

Integrating SAP HANA – Big Data in Memory Technology to IS Education

Ming Wang
California State University, Los Angeles
ming.wang@calstatela.edu

New Trends in DBMS

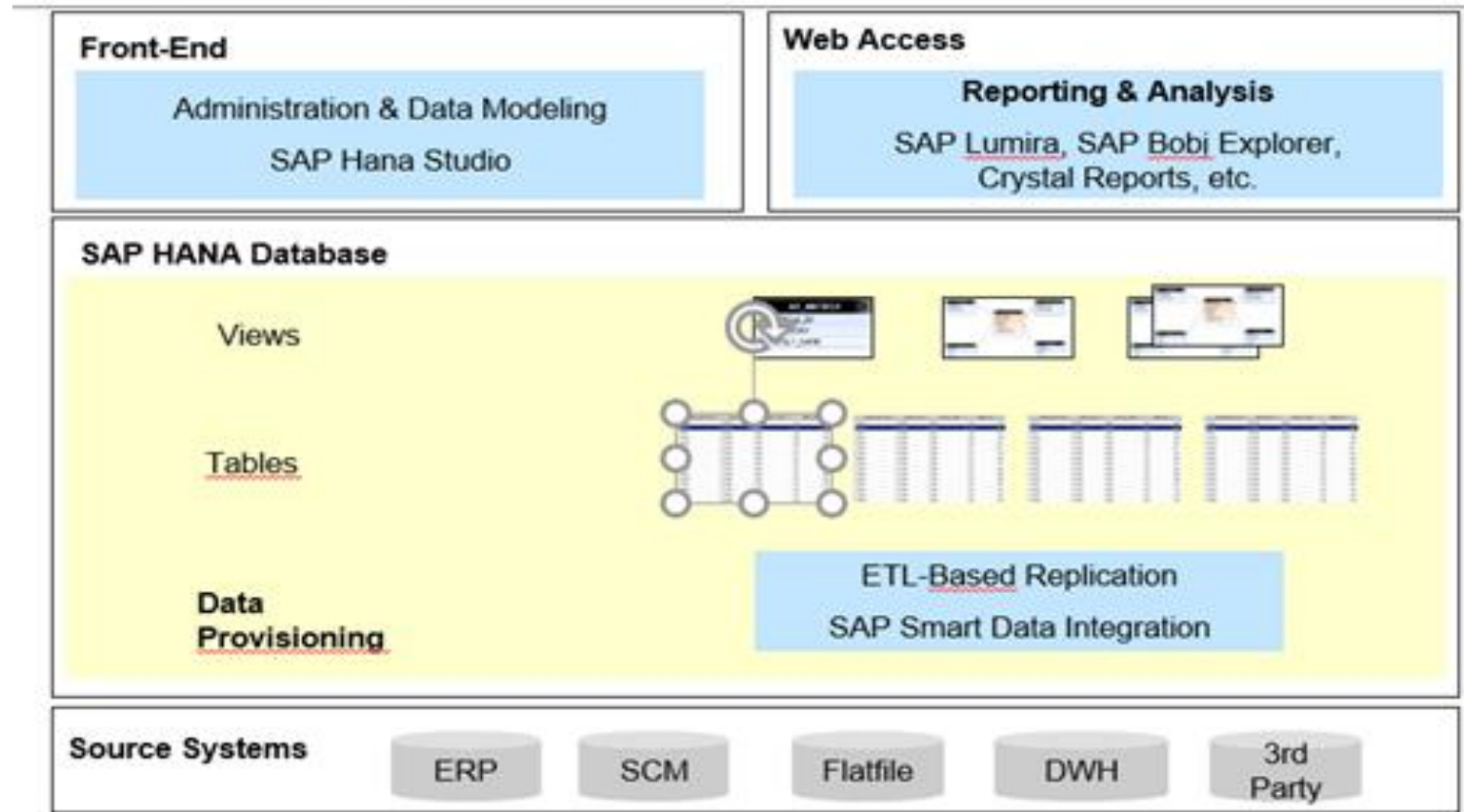
- The databases and data warehouses are two fundamental data sources for business intelligence and data analytics. Recent development of SAP HANA technology has combined the two major data sources into one data in memory platform.
- Some database vendors have recently updated their DBMSs with data-in-memory and columnar data storage features in their new versions. For example, the newer version of Oracle DBMS has made evolutionary changes by allowing data stored in the columnar table and running them in memory.

Features of SAP HANA Platform

- In memory appliance: Allow larger data volumes to be processed in real time
- Columnar data storage: optimized for analytics and data retrieval and high level of compression leading to lesser storage
- Parallel processing: Multi-core-architecture and columnar storage provides parallel processing resulting in fast are response times
- Cloud and On-premise support
- Multi Tenants database support to run more than one application on a single HANA database

Overview of SAP HANA

SAP HANA OVERVIEW



Columnar Data Storage in SAP HANA

SAP HANA 's underlying columnar in-memory relational database is exceptionally fast because the index on the columns

Row-store table

Row 1	USA	3000	DXTR1100	5
Row 2	USA	4000	DXTR1100	21
Row 3	DE	23000	DXTR3100	12
Row 4	DE	17000	DXTR3100	34

Columnar table

Column1	Column2	Column3	Column4
USA	3000	DXTR1100	5
USA	4000	DXTR1100	21
Germany	23000	DXTR3100	12
Germany	17000	DXTR3100	34

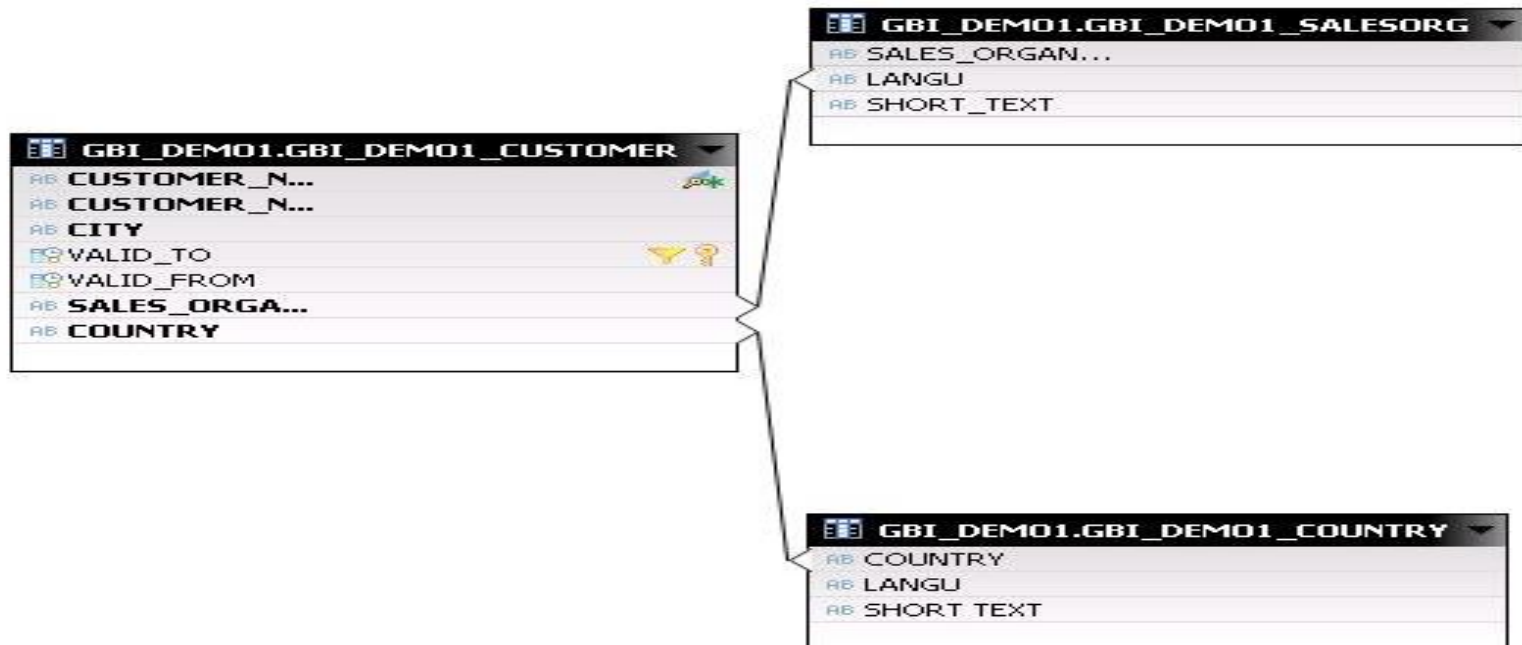
SAP HANA Data Modeling

SAP HANA data modeling for analytic applications uses three information views:

- Attribute views
- Analytic views
- Calculation views

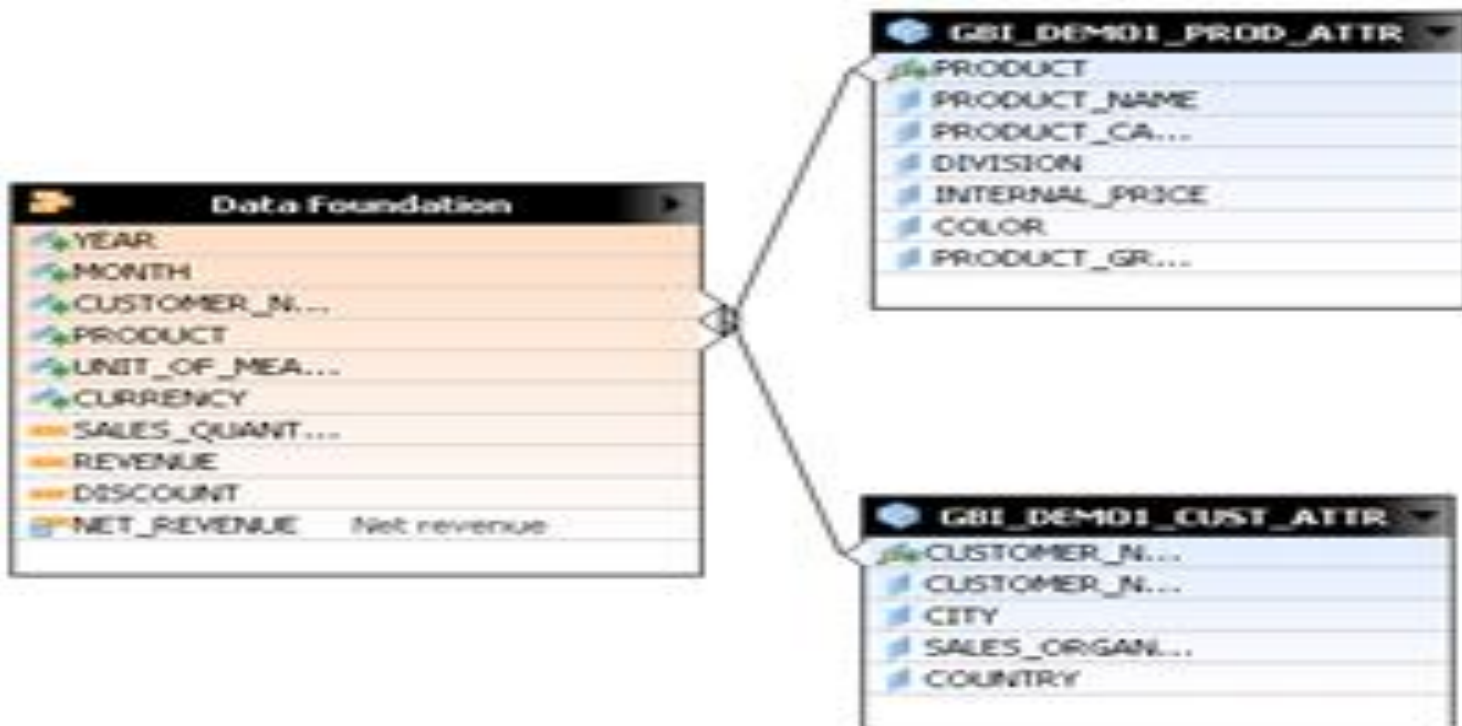
Attribute Views

represent master data (attributes, texts, hierarchies) and provide reusable dimensions for analytic and calculation views, Analytic views



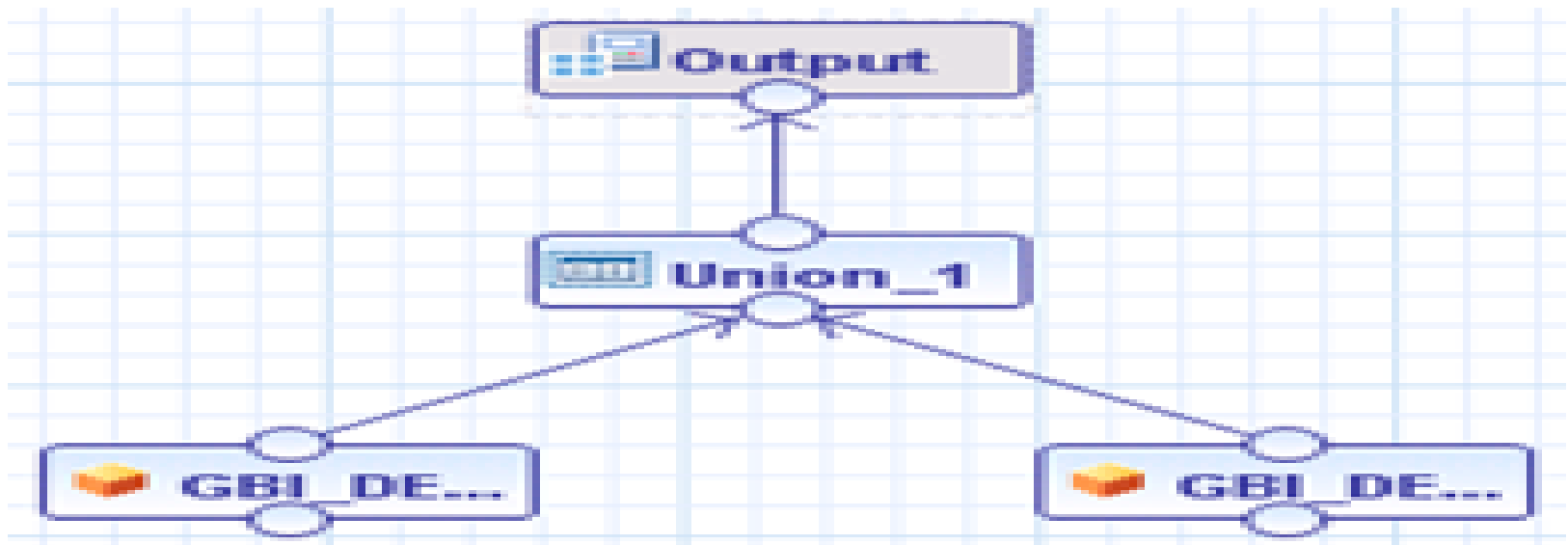
Analytic views

- Analytic views join facts with relevant attribute dimensions



Calculation Views

Calculation views address more complex requirements than analytic views and can include both tables and views.



SQL via OData

-- Display sales

```
SELECT MATERIAL_DESCRIPTION, SALES_ORGNIZATION,  
NET_PRICE, NET_VALUE, QUANTITY, COST
```

```
FROM
```

```
"ERPsim.OData.erp.sim.erp.sim"."entities"."Sales";
```

-- Display plant inventory

```
SELECT PLANT
```

```
FROM
```

```
"ERPsim.OData.erp.sim.erp.sim"."entities"."Inventory";
```

SAP HANA's Pros and Cons

SAP HANA's Advantages

- Better I/O bandwidth utilization
- Higher cache efficiency
- Faster data aggregation
- High compression rates
- Columnar-based parallel processing

SAP HANA's Disadvantages

- Load times can be slow
- Less efficient for transactional processes
- Possibly slower relational interface

Summary

The author currently teaches the Database and Data Warehouse Systems course in the MSIS program. Oracle DBMS is utilized for relational databases and object-relational databases. SAP HANA is utilized in teaching data warehouse. Students modeling data and analyze data on HANA on ERPsim distribution Game. It is important to bring big data-in-memory database technology in the newer edition of database and data warehouse textbooks.

FEATURE STUDY

The author will conduct a comparative study on SAP HANA and newly released version of Oracle DBMS and investigate the revolutionary and evolutionary changes in data modeling, data stores and analytical performance in these two products. Meanwhile the proposer will also develop guidelines for IS educators to teach this dynamically changed technology. The significant findings of such research will benefit both academia researcher, industry practitioners and IS educators.