SAP Data Services in a SAP landscape
SESSION CODE: 0407

Sue Waite, SAP
ASUG Annual Conference – May 2012
Safe Harbor Statement

The information in this presentation is confidential and proprietary to SAP and may not be disclosed without the permission of SAP. This presentation is not subject to your license agreement or any other service or subscription agreement with SAP. SAP has no obligation to pursue any course of business outlined in this document or any related presentation, or to develop or release any functionality mentioned therein. This document, or any related presentation and SAP's strategy and possible future developments, products and or platforms directions and functionality are all subject to change and may be changed by SAP at any time for any reason without notice. The information on this document is not a commitment, promise or legal obligation to deliver any material, code or functionality. This document is provided without a warranty of any kind, either express or implied, including but not limited to, the implied warranties of merchantability, fitness for a particular purpose, or non-infringement. This document is for informational purposes and may not be incorporated into a contract. SAP assumes no responsibility for errors or omissions in this document, except if such damages were caused by SAP intentionally or grossly negligent.

All forward-looking statements are subject to various risks and uncertainties that could cause actual results to differ materially from expectations. Readers are cautioned not to place undue reliance on these forward-looking statements, which speak only as of their dates, and they should not be relied upon in making purchasing decisions.
Agenda

Enterprise Information Management from SAP
Data Services for Analytics
   - Extracting data from SAP Business Suite
   - Load data into HANA
   - Load data into BW, Sybase IQ, ASE
Data Services for Business Processes
   - Data Quality integrated in the Business Suite
   - Data Migrations
   - Information Steward and Master Data Governance

Information based on current versions (SAP Data Services 4.0)
+ additional info on upcoming SAP Data Services 4.1 release
SAP Enterprise Information Management
The Most Complete Enterprise Information Portfolio

- Data Warehouse Management
- Master Data and Reference Data
- Data Asset Maintenance
- Metadata Management

Data Integration and Management:
- Business Warehouse
- Business Warehouse Accelerator
- Rapid Marts
- HANA
- Sybase IQ/ASE
- Data Services (DQ and DI)
- Process Integration
- Data Federator
- Text Analytics
- Information Steward
- Data Migration Services
- Replication Services
- Master Data Management
- Master Data Governance
- Information Steward
- Enterprise Content Management
- Document Access
- Archiving
- Information Lifecycle Management
- Information Steward
- Power Designer

EPM & Analytics:
- Financial Performance Management
- Operational Performance Management
- Industry & Process - Reporting & Analytics

User Interfaces

Composite Applications & Processes

Operational Performance Management

Industry & Process - Reporting & Analytics
SAP Data Services
Your Smart Data Platform

SAP ERP, SAP CRM, SAP Master Data Management (MDM), SAP NetWeaver BW, ...and non-SAP systems too!

Web Services

SAP CRM, SAP ERP, SAP NetWeaver, SAP BusinessOne, SuccessFactors

Sybase, HANA, Oracle, SQL, DB2, HADOOP

PeopleSoft, Oracle Apps, Siebel, SFDC, etc

Files, XML, HTML, Mainframe, Excel, text, etc.

Data Profiling
- Data Services Engine
- Data Cleansing, Enrichment
- Data Validation
- Text Data Processing

4.1

Query, Reporting, Analysis, and Dashboards

Data Migration, Synchronization, Real-Time, ...

HANA, BW, Sybase

Impact Analysis

Shared Metadata

Data Lineage
# SAP Data Services: Enterprise-Wide Data Access

SAP Data Services has a long history for providing broad connectivity to databases, applications, legacy systems, file formats, ...

<table>
<thead>
<tr>
<th>Databases</th>
<th>Applications</th>
<th>Files/Transport</th>
<th>Mainframe (with partner)</th>
<th>Unstructured Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>HANA</td>
<td>SAP Bus Suite</td>
<td>Text delimited</td>
<td>ADABAS</td>
<td>Any text file type</td>
</tr>
<tr>
<td></td>
<td>Extractors</td>
<td>Text fixed width</td>
<td>ISAM</td>
<td>6 languages</td>
</tr>
<tr>
<td></td>
<td>ABAP</td>
<td>EBCDIC</td>
<td>VSAM</td>
<td>Extended to 31 languages</td>
</tr>
<tr>
<td></td>
<td>BAPI</td>
<td>XML</td>
<td>Enscribe</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Idoc</td>
<td>Cobol</td>
<td>IMS/DB</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SAP NW BW</td>
<td>Excel</td>
<td>RMS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>JD Edwards</td>
<td>HTTP</td>
<td>Both direct and changed data</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Oracle Apps</td>
<td>JMS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PeopleSoft</td>
<td>SOAP (Web Services)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Siebel</td>
<td>HADOOP 4.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Salesforce.com</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HP NeoView</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Netezza</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DB2 (LUW, z-series, i-series)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sybase ASE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sybase IQ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SQL Server</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Informix</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MySQL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Teradata</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ODBC</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Reference Data Services PAM for additional details [http://service.sap.com/pam](http://service.sap.com/pam)
SAP Data Services
Integrate and Improve the Quality of All Data

Market-leading, unified solution for enterprise-class data integration and data quality

- Single, easy-to-use user interface to build, test and deploy projects
- Connect, transform and make the information available from virtually any source, structured and unstructured
- Comprehensive data quality solution for cleansing and enriching all types of data
- Scalable solution for high data volumes with high throughput (batch and real-time)

✓ SAP Data Services is a “leader” in the Data Integration and Data Quality Magic Quadrants from Gartner
✓ SAP has 5000+ customers on our Data Integration and Data Quality solutions.
Data Services Only Innovative Solution to Offer Multiple Integrated Data Management Capabilities in One Solution

- One Solution that Provides
  - Data Integration
  - Data Quality
  - Text Data Processing
- One server to execute all capabilities
- One design environment to manage all of the development
- One administration console to monitor all functions

Lower TCO
Development Efficiencies
Reduced Complexities of Environment
### Best for Analytics

**SAP Data Services**

<table>
<thead>
<tr>
<th>Consolidated view of business</th>
<th>Trusted information</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Access and integrate disparate data from virtually any data source (<em>structured and unstructured</em>)</td>
<td>- Cleanse data to improve data quality for more accurate decision making</td>
</tr>
<tr>
<td>- Build a data warehouse or data mart to provide historical trending and analysis</td>
<td>- Understand information context (data lineage) to make more confident decisions</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Integration</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Integration with the SAP BusinessObjects BI platform for lower total cost of ownership (TCO)</td>
<td>- Load SAP and non-SAP data into <strong>SAP HANA</strong> in-memory database, SAP NetWeaver BW, and SYBASE</td>
</tr>
</tbody>
</table>
Native Access to SAP Business Suite data with SAP Data Services 4.0

Benefits of using Business Content extractors (compared to ABAP dataflows)

- Same connectivity to ERP as what BW has provided for years
- Job setup is much easier to deploy and maintain (No ABAP programs generated)
- Delta loads are fully supported (changed data capture)

Pre-requisites: please see note 1522554 - NetWeaver Support Package requirement for Data Services SAP Extractor support (technical requirements and officially released/supported extractors)

4.1 Multi-threaded extractor reading enables scale out
Native Access to SAP Business Suite data with SAP Data Services 4.0

Browse & import extractors in an SAP applications datastore

Use extractors in a data flow like a regular source
Extracting data from SAP ERP/CRM with ABAP dataflows

Key features:

- Metadata browsing
- Support all SAP table types (including cluster, pool and Z-tables)
- Support for hierarchies
- Design ABAP dataflows in a graphical UI – no coding required
- “Generate & Execute” (for Development) or “Execute Pre-loaded” auto-generated ABAP programs
- Support for all SAP ERP versions (from R/3 3.1i until ECC 6)

Data streaming for ABAP data flows to read data from ABAP tables (no staging in files anymore)
- Enhanced function modules: now in their own name space /BODS/ instead of custom Z namespace + additional authorization objects for more granular security.
Loading BW with Data Services

First certified for loading SAP NetWeaver BW in 2002
Browse BW metadata from Data Services Designer
  - Select InfoSources / DataSources to use as a target

Data is loaded into the PSA in the Data Acquisition layer
Data Services load jobs can be initiated from within BW workbench
  - Data Services includes an RFC server integrated in the Management Console to execute BW load requests initiated through an InfoPackage
Create On-the-Fly Jobs from SAP NetWeaver BW 7.3 with Data Services 4.0 to Load from Any Source

Auto-generate extraction jobs for SAP Data Services from within SAP NetWeaver BW to load data in from any source

- Integrated connectivity for SAP Data Services 4.0 in SAP NetWeaver BW 7.3
  - From within the workbench in SAP NetWeaver BW, browse external metadata through connectivity to SAP Data Services
- Native loading in data stores of SAP NetWeaver BW 7 through new staging BAPIs
Extracting from BW with Data Services

BW Open Hub can get data from InfoCubes, DataStores, InfoObjects and InfoSets.

- Open Hub destinations are set up in the BW workbench together with the DTP (Data Transfer Process) to load the data into the Open Hub Destination

Browse BW metadata from Data Services Designer

- Select Open Hub Destinations (with associated process chains) defined in the BW workbench.

Open Hub destinations are controlled from within Data Services through process chains for loading and purging the data.
SAP Data Services’ benefits for customers with SAP ERP & BW

Consolidate all data in the Enterprise Datawarehouse (NetWeaver BW)
- Easy access (no coding) to all data sources through native interfaces

Fix data quality issues
- Harmonize non-SAP data to SAP standards, enrich and cleanse data, remove duplicates

Standardize on one platform for all data extraction, validation and cleansing rules
- SAP and non-SAP data can be combined, cleansed and validated before loading into BW.

Need quick time to value – DWH can take months, but LOB needs data next week
- Over time the datamarts can become sources for the BW EDW

Push data from BW to downstream applications or datamarts
- The EDW becomes is source itself for downstream applications
SAP Rapid Marts
Jump-start analytics implementations

Deliver rapid insight
Increase reporting flexibility
Reduce costs

SAP ERP Sources
- Financial
- Operational
- Manufacturing
- Human Capital Mgmt

Start here without Rapid Mart
Engage Specialized SAP Skills
Extension and Customization
Start here with Rapid Mart

© 2012 SAP AG. All rights reserved.
Data Marts

Set of target RDBMS objects and schemas

Based on best practices for dimensional data modeling

Support for SAP HANA, Microsoft SQL Server, and Oracle as target database

New version 4.0 released April 2012
Available Rapid Marts for SAP solutions:

- Financials Suite
  - General Ledger
  - Accounts Receivable
  - Accounts Payable
  - Cost Center

- Human Capital Management Suite
  - HR

- Operational Suite
  - Sales
  - Inventory
  - Purchasing

- Manufacturing Suite
  - Project Systems
  - Production Planning
  - Plant Maintenance

Please Note: Rapid Marts are jump-start data mart offerings useful for rolling out quick, subject-area specific BI solutions for departmental needs. Rapid Marts’ aim is to simplify and shorten data mart deployment cycles. They are not replacements for an Enterprise Data Warehouse like BW.
SAP Data Services 4.0 for SAP HANA 1.0

**High Performance**
- Highly-scalable engine to move large volumes of data into HANA
- Integrated with HANA’s bulk-loader
- Pushdown to HANA to leverage in-memory where possible (EL-T)

**Wide Connectivity**
- Data and Metadata connectivity to all major enterprise data sources (Applications, databases, files)
- Support for unstructured data with text data processing

**Powerful Transformations**
- Built-in transformations for data quality (*)
- Support for non-relational data formats including text and XML

* = Requires additional license
NEW in Data Services 4.1

Data Services Workbench for data loading into SAP HANA

First release for a new, Eclipse-based user interface for Data Services.

Addressed scenario: replicate data and metadata (table definitions) from 3rd party databases into HANA via a simple wizard.

In three simple steps you can replicate millions of records and hundreds of tables from any source to your HANA target:

- Set up connections to source and target system
- Select the required tables from the source (or simply select all tables)
- Generate and execute the job to move all data with one mouse click

Once the wizard exits, additional features can be added: filters, complex data mappings (expressions and functions), load options, delta loading strategy, …

For more detail:

**ASUG Session 0410**: How to do load data from SAP and non-SAP data sources into SAP HANA in batch mode (ETL) using SAP Data Services
Design your Replication Job

NEW in Data Services 4.1

Step 1: Select source and target connections (database or SAP application)

Step 2: Select tables for replication, target tables will be generated automatically

Step 3: Deploy your Data Services job and execute

This presentation and SAP’s strategy and possible future developments are subject to change and may be changed by SAP at any time for any reason without notice. This document is provided without a warranty of any kind, either express or implied, including but not limited to, the implied warranties of merchantability, fitness for a particular purpose, or non-infringement.
Best for Business Processes
SAP Data Services

<table>
<thead>
<tr>
<th>Data migration</th>
<th>Data quality management</th>
<th>Master data management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Move legacy data from any application into a new deployment of SAP Business Suite</td>
<td>Cleanse data to improve data quality for more effective business processes</td>
<td>Move disparate data into a consolidated master data location</td>
</tr>
<tr>
<td>Ensure that data quality and format meet the needs of the target application</td>
<td>Improve data quality at point of entry or over the life of the SAP ERP, SAP CRM, or MDG application</td>
<td>Cleanse data in preparation for reconciliation to define a single version of data entities (customer, product, supplier,…)</td>
</tr>
</tbody>
</table>
Data migration is a critical element in the success of the ERP or CRM implementation.

Ensure accountability and a robust project approach to data migration from a single strategic vendor.

Enable data governance best practices that live on after the project.

Industry-leading SAP Data Services platform benefits data migration and delivers many additional benefits in information management.

Pre-built framework and tools reduce the need for custom scripts and deliver data insights more quickly and consistently.

Engage business users early with better access to data quality and metrics.

http://service.sap.com/bp-datamigration
Data validation from source data to target data structures

- Data validation out-of-the-box content included with SAP Best Practices includes specific business rules for validating source data against SAP target system.

- Content includes extraction of check tables from the SAP target system for the value mapping.
Efficient loading of cleansed and validated source data to target system

Standard NetWeaver loading mechanisms are used to load data into SAP target systems.
SAP Data Quality Management for SAP Solutions

Prepackaged native integration of data quality best practices within SAP and 3rd party business applications

- Enforces data discipline directly within CRM, ERP, MDG systems for customer, vendor, and business partner data
- Virtually undetectable presence within the application – native look and feel
- Extensive coverage for global postal validation processes - over 240+ countries
- Advanced record matching strategies for duplicate prevention
Automatically Completes and Corrects Entries, then Presents to User for Approval

Incomplete address on new account...

Automatically made accurate!
Previously Unknown Entries are Identified, *Before* another Duplicate can be Created

List of pre-existing entries is presented
Complementary Data Governance Platform

- Powerful Validation Rules enhance data cleansing and remediation within MDG
  - Easy impact assessment of changes in existing rules
  - Predictive assessment for new validation rules to be introduced
  - Implantation of validations which cannot be implemented in MDG
- Data Insight to new additional data sources to be integrated into MDG (“fit for use”)
- Metadata lineage, impact provide context to core MDG Process and data
- Continuous monitoring and integration with MDG processes to remediate data issues
SAP MDG and SAP Information Steward Process Flow

Information Steward

*Technical Data Analyst* reviews DQ Dashboard to see data quality level on data table perspective.
*Technical Data Analyst* reviews the DQ Scorecard to see data quality level on data domain perspective.

MDG

*Data Steward* reviews the same DQ Scorecard directly in Business Client.
*Data Steward* drills down to failed records.
*Data Steward* creates Change Request for the Failed Records.

MDG

*Master Data Steward* reviews his work list with new Change Request.
*Master Data Steward* cleanses the wrong data manually.
*Master Data Steward* closes the Change request.

MDG

*Data Steward* reviews the DQ Scorecard within the Business client and sees that the overall DQ Score has changed as he has manually cleansed failed records.
*Data Steward* drills down to failed records and sees that the number of records is lower.

 nightly data quality assessment
Questions?

Useful links:
SAP BusinessObjects Data Services wiki on SCN: http://wiki.sdn.sap.com/wiki/display/EIM/Data+Services
Sue Waite
EIM Solutions Management
SAP Americas
La Crosse, WI 54601 USA
Email: sue.waite@sap.com
Office: 1.608.793.7313