Master Data Management

MDM and Data Warehousing Complement Each Other

Greater business value from both

Executive Summary

- **Master Data Management (MDM) and Data Warehousing (DW) complement each other**
- **There are areas of overlap**
  - This is a positive benefit, not an unnecessary duplication of effort or data
- **MDM and DW provide quality data to the business but MDM is valuable beyond the DW for 2 reasons**
  - Latency
  - Feedback
- **MDM and DW have different use cases**
  - MDM provides a “golden” source of truth that is used collaboratively for authoring, operationally in the transactional / operational environment and supports the delivery of "quality" Master Data to a DW system
  - DW systems are a multidimensional collection of historical transactional data that may be include than Master Data used to determine trends and create forecasts
  - Introducing MDM enhances the value of existing DWs by improving data integrity and closing the loop with transaction systems

Executive Summary

- "… the MDM environment can be key to the success of a data warehouse or new operational system, such as a CRM, SAP, or BI environment …"

Master Data Management and User Participation

Patty Haines, Chimney Rock Information Solutions

Agenda: Master Data Manager and Data Warehousing

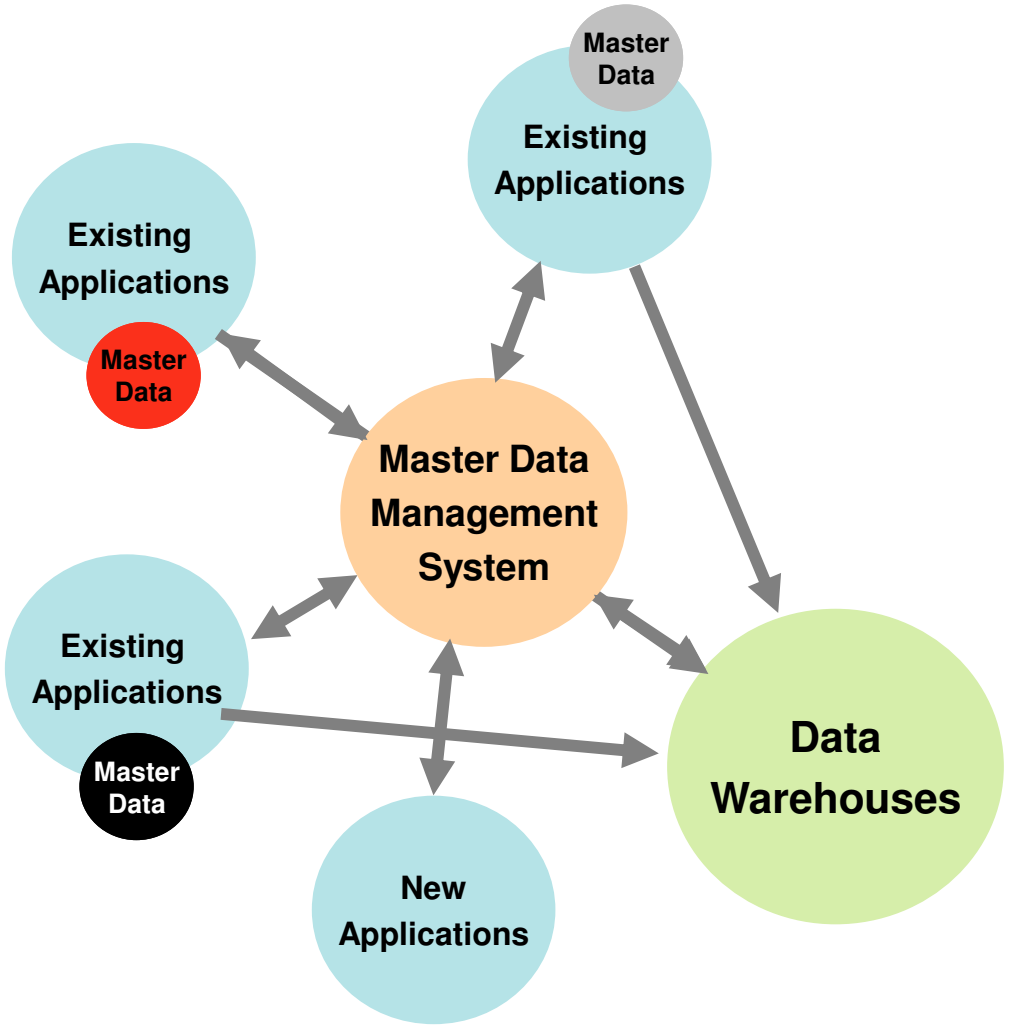
- **MDM Definitions**
- **Latency and feedback**
- **Step through the workflow**
- **Similarities and Differences**
- **Benefits and Unique Value**
- **MDM and DW use cases**
- **Summary**

Master Data Manager and Data Warehousing

- **Latency**
  - MDM can be real-time or near real time
    - A true stand-alone MDM platform can provide the robust, scalable platform required for real-time data correction
  - DW can have a time delay
    - DW must wait until information collected
- **Feedback**
  - MDM ensures new data is correctly entered *initially*
    - MDM systems become essential components of the operational applications
  - Many DW application correct data but do not feedback the corrected data to the original applications

Benefits from Master Data Management

- **Decouples master information from individual applications**
- **Becomes a central, application independent resource**
- **Simplifies ongoing integration tasks and new app development**
- **Ensure consistent master information across transactional and analytical systems**
- **Addresses key issues such as latency and data quality feedback proactively rather than "after the fact" in the data warehouse**



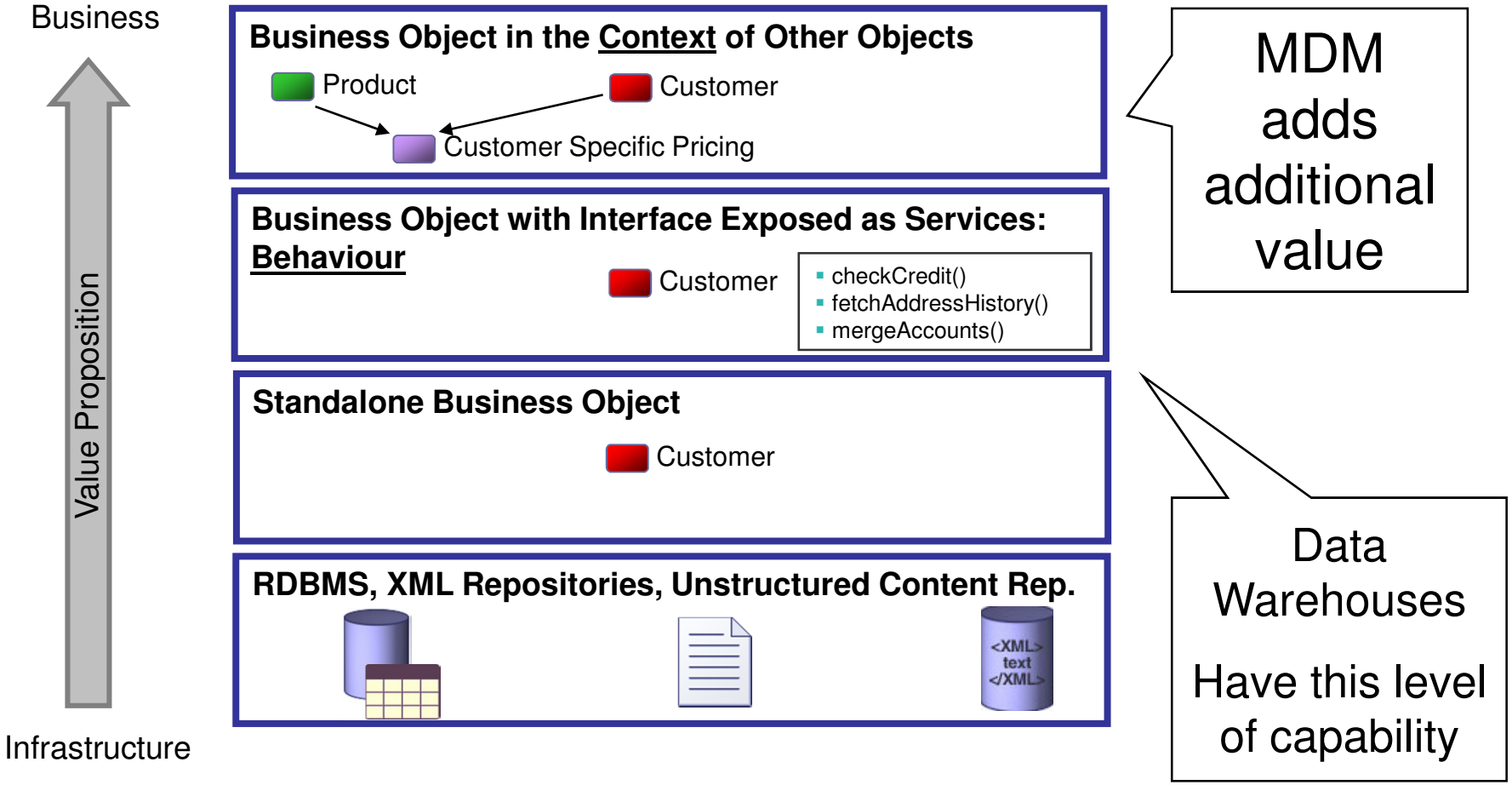
Master Data Management Systems: Methods of Use

- Collaboration
  - Workflow and check in / check out services to control Master Data creation, management and quality
  - Generally focused on Definition Master Data, but not exclusive

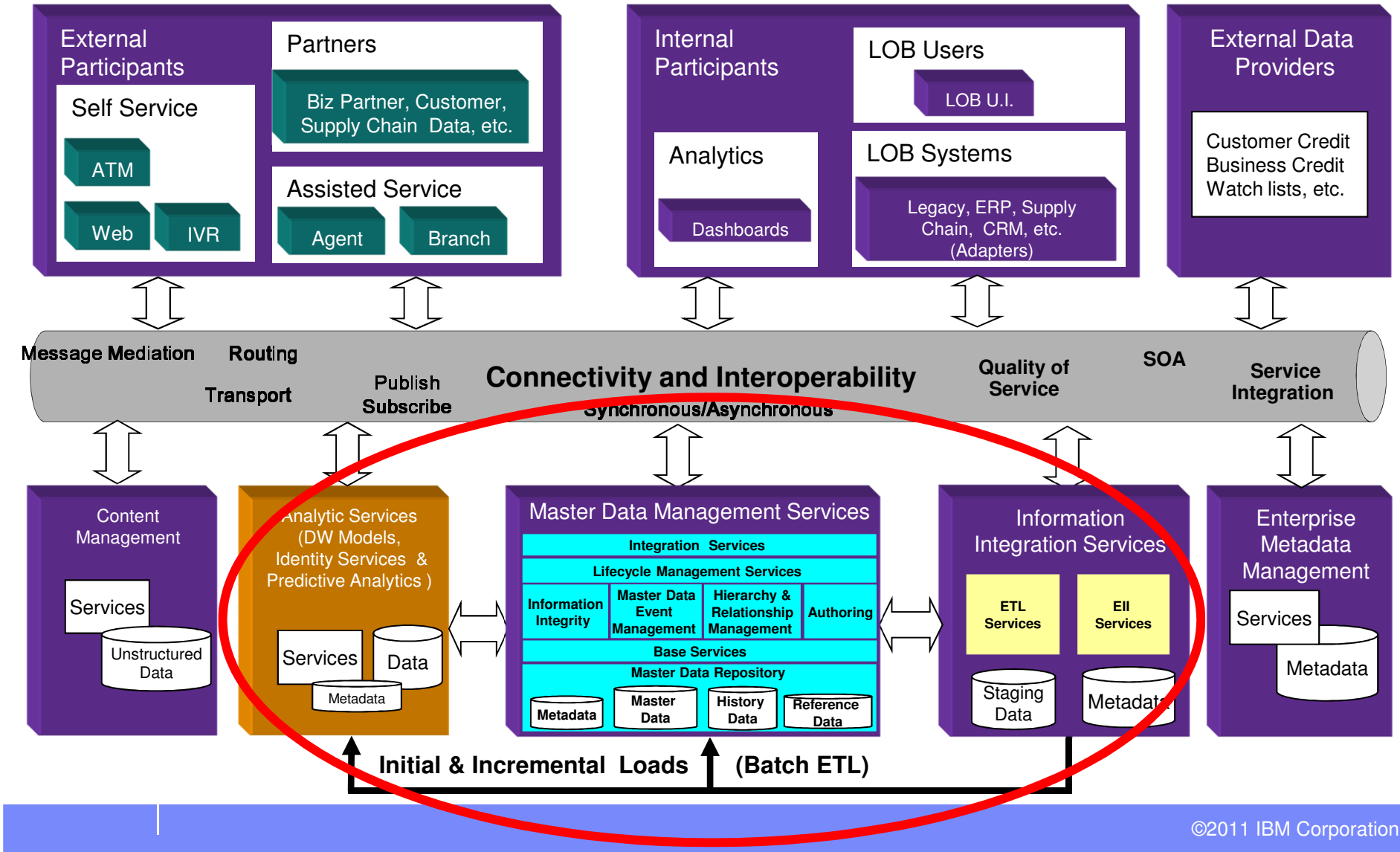
- Operational
  - SOA services control Instance Master Data creation, management, quality and access
  - Master Data is leveraged by other systems via real time SOA services
  - Consuming systems are dependent on the MDM System to perform transactions.
  - Instance data is considered the “system of record”

- Analytical
  - Master data is a fundamental and important source for analytical environments
  - Master data is key to simplifying the input to analytic environments and improving the quality

MDM Builds On Infrastructure and Provides *Context*



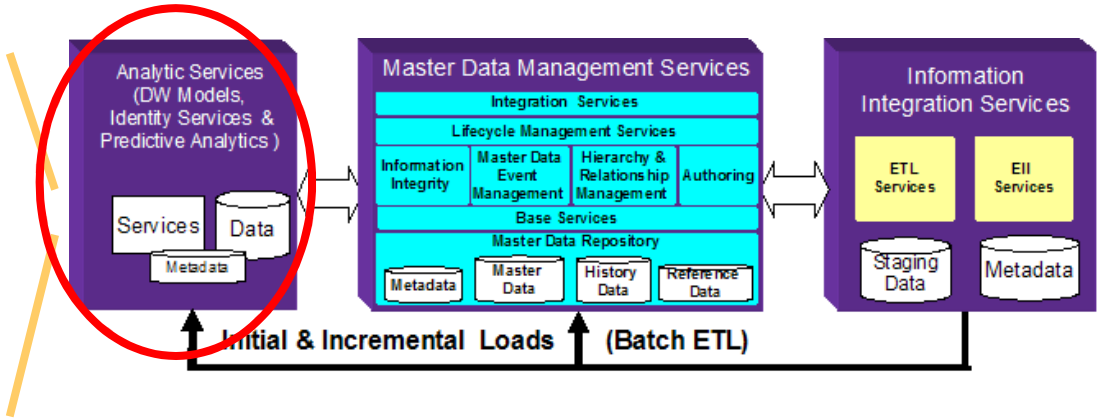
MDM Logical Architecture



MDM Technical Architecture with Data Warehousing

Data Warehouse

The data warehouse provides historical analytical capability for business analysis and corporate dashboards. The DW holds master data copies and historical transaction detail. The information is aggregated and for use in specific analytical applications such as forecasting and budgeting systems. MDM provides clean information to the DW. MDM also supplies analytical information for master data objects. The same infrastructure can be used to populate the DW as is used for the MDM system initial loads and batch load cases.



Data Warehouse Services

- MDX
- SQL
- Web Services
- XML

Data Warehouse Uses

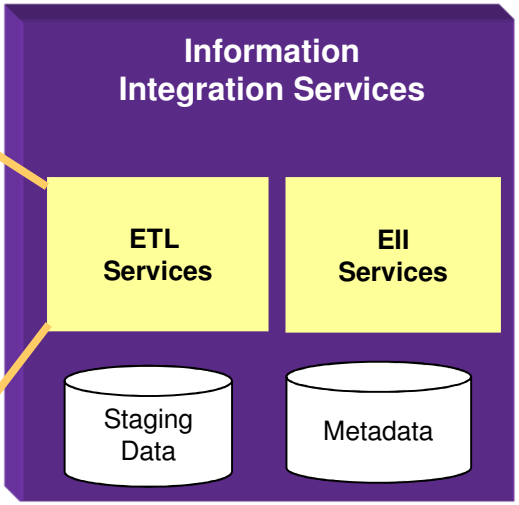
- Historical transaction detail
- Data mining
- Forecasting
- Trending
- Dash boards

MDM and Data Warehousing share Information Integration Services

Information Integration ETL Services Overview

Information Integration ETL Services provide services that support the loading of bulk data and near real-time replication of data from one or more source systems into a target database such as a Master Data Repository or Enterprise Data Warehouse. The loading of bulk data refers to loading a large volume of data as part of the initial or incremental load of data into a target database. Services are available to support the extraction of large volumes of data form a source system, detect changed data for the replication of data to a target system, cleanse and standardize data, match data from multiple sources, transform and load data into a target system. Data Profile and Analysis services provide the ability to understand and model source system data to determine the rules necessary to match, transform and load the data into the target database.

The Metadata Repository provides a business glossary, rules for data cleansing and data transformation. The Staging Database provides a storage area for the cleansing, merging and transformation of data.



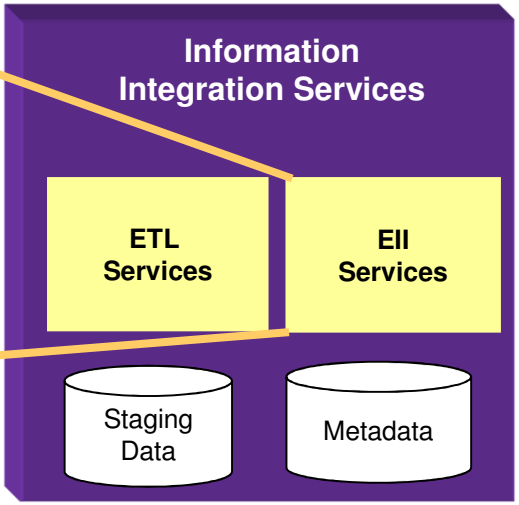
ETL

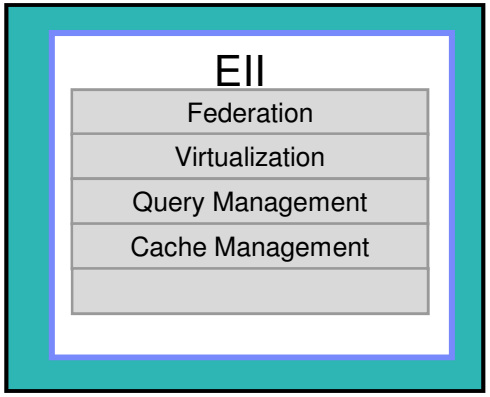
Extraction	Detection
Translation	Profiling
Transformation	Analysis
Load/Apply	Cleanse/Standardize
Replication	Match

MDM and Data Warehousing share Information Integration EII Services

Information Integration EII Services Overview

Enterprise Information Integration (EII) Services provide the capabilities to access structured and unstructured content contained in disparate data sources across multiple LOB systems. Cache Management and Query Management Services are used to optimize query performance for locating, accessing and aggregating query results.

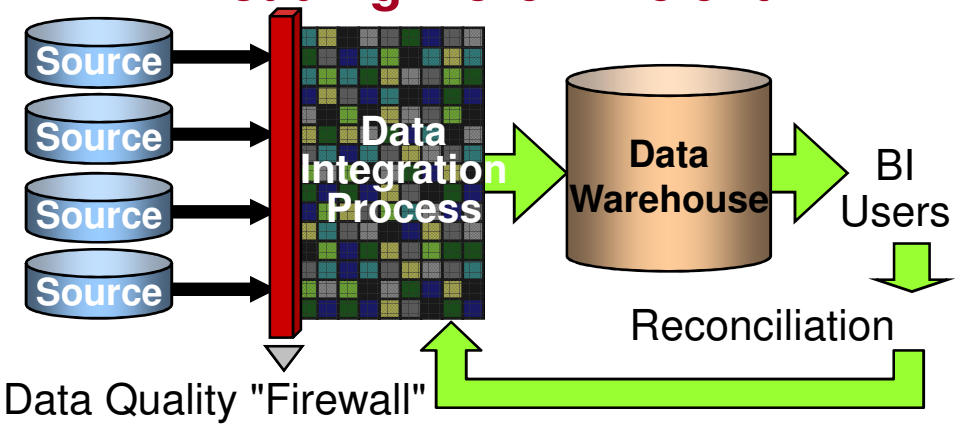




Workflow MDM to Data Warehousing

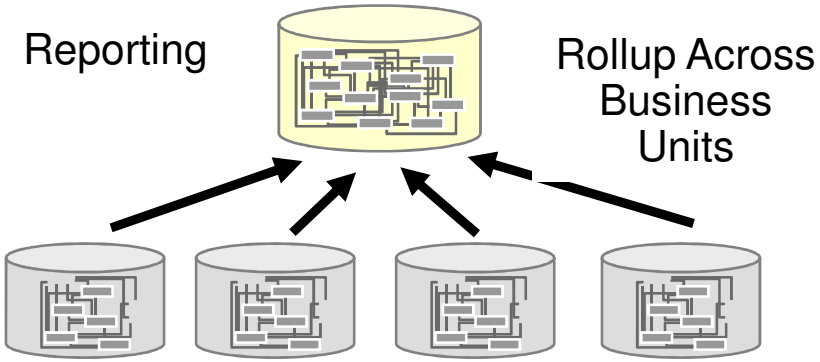
How Does MDM Fit In With Analytical (As Opposed to Operational) Systems?

Improving the Data Quality of the Data Warehouse and Making Loading More Efficient



- Extracting, transforming and loading data into data warehouses is hard work
- We are continually cleaning up data pollution creating upstream in the operational systems
- A focus on MDM in the operational world should improve
  - the efficiency and cost of loading the data warehouse
  - the quality of the data warehouse

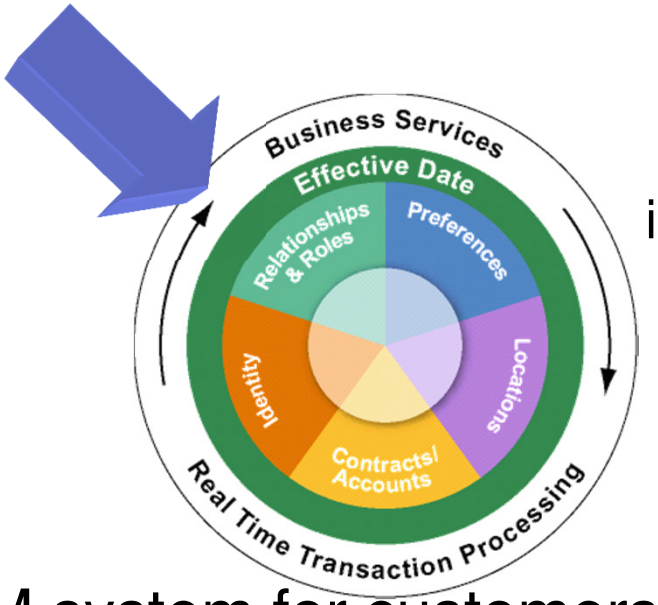
Providing Consistency for Performance Management



- There can be heterogeneity in the Business Intelligence world as well
- There may be inconsistency across data marts held at the national or line of business level
- Rollup for reporting purposes may not be possible without a form of MDM
- MDM for performance management is mainly about metadata

Workflow MDM to Data Warehousing

New customer is created In Call Center application

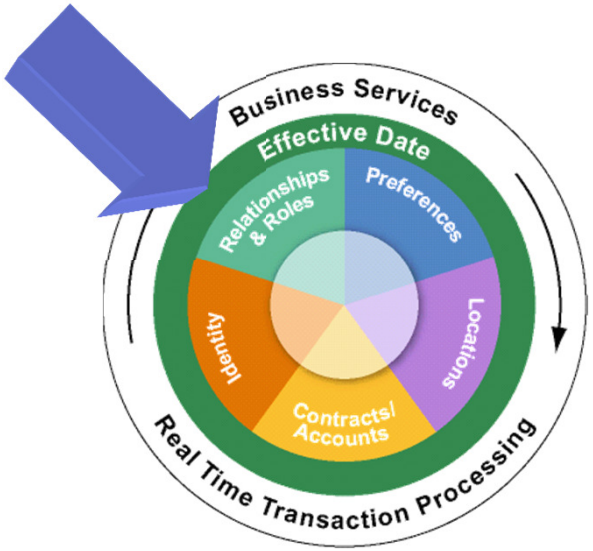


MDM verifies address and other demographic information, ensures this is a not a duplicate, and follows the business rules for new customers

MDM system for customers (CDI)

Workflow MDM to Data Warehousing

New customer is created In Web based application



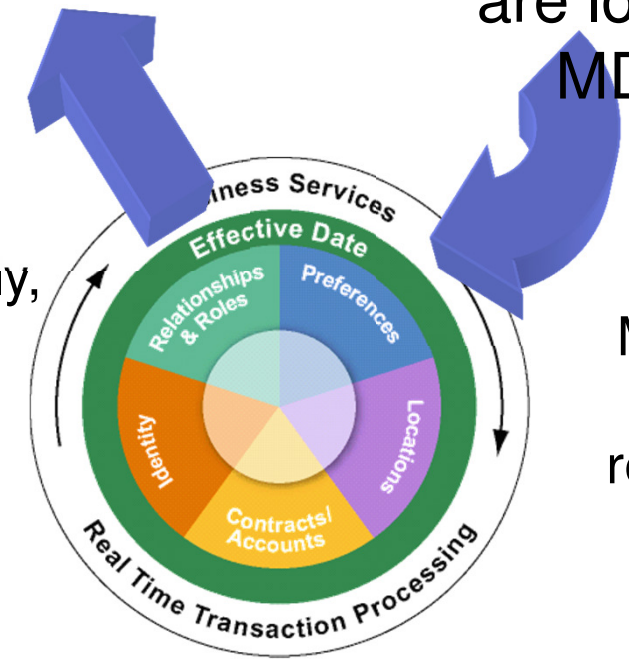
Again MDM verifies address and other demographic information, ensures this is a not a duplicate, and follows the business rules for new customers

MDM system for customers

Workflow MDM to Data Warehousing

All new customers this week are contacted via the CRM application

Results of the contacts made by the CRM application are loaded into the MDM system



MDM provides a timely, clean list of new customer from all systems

MDM now holds “no call” information, responses marketing channels

MDM system for customers

Workflow MDM to Data Warehousing

Employees and parts suppliers add information about tasks throughout New Product development process

Information about the new product is available to other applications. For example a customer-focused web app that lists the languages available for the product

MDM data during all the steps required to create a new product

MDM holds all data about the product and can track access and usage

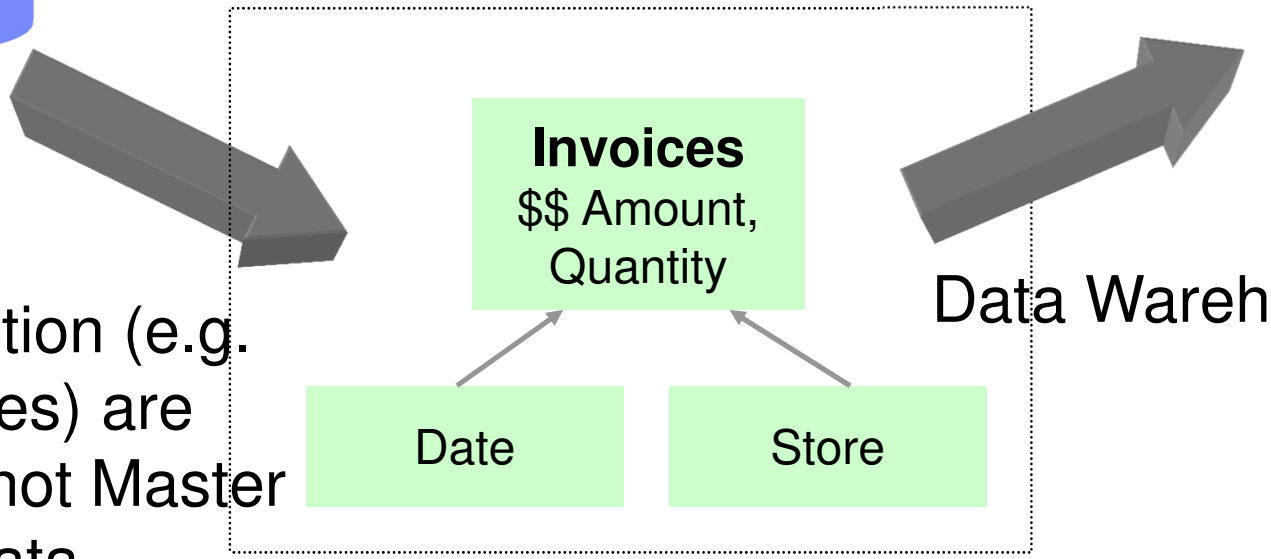
MDM system for products (PIM)

What is NOT Master Data

Managers and Analysts access the data warehouse to forecast business trends and manage their business areas

Application provides DW transactions such as invoices

Billing system



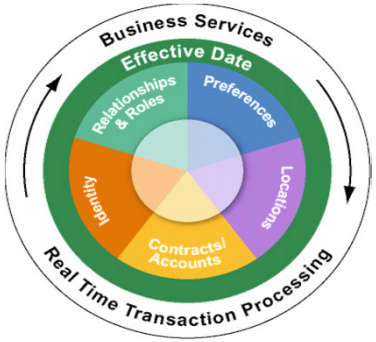
Data Warehouse

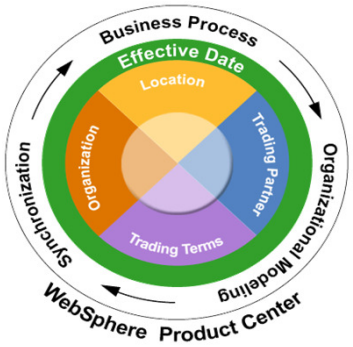
Transaction (e.g. invoices) are typically not Master Data

Intersection of MDM and DW

Application provides DW transaction data

Billing system





MDM maintains DW dimensions of customer and product

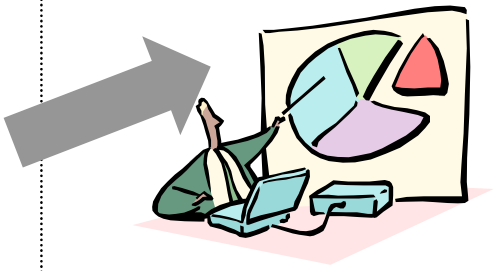
Customers

Products

Invoices
$$ Amount, Quantity

Date

Store



Data Warehouse

(Now complete with dimensions managed by MDM)

Similar value to the organization

- **Single version of the truth**
- **Reduce redundancy and data discrepancies increase consistency**
- **Centralize management of function**
- **Increase insight**
- **Enable analysis across departments**

Similar enabling technologies

- **Data quality is a prerequisite**
- **Services infrastructure, data movement and loading technologies must be in place**
- **Metadata is required for each**

Similar enabling processes

- **Data governance, data stewardships**
- **Inter organizational cooperation**
- **IT builds the systems but other departments must sponsor the projects**
- **Unified security**
- **Unified modelling of the business processes**

But different

- **MDM is more for Transactional purposes vs. DW which are more for analytical purposes**
- **MDM is used by transactional applications**
- **DW's are used by managers and analysts**
- **MDM captures business rules for entities**
  - Not the results of the business processes, does not hold transactions
- **DW's capture and *analyse* historical facts**
  - Do not hold business rules
  - Historical transaction data does not change, user do not usually update transactional data
- **MDM captures and enhances customer data, product data, etc.**

But different (continued)

- **MDM hub is optimised to support transactional systems**
  - Data models are more normalized
  - Will not grow faster than the business, for example will grow at the rate new customers are added
  - Users are transactional systems, call centers, inventory systems, ERP systems
- **DW is optimized to support analytical functions**
  - Data models are de-normalized star schemas
  - Can grow to be very large by holding historical transaction data
  - Users are analytical and forecasting systems, analysts and managers.

Unique benefits from DW

- **Historical reporting beyond the capabilities of the online systems**
  - Can store and analyse data pat the retention limits of online systems. Most online system keep 6-18 months of data only
- **Ability to report across multiple applications**
  - For example “what was the average commission of sales people who worked in the top 10 stores by revenue?”. This query crosses payroll and invoicing systems.
- **Single version of the truth about business “facts”**
  - Example: Total number of phone calls made IBM
  - Example: Total value of all invoices last week
  - Transaction data is not considered Master Data

Unique benefits from MDM

- **Write back to ERP and other systems clean data**
  - This is what most DW's cannot deliver
- **Consistent treatment of customers, products, suppliers**
  - All systems have the same information about each customer
- **Reduces application development time for new analytical and online applications**
  - This can be the most valuable aspect of an MDM system
- **Reduces "chatter" between apps**
  - The spiders web of interconnection is reduced through the use services

Benefits to DW from MDM

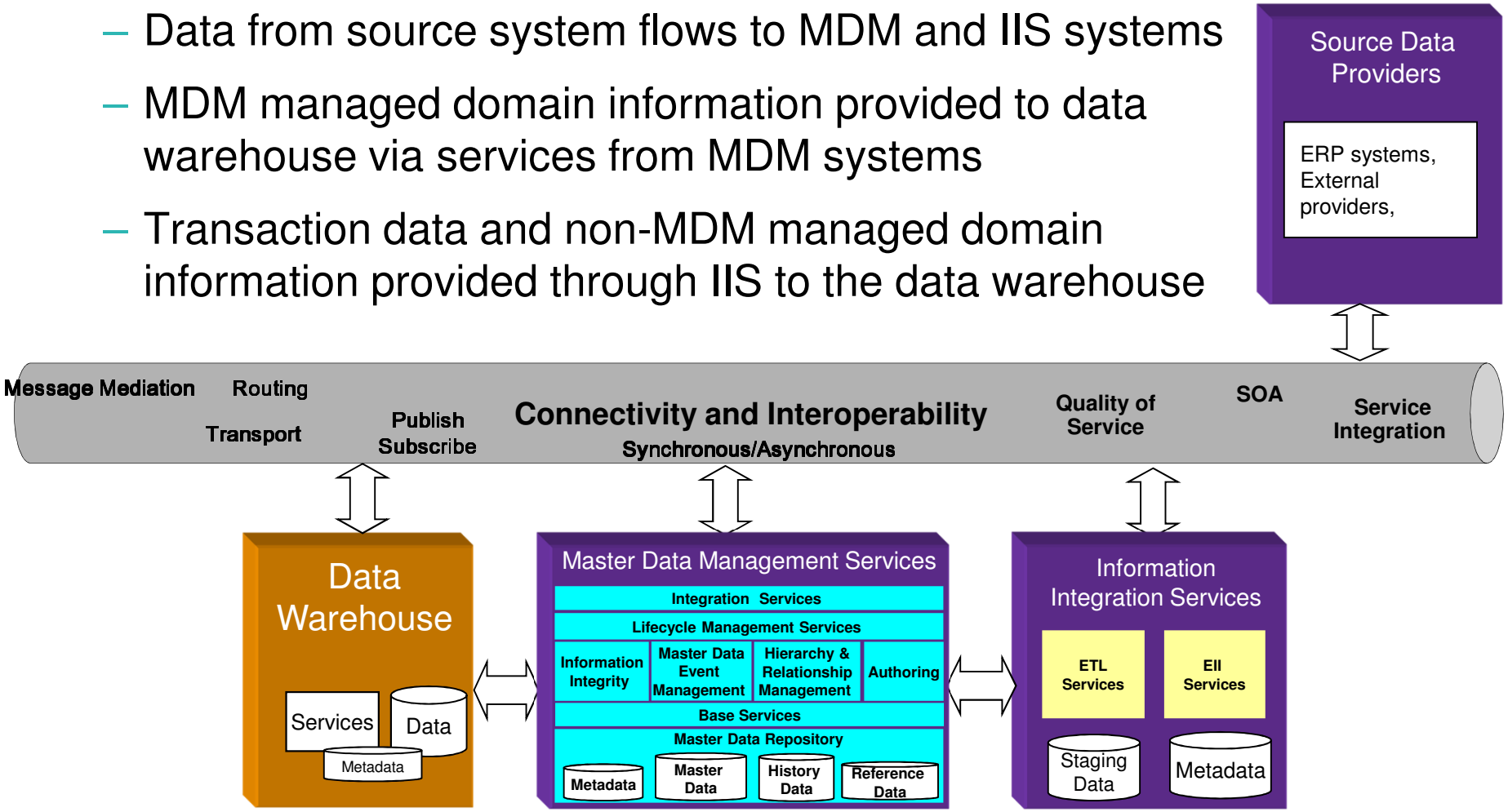
- **Organizational focus on MDM and governance directly benefits DW usability and acceptance**
- **Current and accurate dimensions**
  - From dimensions managed by MDM hubs
- **Off loading of some dimension maintenance from the DW ETL process**
  - Dimensional data is managed outside of the DW but some processing is still required

Use Cases for Data Warehousing and MDM

- **Operational Data Store / Dynamic / real-time warehousing**
- **Periodic or batch loading of MDM managed domains**
- **Vendor specific solutions**

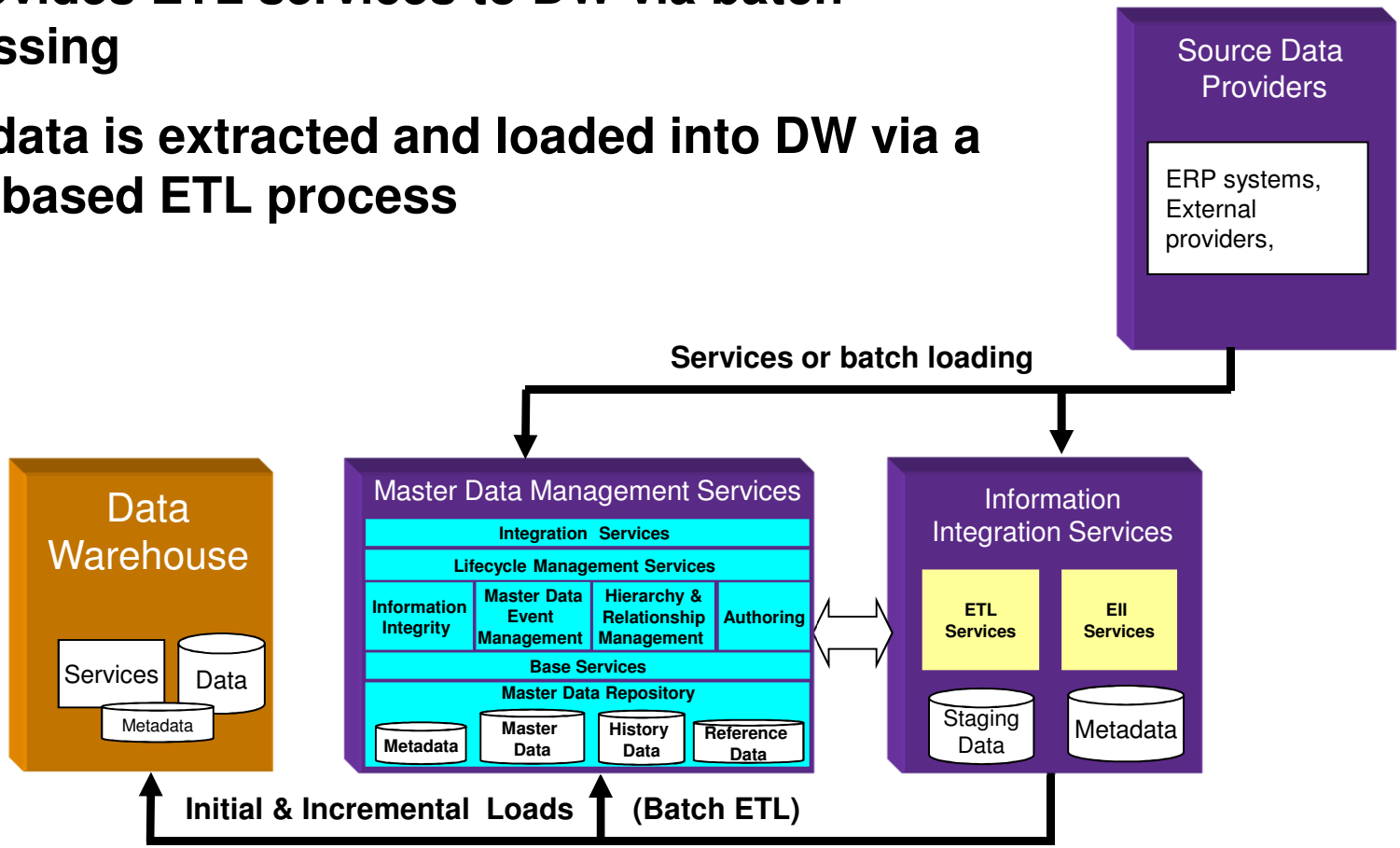
Operational Data Store / Dynamic / real-time warehousing

- **When an update occurs new source data sent to MDM and IIS**
  - Data from source system flows to MDM and IIS systems
  - MDM managed domain information provided to data warehouse via services from MDM systems
  - Transaction data and non-MDM managed domain information provided through IIS to the data warehouse



Periodic or batch loading of MDM managed domains

- **Traditional approach via periodic extracts from source systems and MDM systems**
- **IIS provides ETL services to DW via batch processing**
- **MDM data is extracted and loaded into DW via a batch based ETL process**

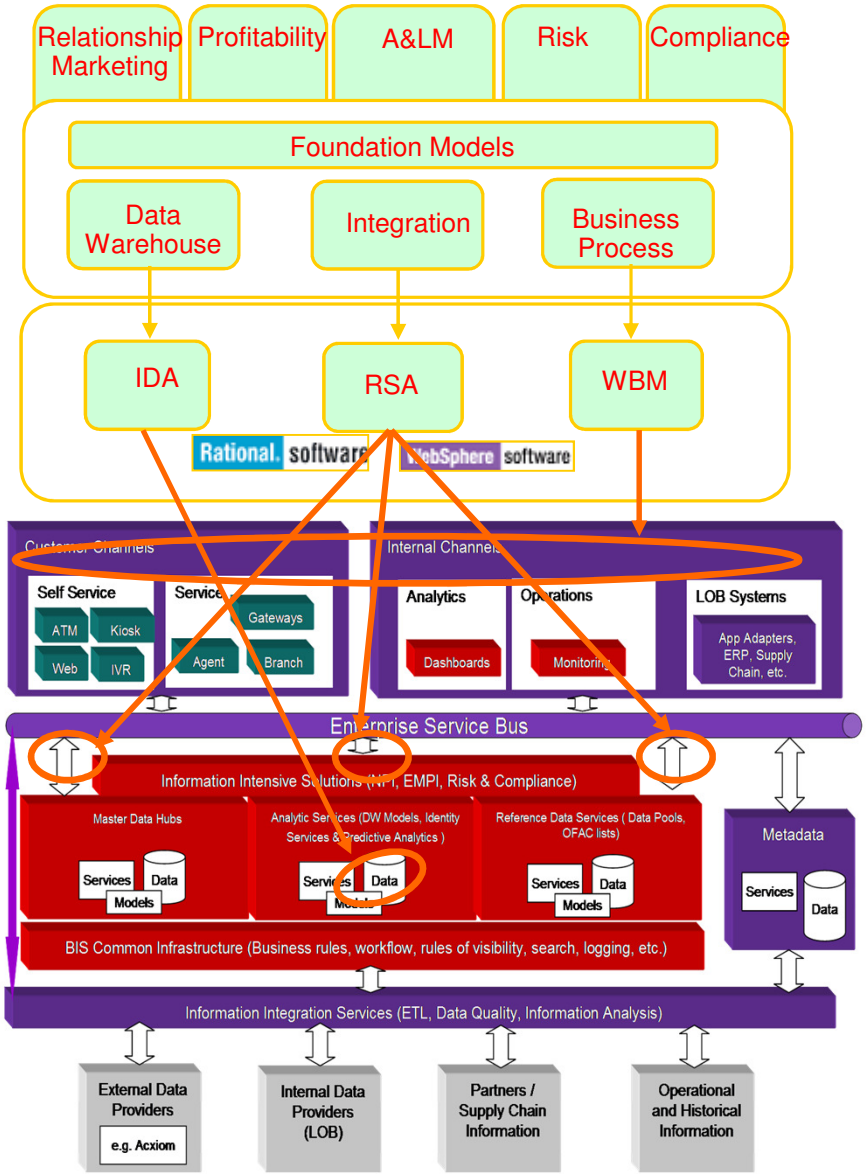


Data Warehousing and Master Data Management

Collaboration can Gain Control Enterprise Data and Business Process Needs

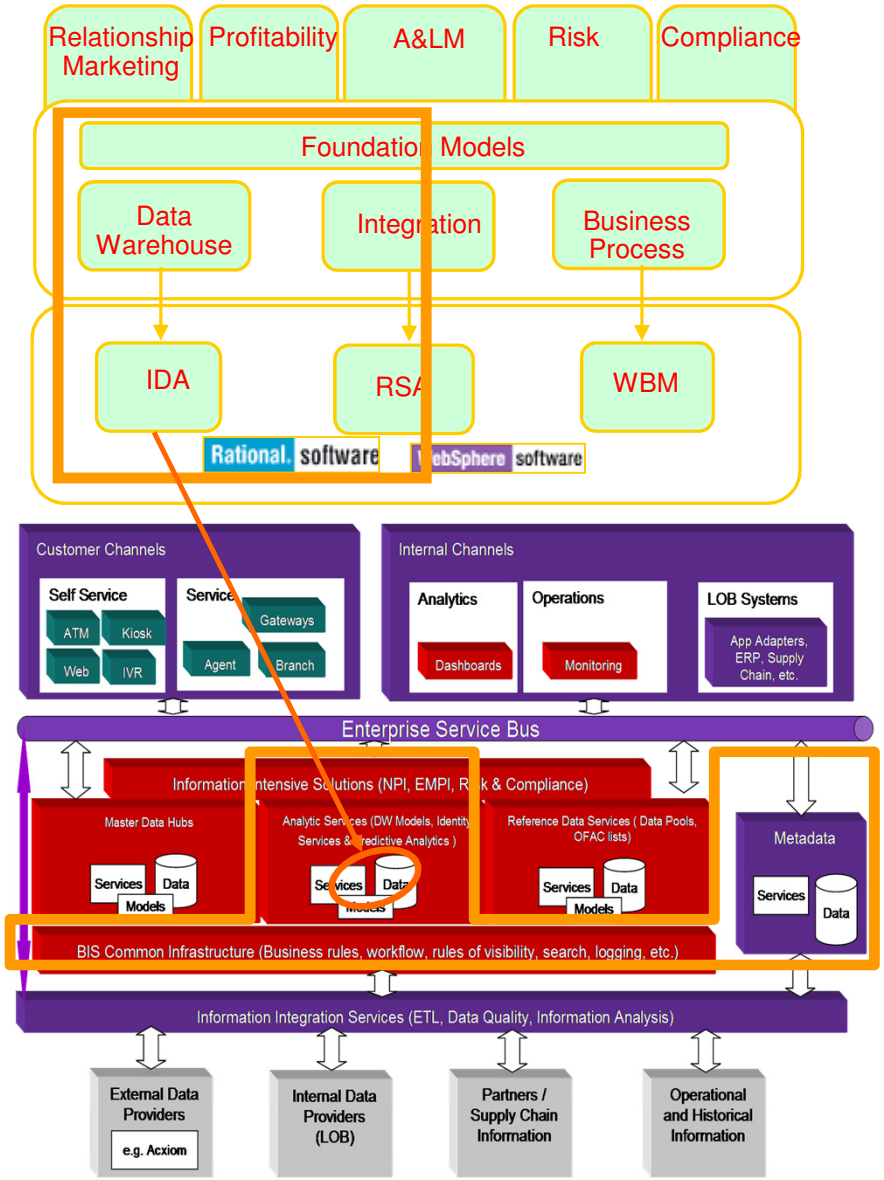
- Shared information infrastructure
- Shared data governance and stewardship policies
- Shared development environments
- Different users and use cases
- Different platforms and presentations layers

Models and Tooling – A single deployment platform



- The Models combined with integrated tooling – single enterprise deployment platform
- Models housed in standard IBM Modelling tools (supplemented by specialist tooling)
  - Generation capabilities enable to deployment of runtime artefacts (DDL, WSDL, BPEL)

Deploying for Business Intelligence

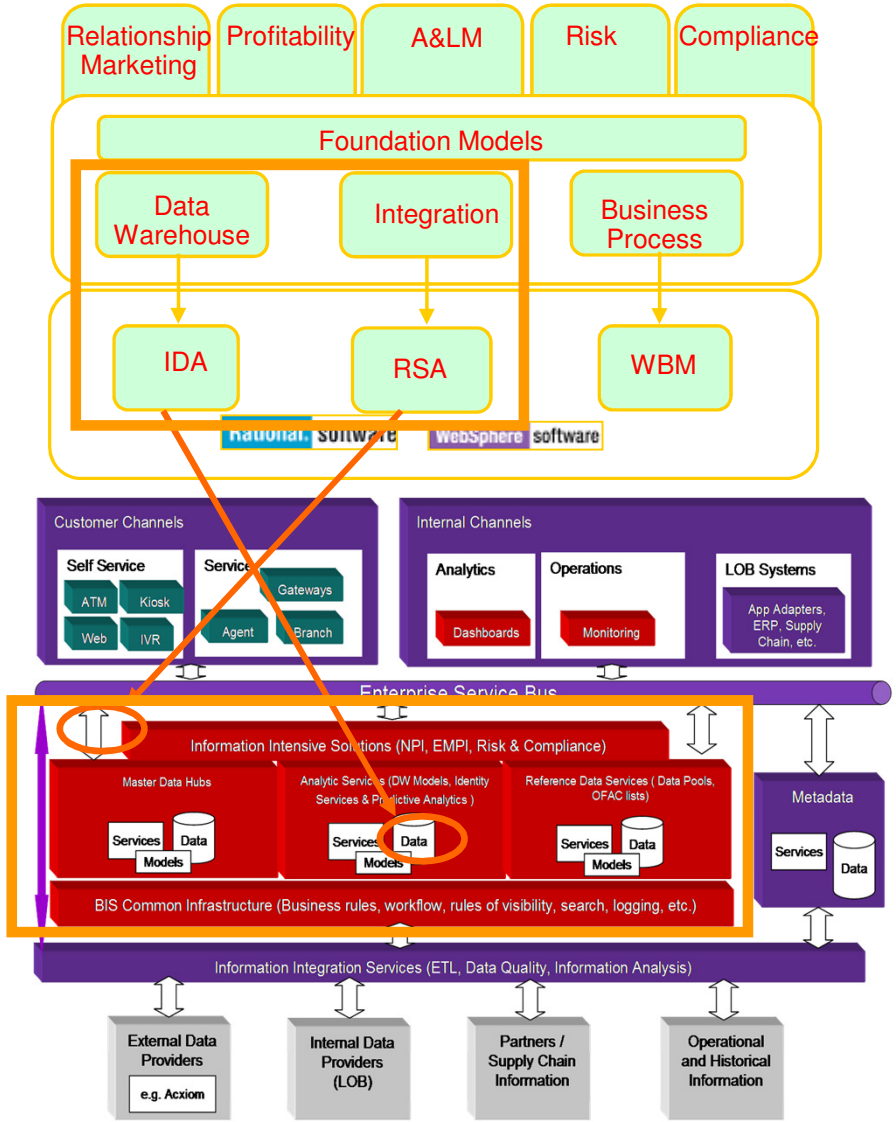


- Tighter and ongoing integration with IDA
  - Improved BDW Implementation Guidelines/Methodology

Future

- Initial Support for WebSphere Metadata Server
  - Websphere Business Glossary –
- DB2 Physicalization Guidelines for Data Warehouse Models
- Tighter integration with XBRL capabilities in DB2

Deploying for Master Data Management



- MDM products provide the “anchor” hub components for Customer and Product Master Data and other hubs as they appear
- Industry Models combined with IBM tooling provide wider infrastructure capability
- Initial Mapping activities completed

Future

- Development interlock
  - Model content enhancements to aid MDM-integration included in 2009 release
- Future Tooling integration
  - Achieve Dynamic mappings by incorporating MDM models in the Tooling

Summary

- **Additional value is added to data warehouses when MDM is implemented**
- **MDM and data warehouses complement each other**
- **A more sophisticated, enterprise wide view of the organisation is provided through the combination of MDM and DWs**