Enterprise MDM: Complementing & Extending the Active Data Warehouse

Mark Shainman
Global Program Director, Teradata MDM
Agenda

- MDM and its Importance
- MDM, The Enterprise Data Warehouse and Data Mart Consolidation.
- Teradata’s MDM Solution
- Case Study
MDM – What Is It?

Master Data Management refers to the methods by which clean, accurate, and consistent master data (e.g., customers, products) are managed, referenced, and synchronized across the enterprise and made available to users as required.

- It is used to classify and define master data, through the use of a centralized integration manager, sometimes referred to as a hub.
- It leverages policies and procedures for access, update, and overall management of this central resource and its coordination with other participating systems across the enterprise.
- Areas such as customer data integration (CDI) – management of customer reference data and product information management (PIM) – management of product and supplier reference data, are domain-specific subsets of MDM.
# Examples of Master Data

<table>
<thead>
<tr>
<th>Organization</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sales and Marketing</strong></td>
<td>Customer Record</td>
</tr>
<tr>
<td><strong>Service</strong></td>
<td>Warranty Codes, Service Codes</td>
</tr>
<tr>
<td><strong>Product Development, Sourcing</strong></td>
<td>Product, Supplier, Item, BoM, Product hierarchies</td>
</tr>
<tr>
<td><strong>Finance</strong></td>
<td>Dept. Codes, Cost Centers, Financial hierarchies, Chart of accounts</td>
</tr>
<tr>
<td><strong>HR</strong></td>
<td>Employee, Org hierarchies</td>
</tr>
<tr>
<td><strong>Misc.</strong></td>
<td>Code masters for common entities like currency, UoM etc.</td>
</tr>
</tbody>
</table>
Master Data Examples and Importance

**Master Data**, also called “reference data,” classifies business data (customers, vendors, product offerings, etc.) consistently across systems.

**MDM Risk Area:**
**Inconsistent customer identifiers**
- ✘ What’s the full picture of each customer’s ‘value’?
- ✘ Does this include travel w/subsidiaries?

**MDM Risk Area:**
**Inconsistent product identification**
- ✘ Which seats are preferred by frequent flyers and should be reserved for them?

**MDM Risk Area:**
**Other categorization**
- ✘ How has travel in/out of certain airports changed over time?

**Master Data Management** (MDM) refers to methods by which clean & consistent reference data is managed, referenced, and synchronized across the enterprise.
Master Data Management (MDM) in context off all data within the EDW

Record of business events
- Orders
- Forecasts
- Inventory
- Sales History...

They occur at a discrete point in time

*Once recorded, transaction data, never changes and requires no maintenance.*

Transaction Data

Master Data
- Master Reference Data
- Relationship Data

Defines business objects
- Product list
- Customer list
- Locations...

*Once recorded, Master data, constantly changes and requires on-going maintenance.*

Information about data
- Data type
- Definition
- Constraints
- Relationships
- Location...

Relationships among master data entities
- Product - Location
- Customer hierarchy
- Product hierarchy...
What is needed for a successful Enterprise MDM solution

A good process

A good framework and solution components

A enterprise data repository & data model
Thus...MDM Enables more **Accurate** Active Enterprise Intelligence

**STRATEGIC INTELLIGENCE**

- Which vendors contribute the least to profit?
- Am I sure that I am seeing all my vendor division’s spend accurately
- Am I sure that I am rolling up my Product Revenues correctly
- Target customers to acquire? Retain? At what cost?

**OPERATIONAL INTELLIGENCE**

- Expedite overnight for 10:00 a.m. delivery?
- Is the new promo driving sales this morning?
- Do I give this customer a discount?
Agenda

• MDM and its Importance

• MDM, The Enterprise Data Warehouse and Data Mart Consolidation.

• Teradata’s MDM Solution

• Case Study
MDM and the Data Warehouse

• Data Warehouses are moving to be more active and operational to the needs of companies

• There are common elements in terms of leverage
  > Database technology
  > Data models
  > Quality
  > Governance
  > Integrations
  > Others..
  > And of course THE DATA!
MDM Has Been Happening in Bits and Pieces Within the DW for Years

Where did you first deploy an MDM solution?

Data warehousing or business intelligence | 61%
CrM or CDI | 13%
Supply chain or product-related task | 13%
Regulatory compliance | 4%
Other | 9%

Figure 14. Based on 148 respondents.

...Hence, data warehousing and BI professionals tend to be combat-hardened veterans of master data management – though few of them use the term. Most see MDM and MDM-like practices as part and parcel of data warehousing’s individual layers, namely data integration, metadata management, data modeling, and report design. Whatever you call it, managing master data across the many layers of the technology stack is required for a deep and rich data warehouse.

And Data Warehouses Overwhelmingly Need This Data Repeatedly...
The Master Data Problem Within the Warehouse

> Organic growth and acquisitions produced arrays of systems with overlapping data
  - Customers, suppliers
  - Offerings, locations, personnel
> Inconsistent means of identification and classification (reference or master data) make achieving a single view of the business difficult...
> ...even when all data is loaded into a data warehouse
> Root Cause: decentralized data governance

“Management of master data is the most frequent source of inconsistency and organizational conflict in data warehousing.”

– Forrester Research, J. Paul Kirby, Jan. 2007
LDMs Enable the “Enterprise” in EDW

- Integrated Logical Data Model ... Store Once / Use Many
  > What data do I have? What data do I need?
  > Where is the data? Where did it come from?
  > Is it reliable? How fresh is it? What is its level of quality?

LDM’s are the EDW roadmap, but the data still has to be physically integrated and maintained!
MDM is necessary to have a truly integrated DW!

- The goal of a centralized EDW is to have truly integrated view of the business and the data.
- A logical data model is a necessary infrastructure piece to map the business and business needs to the data.
- MDM is necessary infrastructure piece to truly physically integrate the data!
- Different operational systems and divisions views of customers, products, and suppliers must be rectified before data can be truly integrated and true “single version of the truth” or “single view of the business” can be achieved.
- MDM not only allows for the true integration of data within the warehouse, but improves overall data quality.
- One key thing is that unlike other traditional integration and DQ methods, MDM enables the continued maintenance of the data not by just IT, but by the domain experts of the data, business!
Many existing Teradata customers do some type of MDM in the warehouse today – maintaining a subset of master data for analytical purposes.

Problem is, the process is fragmented, complex, incomplete and un-maintainable!
MDM Allows Customers to Manage Existing Data Warehouse MDM Processes Holistically!

- Work flow driven, repeatable, processes.
- Agile and flexible
- Business supported

Teradata MDM

Business Rules and Workflows, Process and Data Management

Teradata Warehouse

ETL Tools

Third Party Data Quality

Improving data quality and enabling data integration!
MDM Enables on-going Management of Master Data.

Unlike traditional attempts to rectify disparate master data by hard coding cross referencing in ETL jobs or custom code, Teradata MDM allows for the continued easy maintenance and upkeep of the data!
MDM With the EDW Enables Continued Updated, Accurate, and INTEGRATED View of Enterprise Data.
Teradata MDM/EDW Architectural Approach

- Fundamental architectural difference is the separation of the management of the master data from the transactional data!
- Separate ETL streams to allow for the master data to be loaded and managed differently than the transactions
- With the master data managed separately, the transactional data can be viewed in relation to the master data at any point in time without ever having to change the data! - viewed in past, and present hierarchies, multiple hierarchies, previous names, and addresses etc.
MDM and Data Mart Consolidation

Data Mart Consolidation is the process of centralizing data from multiple, disparate data marts into a single data warehouse that is accessible by all information seekers within the organization.

- **Cost saving initiative**
  > Consolidating marts into centralized data warehouse reduces the cost of hardware, software, tools, processes and personnel.

- **Information improvement initiative**
  > Consolidating marts will provide a single version of the truth, providing more accurate information and better decisions.

- **Cost avoidance initiative**
  > As demands arise for more information and applications, a centralized data warehouse will provide solutions faster and more cost effectively than building an additional data mart.

DMC without some type of MDM can create “good information” at the business unit level, but might create “misinformation” at the enterprise level.
Data Mart Consolidation Without MDM

In a data mart consolidation master data and transactional data is moved to the EDW, but still exists in silo’s.
Data Mart Consolidation With MDM

Customer

Order Entry System 1
Order Entry System 2
Order Entry System 3

Teradata MDM

Master Data (Customer, CST, Cust)

Customer $500
CST $200
Cust $100

Transactional Data

Data Mart 1

Data Mart 2

Data Mart 3

Customer Prime
$800
Teradata Warehouse

Customer $500
CST $200
Cust $100
Master Data Management (MDM) with Data Mart Consolidation

Adding Business Value and Reducing TCO

- DMC without MDM = Data mart data within a EDW platform
- DMC with MDM allows customers to achieve the business value of truly consolidated data, and a single copy of customer, product, supplier etc.
- MDM becomes a key component of achieving the vision of a single view of the business.
- MDM enables companies to more rapidly implement analytical applications using enterprise MDM data = faster time to market and increased ROI
- DMC is an ideal time to implement Teradata MDM, which can create and maintain application-specific data; consolidate data from disparate systems; cleanse, transform and validate data; and filter unwanted data
The true business value of DMC is not just to re-platform the data but to integrate the data—what MDM does!!
Teradata MDM’s Enterprise Value Approach:
A single enterprise platform can manage all master data!

By merging the MDM data from multiple sources, and multiple subject areas onto a single platform, you can retain the functional capabilities of the originals while broadening the business value to other areas of the organization. A single infrastructure investment can manage all domains!!

**FROM current multi-hub local based approach**

- App 1: Siebel CRM
- App 2: SCM
- App n: SAP FI/CO
- Customer
- Product
- Finance
- Reporting
- Analy. Apps
- Mining

**TO a centralized integrated Approach**

- App 1: Siebel CRM
- App 2: SCM
- App n: SAP FI/CO
- MDM
- Reporting
- Analy. Apps
- Mining
Value of a Common MDM and EDW Platform

- **One Enterprise-wide Data Model** – single view of reference data for operational and analytical purposes
- **Cost Savings** – limits separate MDM infrastructure (platforms, integration components, etc.). Leverages existing EDW investments, including those for high availability and scalable performance.
- **Prevents MDM Platform/Data Mart Proliferation** – e.g. customer master data repository, financial master data repository... and overhead to support.
- **Skill Sets** – existing data management and integration skills exist within the EDW environment. These can jump-start your MDM initiative.
- **Master Data Analysis** – changes or trends in the master data itself can be analyzed without moving it
- **Managed Growth** – start small with analytical MDM in limited subject areas, without changing operational systems. Add more subject areas and operational MDM after value has been proven.
Agenda

• What is Master Data, Master Data vs. Metadata, MDM and its Importance

• MDM, The Enterprise Data Warehouse and Data Mart Consolidation.

• Teradata’s MDM Solution

• Case Study
Teradata MDM Solution Today

- **Data Management Platform with Core MDM Services**
  > Hierarchy management, cross-referencing, business rules, data process definitions, metadata management, etc.

- **Teradata Industry Logical Data Models**
  > Financial, Retail, Manufacturing, Telecom, Transportation, etc.

- **Teradata Warehouse**
  > Proven scalable data management technologies for active enterprise intelligence and mixed workload environments

- **Teradata Consulting and Implementation Services**
  > Consulting services providing data strategy and architecture
  > Implementation services supporting assessments, data quality, cleansing, etc.
MDM Solution

MDM Design Studio

Meta-modeling Environment
Data Models, Workflows, Rules, Validations and UI’s

Design and Build Domain Specific IP

Customer Data Integration

Reference Data for Risk

Product Information Mgt

Applied MDM Solutions

Editing, Visualization, Search etc.

UI Layer

Application Layer

Database Layer

Overall Workflows and Policies

Security and Governance
 Mapping/X-Referencing
 Code Mastering

Validation
 Publishing Service
 Hierarchy Mgt

Meta Data Mgt
 Events/Alerts
 Auditing

MDM Platform

Teradata Database

MDM

Publish/Outbound
 Master
 Net Change
 Errors
 Inbound Staging

Teradata Database

EDW

Integration EAI, ETL, Web Services

External Services
- Data Quality (Trillium, Firstlogic…)
- Metadata Services
- D&B Information
- Trading Partners
- Data Analytics
- ...

Enterprise Systems

Operational Systems

Analytical Systems

Teradata

Raising Intelligence
Teradata MDM Design Studio
Value of a Workflow Based Approach to MDM

- In MDM every customer has requirements that are unique
- Configurable workflows allow the solution to map to customer’s best practices vs. being a constraint
- This approach also allows us to support any data asset and model thus supporting an Enterprise MDM strategy for our customers
Teradata Customer Data Integration (CDI)

Sources of Customer Data:
- ERP
- Legacy
- Other

Data Acquisition

Data Cleansing & Validation

Address Cleansing
- missing address line1
- missing city
- missing state
- missing zip code

Address Standardization
- Road, Rd, Rd.
- St, Street, St.
- Blvd, Boulevard, Blvd.
- 2 letter State Code (Texas to TX)

Other Data Validations
- "Not Null" checks
- "Lookup" checks
- Auto Updates
- Filter data
- Etc.

Matching / Duplicate Recognition

Matching / Duplicate Recognition

Automated Merging / De-Dup

Automated, Rules Based Merging or De-Duplication
- Cross-Referencing

Interactive Data Steward Workflows

Data Steward Workflows for Merging
- Cross-Referencing
- Search
- Audit Trail

Customer Enrichment Workflows

Customer Data Enrichment
- New Customer Introduction

Data Acquisition

Data Cleansing & Validation

Address Cleansing
- missing address line1
- missing city
- missing state
- missing zip code

Address Standardization
- Road, Rd, Rd.
- St, Street, St.
- Blvd, Boulevard, Blvd.
- 2 letter State Code (Texas to TX)

Other Data Validations
- "Not Null" checks
- "Lookup" checks
- Auto Updates
- Filter data
- Etc.

Matching / Duplicate Recognition

Matching / Duplicate Recognition

Automated Merging / De-Dup

Automated, Rules Based Merging or De-Duplication
- Cross-Referencing

Interactive Data Steward Workflows

Data Steward Workflows for Merging
- Cross-Referencing
- Search
- Audit Trail

Customer Enrichment Workflows

Customer Data Enrichment
- New Customer Introduction

Sources of Customer Data:
- ERP
- Legacy
- Other
Teradata Product Information Management (PIM)

- New Product Introduction
- New Packaging Variant Introduction
- Product Versioning and Authorization
- Enterprise Hierarchy Management
- Product Hierarchy Management
- Promotions Management
- Maintaining Product Packaging
- Update/Cancel/Discontinue/Delete/Correct Product
- Distribute Product by Trading Partner or Target Market
- Trading Partner Specific Price Maintenance
- Product Attribute Management
- GDSN Connectivity Adaptors
Initially MDM Tables can be dedicated tables in their own database, but in time they may become simply views into Tables in the larger EDW.
Agenda

- What is Master Data, Master Data vs. Metadata, MDM and its Importance
- MDM, The Enterprise Data Warehouse and Data Mart Consolidation.
- Teradata’s MDM Solution
- Case Study
MDM Case Study—Corporate Overview & Industry Challenges

- Among the largest bank holding companies in the US. Also a “top 10” US full-service brokerage firm.
- Customer base: over 8,000,000 households and businesses
- Over 75,000 employees
- Avg. 4-5 acquisitions per year

**Industry Challenges**

- Increased risk of fraud
- Aggressive regulatory environment
- Margin and cost pressure
- Increased competition from traditional and nontraditional sources
- Demanding clients
- Industry consolidation
- Globalization
- Rapidly scaling technology capabilities
- Increased consumer credit risk
- Product commoditization

Source: The Tower Group, Jan. 2007
MDM Case Study - Management Objective

- Achieve Basel II compliance
  - Requires holistic view of each customer with respect to credit exposure (credit card + mortgage + other loans + ..)
  - Can no longer calculate exposure & corresponding capitalization needs from a “product silo” perspective
  - Must not only roll up to single customer view but also unite related product types
MDM Case Study-Project Overview

Two related projects, both related to the credit risk-focused data warehouse and using Teradata MDM

<table>
<thead>
<tr>
<th>Reference Data Management</th>
<th>General Ledger Reconciliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Maps “dirty” source system reference data to “clean” master data values in Teradata</td>
<td>• Enables monthly reconciliation of account balances in the risk DW to general ledger (G/L) balances. Must match!</td>
</tr>
<tr>
<td>• Provides workflow and UIs to allow risk dept. (business) “stewards” to do the work themselves</td>
<td>• Occurs during the monthly G/L closing period</td>
</tr>
<tr>
<td>• Enables reporting &amp; analysis of data trends and trends in data exceptions</td>
<td>• Supports regulatory reporting requirements</td>
</tr>
<tr>
<td></td>
<td>• Identifies variances and enables research &amp; explanation, including reporting on trends compared to earlier months</td>
</tr>
<tr>
<td></td>
<td>• Provides adjustment entry and control of approvals and of the overall process</td>
</tr>
</tbody>
</table>
MDM Case Study - Previous Environment

- No effective means of dealing with messy source system data
- Explanation & reconciliation of variances more difficult and manual
- Unable to quickly and fully address audit-related needs
MDM Case Study - Groups Involved

- Risk Management – sponsors of the projects (business)
  > Includes “reference data stewards” fully responsible for different types of master data and empowered to make all changes to it, including mapping newly received values to approved values used for reporting

- IT: implementation & support

- Master data workflow includes approval processes. Business (risk) users also review and provide final approvals before updates are fed into ETL processes for use during the next DW load.
MDM Case Study—Reference Data Management

**Source Data**

- ETL processing uses master data from Teradata MDM when loading data to EDW
- Data stewards notified of exceptions. They then address based on governance rules
- Approval process for new reference & cross-reference data
- Source system “dirty” reference data is not updated, just mapped to clean master data
- Includes versioning of master data, a Basel requirement

**Data Integration**

- Cleanse incoming data
- ETL the data into Teradata data warehouse
- Source system ref. data mapped to master data maintained via Teradata MDM
- Unidentifiable values mapped to temporary values for review

**Teradata Warehouse**

- Product & customer reference data
- Cross-references between source and target values
- Data enrichment
- Holds MDM data versioning (Basel requirement)
- Data Governance profiles

**MDM Users**

- Add reference & cross-reference data
- Approve master data
- Establish govern. roles by attribute & source

**Reporting**

- Exposure/Sub Exposure reporting
- Data trends
- Data Exception Trends
- SAS computational engine
MDM Case Study - G/L Reconciliation Solution

- ETL processing uses master data from Teradata MDM when loading data to EDW
- ETL generates autoexplanations for variances below established thresholds
- GL Data processing cycle occurs monthly, driven by GL reporting cycle
- Source system “dirty” reference data is not updated, just mapped to clean master data

G/L Source Data

Data Integration
- Cleanse incoming data
- ETL the G/L data into Teradata data warehouse

Teradata Warehouse

Teradata MDM
- G/L Reference data
- Threshold values
- Variances
- GL Data Governance profiles
- Audit trail of adjustments

MDM Users
- Threshold maintenance
- Adjust/explain variances
- Certify G/L adjustments

Reporting
- Variances
- Account Details
- Unclassified Adjustments
- Adjustment Status
- Adjustment Trends

Data Integration
- Cleanse incoming data
- ETL the G/L data into Teradata data warehouse
Business Impact

• Able to conduct credit risk analytics across all imported data due to mapping of source system reference data to “clean” values in the risk DW

• Able to comply with statutory/regulatory requirements, including those specified in Basel II

• Can enter an audit with higher confidence and ability to explain the data, including deep history

• Decisions and actions on issues such as cross-referencing made more efficiently in one place, as the risk management team is directly involved in day-to-day management of reference data

• Infrastructure & processes in place to ensure extensibility of the data model as needed in the future
Questions?