.g. Ebay, Amazon, …) generating huge amounts of (sensor) data
- Fairly modest challenges regarding semantics, consolidation, harmonization, integration with other data

Data Mart
- Few data sources
- Data mart type of setup or operational (OLTP) analytics
- Moderate number of tables
- Moderate (need for) integrations between data models

Extreme large EDW
- Extremely high number of scenarios and combinations of scenarios
- Extremely high data volumes

EDW
- Mix of scenarios with small and large amounts of data
- Many (100s to 10000s) of data models
- Many (100-1000s) different data sources

Complexity: number of scenarios, data models, sources, …
How does SAP address the needs?
SAP RTDP addresses the needs for a packaged EDW framework with SAP BW, as well as custom built EDW and Data Marts

- SAP NetWeaver BW as strategic offering for packaged EDW frameworks
- SAP HANA as an in-memory foundation for custom built solutions and in-memory database platform for BW
- SAP Sybase IQ as disk-based database for custom built petabyte scale EDW and cost effective near-line storage solution (NLS) for BW on HANA
SAP HANA, BW on HANA, Sybase IQ
Strong individual solutions

**SAP HANA**
- Data Platform for business- or mission-critical enterprise applications
- In-Memory database for transactional and analytic workloads
- More than 1000 customers with standalone, BW or Business Suite scenarios

**SAP NetWeaver BW**
- More than 14000 customers referring to > 17000 productive systems
- Vast majority: Central EDW, harmonizing many source systems
- Embedded into mission critical business processes

**Sybase IQ**
- Recognized leader in EDW market by Gartner and Forrester Group
- Industry leading performance and scale benchmarks with 4500+ installations in 2150 accounts
- Greater than 96% customer satisfaction rates - consistently
EDW with SAP BW on HANA

Data Modeling
- Layered scalable EDW model
- Tight integration with native HANA Data Marts
- In-Memory based OLAP features and Planning engine
- Agile Data Marts with BW Workspaces

Scheduling & Monitoring
- Process Chains enable workload management for data load processes across systems
- Rich monitoring capabilities via Admin Cockpit
- Support of 3rd party tools such as UC4, Tivoli

Supportability, Traceability
- Easy extensible BADI framework for data staging, reporting and authorization
- Access statistics including Identity handling
- Possibility to trace on application logic, database access and effect of authorizations

Reliable Data Acquisition
- Open for SAP – and non SAP systems
- Embedded ETL and Streamlined In-Memory Data Staging processes
- Sophisticated Error Handling
- Incremental Data Provisioning capabilities

Lifecycle Management & Security
- Hot- / Cold data handling for optimized resource usage
- Compliant NLS and Archiving
- Fine-grained security model with mass handling capabilities

Business Content
- Enables fast implementation
- Leverages SAP’s proven knowledge in business applications and industries
- Model-driven approach
EDW with SAP HANA

Data Modeling
- Entity-Relationship models
- Procedural extensions for complex views
- Support for planning operations on views
- Real-time transformation 3NF to analytic view

Scheduling & Monitoring
- DB level monitoring in SAP HANA Studio
- Explain plans for SQL statements
- No scheduling capabilities, requires additional tools such as SAP Business Objects Data Services

Supportability, Traceability
- Technical tracing of database kernel operations
- Alerting for vital database parameters
- Integration with SAP Solution Manager

Reliable Data Acquisition
- Support for virtually any kind of data
- Support for SAP or 3rd party ETL tools
- Real time replication and event stream connection

Lifecycle Management & Security
- Repository for design time artifacts
- Database security with user and role concept
- Privilege concept for design time and runtime
- Content lifecycle management

Business Content
- Business Suite data models available as content packages (SAP HANA Live, RDS)
- Application content available as delivery units
Data Modeling
- Leading open modeling tool: PowerDesigner
- High flexibility in schema definition
- Ad-Hoc SQL, structured and unstructured data support, and integrated full text search
- In-database analytics, MapReduce API, Hadoop and R-Integration

Scheduling & Monitoring
- Web-based SAP Sybase Control Center as DB administration & monitoring tool
- ETL scheduling with DataServices
- Open and certified support for 3rd party monitoring solutions

Supportability, Traceability
- Database auditing: logins, data update timestamps, and administrative actions
- Support of multiple trace and debug levels
- Extensive set of system stored procedures for diagnostics

Reliable Data Acquisition
- Open for SAP – and non SAP systems
- Data Services as powerful, approved ETL tool
- Open and certified support for 3rd party ETL solutions
- Real Time Data Loading via ESP and SRS

Lifecycle Management & Security
- Built-in disk–based Information Lifecycle Management
- Role-based access control, LDAP authentication, strong database, column, and transport layer encryption (FIPS certified)
- New generation petabyte-scale store
- Resilient Multiplex grid architecture

Business Content
- Prebuilt vertical Industry Warehouse Solutions
- Open and certified support for many 3rd party applications
Integration scenarios:

- Tight integration between BW on HANA and native HANA scenarios combines the strengths of both worlds
- Consumption of BW data models in HANA and consumption of HANA models in BW
- Data staging from HANA to BW and from BW to HANA
- IQ supports NLS (multi temperature Data) for BW on HANA and remote data access from HANA for smart query processing
- Lower TCO by reduced data volumes in HANA and shortened backup time frames
- Cost – and Performance optimized EDW
- Mature EDWs will leverage the capabilities of all three components where appropriate
Conclusions – Isolated view

**Customer requires standalone Data Marts**
- HANA provides maximum agility for custom built data marts
- With increasing BI maturity* customers are able to grow from HANA based Data Marts to an EDW on HANA

**Customer requires operational analytics directly on the original transactional data**
- SAP HANA Live for SAP Business Suite (see appendix for more details)

**Customer requires an EDW**
- BW on HANA is SAP’s strategic EDW solution and should be considered as a cornerstone for customers’ EDW strategy

**Customer requires an EDW for very large data volumes**
- Sybase IQ recommended for Petabyte scale

**Customer requires an NLS solution**
- Sybase IQ provides very low TCO for advanced data lifecycle management capabilities, e.g. NLS

* TDWI BI Maturity Model
Conclusion – The HANA EDW
Combining the strengths of the different worlds

- BW on HANA and HANA Agile Data Mart scenarios perfectly address the needs of Central EDWs and Data Marts
- IQ is SAP’s native NLS for BW on HANA and is capable of storing terabytes to petabytes of structured & unstructured data
- Mature customer architectures will leverage the capabilities of all three components where appropriate
Thank you

Contact information:

Product Management: Lothar Henkes, Courtney Claussen, Brian Wood
Solution Management: Daniel Rutschmann
Consulting: Andreas Scholl
COE: Lars Jakob
CSA: Rudolf Hennecke
Further Information

Find additional information

- [www.SAPHANA.com](http://www.SAPHANA.com)

Link to EDW Positioning presentation:

- [https://scn.sap.com/docs/DOC-41326](https://scn.sap.com/docs/DOC-41326)