



helping build the smart business

Using Master Data in Business Intelligence

**Colin White
BI Research**

**March 2007
Sponsored by SAP**

TABLE OF CONTENTS

THE IMPORTANCE OF MASTER DATA MANAGEMENT	1
What is Master Data Management?	2
The Business Benefits of MDM	2
The Ultimate Goal: Enterprise MDM	3
Business Area MDM versus Enterprise MDM	5
IMPLEMENTING MDM	6
Operational Master Data	6
Business Intelligence Master Data	6
Where Should Historical Master Data be Managed?	8
An Evolutionary Approach to MDM	9
THE SAP NETWEAVER MDM SOLUTION	11
SAP Support for Master Data Management	11
Using SAP NetWeaver MDM	13

Brand and product names mentioned in this paper may be the trademarks or registered trademarks of their respective owners.

THE IMPORTANCE OF MASTER DATA MANAGEMENT

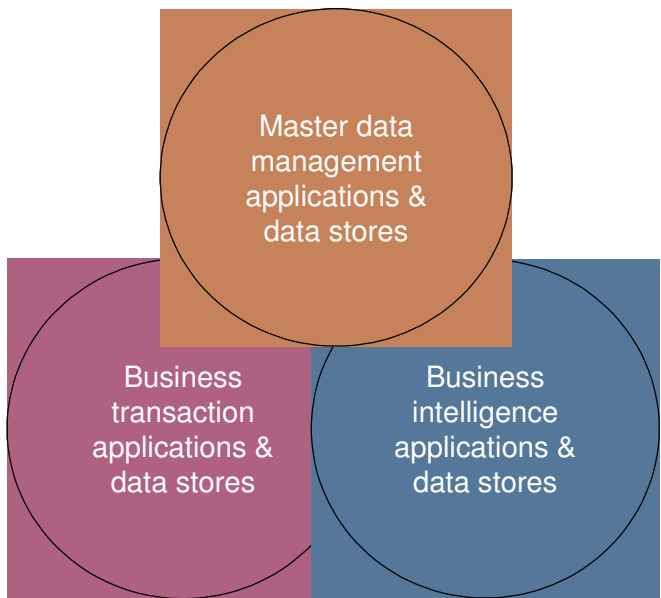
The current industry focus on master data management (MDM) creates the impression that MDM is a brand new technology, but this not the case. Companies have been building their own custom-built business transaction (BTx) solutions for managing operational master data for many years, and business intelligence (BI) applications and their underlying data warehouses have been providing a single view of key master data entitles such as customers and products for analytical purposes for some time.

Master data should be managed separately from business transaction and BI systems

What is new about MDM is that enterprises are beginning to realize that master data needs to be managed and integrated outside of the traditional business transaction and business intelligence environments, and then used to supply BTx and BI applications with the master data they need (see Figure 1). Another important development in MDM is that vendors are providing packaged solutions that ease and reduce the effort involved in deploying an MDM environment and its associated applications.

This paper reviews the current status of MDM, and offers suggestions for planning, building and deploying an MDM environment. The paper proposes an evolutionary, rather than revolutionary, approach to MDM, and examines how the MDM environment can be used in conjunction with BI applications and an underlying data warehousing environment to improve the decision making process. It also presents an overview of the SAP MDM solution and how it supports many of the ideas presented in this paper.

Figure 1.
The three types of IT processing



WHAT IS MASTER DATA MANAGEMENT?

Master data is reference data about an organization's core business entities. These entities include people (customers, employees, suppliers), things (products, assets, ledgers), and places (countries, cities, locations). The applications and technologies used to create and maintain master data are part of a master data management (MDM) system. MDM, however, involves more than applications and technologies, it also requires an organization to implement policies and procedures for controlling how master data is created and used.

The system of record is the gold copy of master data

One of the main objectives of an MDM system is to publish an integrated, accurate, and consistent set of master data for use by other applications and users. This integrated set of master data is called the master data *system of record* (SOR). The SOR is the *gold copy* for any given piece of master data, and is the single place in an organization that the master data is guaranteed to be accurate and up to date.

The goal is for the MDM system to also be the primary system of entry

Although an MDM system publishes the master data SOR for use by the rest of the IT environment, it is not necessarily the system where master is created and maintained. The system responsible for maintaining any given piece of master data is called the *system of entry* (SOE). In most organizations today, master data is maintained by multiple SOEs. Customer data is an example. A company may, for example, have customer master data that is maintained by multiple Web store fronts, by the retail organization, and by the shipping and billing systems. Creating a single SOR for customer data in such an environment is a complex task. The long term goal of an enterprise MDM environment is to solve this problem by creating an MDM system that is not only the SOR for any given type of master data, but also the SOE as well.

MDM then can be defined as *a set of policies, procedures, applications and technologies for harmonizing and managing the system of record and systems of entry for the data and metadata associated with the key business entities of an organization.*

THE BUSINESS BENEFITS OF MDM

MDM help solve four key master data problems

MDM helps organizations handle four key issues:¹

- **Data redundancy:** Without MDM, each system, application, and department within an organization collects its own version of key business entities. This leads to redundant master data and poor data quality.
- **Data inconsistency:** Enterprises spend enormous resources trying to reconcile master data, often with limited success. Furthermore, this reconciliation process is repeated over and over because there is no mechanism to capture the data assets garnered from the first or succeeding reconciliations.

¹ Colin White and Claudia Imhoff, "Master Data Management: Creating a Single View of the Business." Business Intelligence Network Research Report, September 2006.

- **Business inefficiency:** Redundant and inconsistent master data leads to inefficient supply chain management, inconsistent customer support, customer dissatisfaction, and wasted marketing efforts. Fractured master data in business processes causes ineffectiveness and inefficiency.
- **Supporting Business change:** Organizations are constantly changing as new products and services are introduced and withdrawn, companies are acquired and sold, and new technologies appear and reach maturity. These disruptive events cause a constant stream of changes to master data, and without a way of managing these changes, the issues of data redundancy, data inconsistency, and business inefficiency are exacerbated.

THE ULTIMATE GOAL: ENTERPRISE MDM

There are many different approaches to integrating and managing master data. Before discussing these approaches, let's first identify the ideal architecture of an enterprise MDM system. This architectural goal is achieved through an iterative and evolutionary application development process.

Enterprise MDM maintains and publishes all master data

In an enterprise MDM system, all master data is maintained and published to business users and other IT systems using MDM applications. These applications handle master data and metadata changes, and maintain a historical record of those changes. An MDM application could, for example, manage and track customer account data such as account identifiers, customer names and addresses, credit ratings, etc.

MDM does not handle standard operational business transactions

The enterprise MDM system propagates master data to other internal and external IT systems as required. It also provides business views of master data that can be used by business users and applications to directly access master data in the MDM system itself. MDM applications do not handle or manage other types of business transaction data such as customer account deposits and withdrawals. This data is managed and distributed by other business transaction applications.

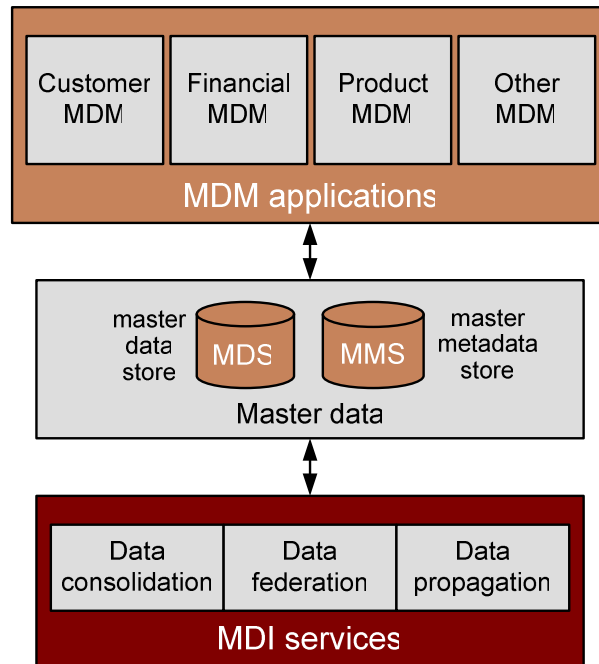
Figure 2 shows the main components of an enterprise MDM system. These components include:

- **MDM applications** for managing and publishing master data and metadata.
- **A master data store (MDS)** containing consolidated master data.
- **A master metadata store (MMS)** containing the master data business model, and master data rules and definitions. The master data business model documents master data entities, attributes, relationships and their business meaning.
- **A set of master data integration (MDI) services** for consolidating, federating and propagating master data.

Business users employ custom-built and/or packaged MDM applications to access and maintain master data in the MDS. The MDS represents the SOR for enterprise-wide master data. Information about the SOR is documented and maintained in the

MMS. As master data is created and maintained, MMS business rules ensure that the master data conforms to the business practices of the organization.

Figure 2.
MDM Components



Business users employ custom-built and/or packaged MDM applications to access and maintain master data in the MDS. The MDS represents the SOR for enterprise-wide master data. Information about the SOR is documented and maintained in the MMS. As master data is created and maintained, MMS business rules ensure that the master data conforms to the business practices of the organization.

Master data integration services should be a component of an enterprise integration architecture

Underlying master data integration (MDI) services provide the capabilities for MDM applications to integrate and copy master. These integration services should be a key component of an organization's enterprise integration architecture. Key requirements here include:

- Data quality management
- Metadata integration and propagation
- Synchronous and asynchronous data propagation with guaranteed delivery
- Change data capture and data transformation
- Data federation
- A service-oriented architecture (SOA)

An MDM system should act as an intelligent source of master data that drives other IT systems.

In a fully compliant enterprise MDM environment, all master data and metadata is managed by the MDM system, i.e., the MDM system is also the SOE for all master data.

An MDM system and its applications and services are often implemented as tactical extensions to existing business transaction and business intelligence projects. To be successful, however, a strategic MDM initiative should be approached as an independent enterprise-level project with strong executive backing. An MDM system should act as an intelligent source of master data that drives other IT systems. It should not simply consist of a set of adjunct IT applications that gather and integrate existing master data to overcome the problems caused by dispersed master data management.

Bottom-up tactical stealth projects driven by IT groups may be a way to get started in MDM, but an enterprise must develop a strategic MDM plan if it is to be successful in managing master data over the long term by avoiding solutions that simply alleviate the symptoms of master data problems, rather than cure them.

BUSINESS AREA MDM VERSUS ENTERPRISE MDM

There is a risk of multiple master data silos being deployed in an organization

Some MDM initiatives are targeted at addressing a specific business need such as creating a single view of the customer or a single parts catalog. The issue here is that although these projects are faster and less costly to implement than full enterprise MDM, there is a risk of multiple master data silos being deployed in an organization. This is somewhat analogous to the data mart and enterprise data warehouse issues that occur in business intelligence.

The long term goal should be to develop a consolidated master data business model and to have integrated master data and metadata stores. All tactical master data project managers should keep the long-term MDM objective in mind when designing and deploying master data applications.

The best approach to balancing the needs of short- and long term master data requirements is to have a master data practices group that is responsible for helping support the strategic master data objectives of the organization.

Now that we have a good understanding of the objectives of MDM, and what an ideal enterprise MDM should look like, we are in a position to discuss different approaches to deploying MDM solutions.

IMPLEMENTING MDM

OPERATIONAL MASTER DATA

Four approaches to managing master data

In a traditional operational environment master data is dispersed across multiple operational systems and intermingled with other types of business transaction data. There are four approaches that can be used to create and maintain a consistent view of operational master data in such an environment (see Figure 3).

1. **Consolidate** the master data from multiple operational systems of entry in a single master data store (MDS) that becomes the system of record for master data.
2. **Propagate** and synchronize master data changes between operational systems of entry so that master data in all systems of entry is kept consistent.
3. **Consolidate and propagate** master data using a combination of Approaches 1 and 2.
4. **Migrate and centralize** operational master data systems of entry to a new enterprise MDM system that acts as both the system of record and system of entry for master data.

Companies can start with Approach 1 and gradually evolve to Approach 4

Companies may start with the Approach 1, gradually implement Approaches 2 and/or 3, and then finally migrate to Approach 4.

In some cases, the fully compliant enterprise MDM system in Approach 4 may not be possible for political or technology reasons, and compromises may have to be made to allow exceptions to migrating certain systems of entry to the enterprise MDM system. The long term objective, however, is to evolve to the enterprise MDM environment of Approach 4.

BUSINESS INTELLIGENCE MASTER DATA

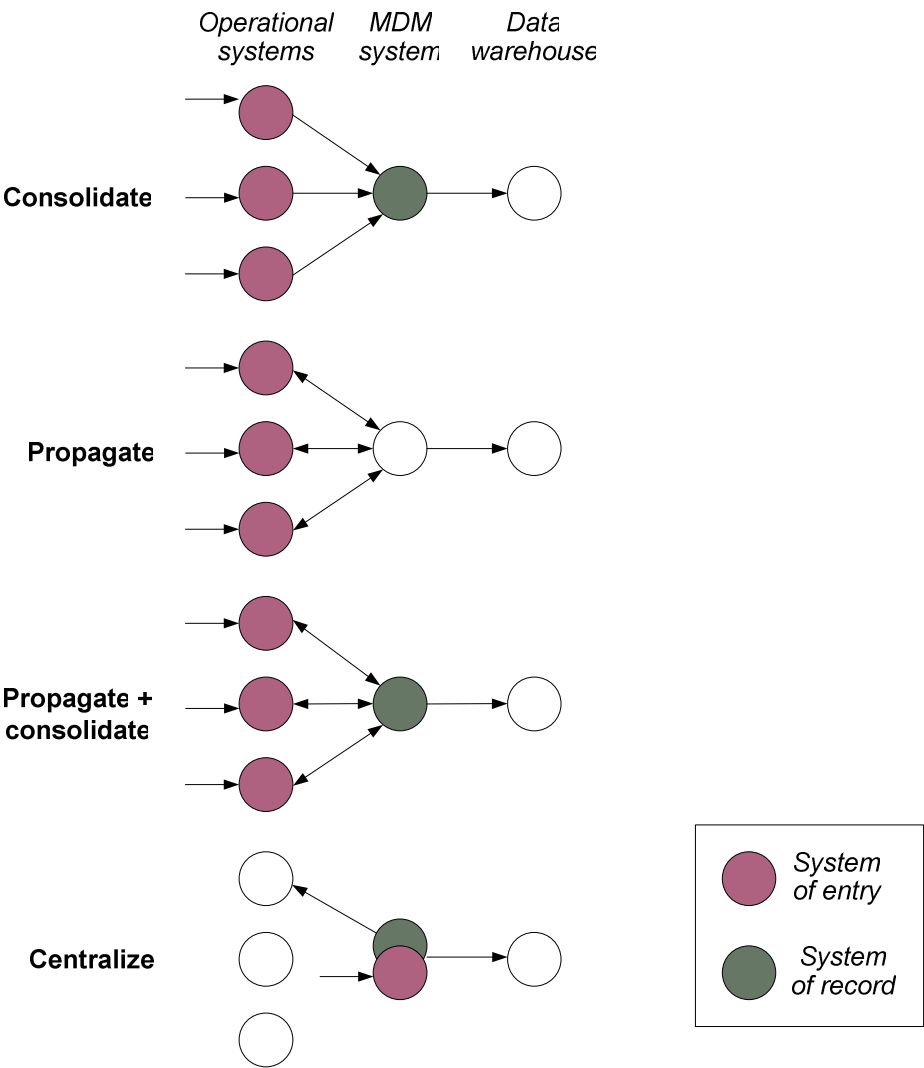
Master data can be used in BI for both historical and predictive analysis

For BI processing, master data documents how the data has changed over time. In the case of customer master data, a record could be kept of the various addresses customers have had, how their credit rating has changed over time, and so on. This historical master data can be combined with other historical business transaction data to produce analytical reports. Information about customers and their purchases could be used, for example, to produce reports identifying key customers, or for illustrating how customer buying patterns vary over of time based on their credit rating.

Master data can also be used in a BI environment for forecasting. If a company is proposing to reorganize its sales regions it could create a new set of master data to reflect the new sales organization and then combine this new master data with historical sales data to predict the effect of the changes on sales.

How historical master data is managed in a BI system depends on how that master data is managed in operational systems. Let's examine each of the four approaches listed above for operational master data management, and then look at how each approach affects the way master data is handled in a BI system (see Figure 3).

Figure 3.
Managing operational
and BI master data



**A master data store
can contain
operational master
data and/or historical
master data**

Approach 1: With this approach operational master data is consolidated into a single operational MDS. The operational MDS contains master data that has a zero or low latency compared with the data in the operational systems of entry from which it came. The operational MDS becomes the SOR for its associated master data. In some projects, the operational MDS is used to correct errors in source systems. For BI processing, there are two options for handling the master data in the operational MDS. The first option is to use the operational MDS as a data source for the data warehouse. The second option is to use a single integrated hybrid data store for both operational and historical master data.

Approach 2: With this approach, the operational master data is kept consistent, but it is still dispersed across multiple operational systems and intermingled with other types of business transaction data. As the master data is propagated between systems of entry, it can also be propagated to a data warehouse. Another option is the master in systems of entry can be extracted and integrated into a data warehouse in the same way that any type of operational data is brought into the BI environment. During the data integration process, however, data reconciliation should be simpler because the master data sources are kept consistent in the operational environment.

Approach 3: BI master data with this approach is handled in a same way as it is in the two options presented above for Approach 1.

Approach 4: With Approach 4, the enterprise MDM system acts as both the SOE and SOR for the master data. The master data store in this environment can be an operational MDS that becomes a data source for the data warehouse, or it can be a hybrid operational and historical MDS.

WHERE SHOULD HISTORICAL MASTER DATA BE MANAGED?

The management of historical master data is a controversial topic

We can see from the above discussion that historical master data can be managed in an historical MDS, in a hybrid historical and operational MDS, or in a data warehouse. Which of these options is best to use is a controversial and sometimes hotly debated topic. This controversy is made more confusing by the participants often failing to distinguish between logical and physical concepts. Also, the fact that many companies have used their data warehousing systems to provide a single view of a business entity, such as customers, leads to a position where the data warehouse is often seen as a good place to start a master data project. In the longer term this position may prove to be a bad one.

Master metadata management is the real issue in MDM

Managing master data is a logical data problem, rather than a physical one. In fact, it is master *metadata*, rather than master *data*, that is the real issue here. The models and definitions of most business entities are very complex and are constantly changing as businesses evolve. Take a look at a data model of a customer entity and you will see what I mean. Organizations need to track not only current master metadata and data, but also how this metadata and data changes over time. This is not only for analysis purposes, but also frequently for legal reasons as well. A facility to track and record relationships between different business entities is a requirement for many organizations. The ability to relate customers to the products they buy is a good example here.

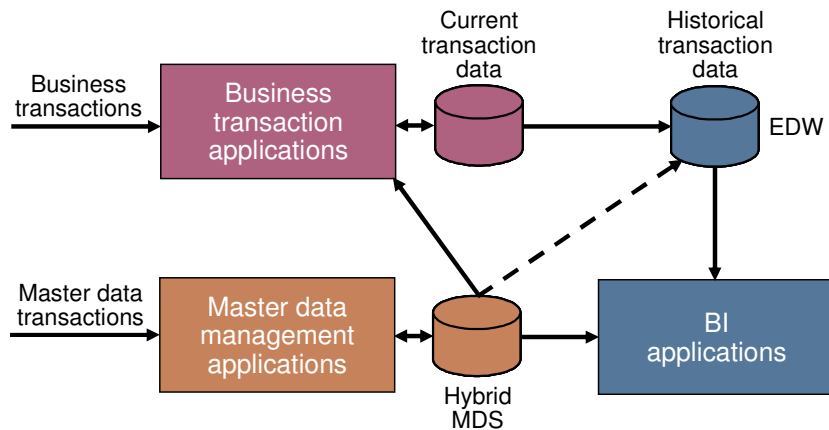
Current data warehouse design techniques are not suitable for handling historical master metadata

Managing complex master data hierarchies and relationships is best done outside of the data warehousing environment. Current data warehouse design techniques (such as slowing changing dimensions, for example) may be able to handle subsets of master data and master data relationships, but they are totally inadequate for supporting a complete picture of an organization's master data entities and relationships, and how they change over time.

Maintaining master data outside of the data warehouse environment also makes it easier for organizations to evolve to a full enterprise MDM environment with its own master data store. This store can be used for managing both current and historical

master data. In this environment, BI applications access master data from the hybrid MDS and historical business transaction data from a data warehouse. For ease of access and performance, a subset of the MDS data may be copied at regular intervals into the data warehouse. In a multidimensional data warehouse, this master data subset would be used to populate the dimension tables of the data warehouse (see Figure 4).

Figure 4.
An enterprise MDM environment



MDM applications will be business area specific

How MDM applications maintain and manage master data in a hybrid MDS will vary by the kind of master data involved. Processing customer master data is often very different from dealing with product master data. Business rules and models for defining customer name and address data, for example, are usually well understood by companies. Product data, however, is more complex and less standardized, and requires more semantic understanding of the data in areas like product naming and product codes. MDM applications, therefore, will tend to be business area specific, and it is unlikely that a company can create a one size fits all MDM application.

AN EVOLUTIONARY APPROACH TO MDM

MDI solutions do not address the issues that cause master data problems

Implemented correctly, master data management can provide significant business benefits in terms of improving productivity, reducing risk and increasing revenues. Many companies build MDM solutions by deploying master data integration (MDI) applications that are targeted at integrating current master data. These solutions typically do not track master data changes, nor do they provide capabilities for managing the constant changes that occur in master data hierarchies. These MDI solutions may help solve the symptoms of master data problems, but they rarely address the factors that cause them.

Enterprise MDM is the answer to solving master data issues

An enterprise MDM solution involves business users in the MDM process, and extends MDI by adding capabilities to manage, track and audit constantly changing master data and metadata. Enterprise MDM is a separate component that manages the system of record and system of entry on behalf of other IT components. It works in conjunction with, and supplies the master data to, business transaction and business intelligence applications. The master metadata can also be used as a source

of information for helping drive the design of the data warehousing environment that forms the underpinning of business intelligence application processing.

Full enterprise MDM, however, is a long-term project that organizations should evolve to using the four approaches described in this paper.

THE SAP NETWEAVER MDM SOLUTION

SAP MDM is an integrated component of SAP NetWeaver

SAP master data management is a component of SAP's NetWeaver application platform, which provides a complete environment for designing, building, integrating and deploying business processes. SAP NetWeaver provides the technology foundation for SAP, SAP partner and custom-built applications. Key components of SAP NetWeaver are outlined below.

- The **SAP NetWeaver Portal** provides users with single sign-on and a single view of enterprise-wide information and applications. SAP NetWeaver Portal works in conjunction with the SAP NetWeaver Collaboration and Knowledge Management capabilities. The Collaboration facility includes shared e-mails, calendars, and threaded discussions, and has a shared document store. Knowledge Management provides search tools, taxonomy development, content management, publishing, and workflow management. SAP NetWeaver Portal can also be used with SAP Multi-channel Access and the SAP NetWeaver Mobile Infrastructure to deliver information to mobile users.
- **SAP NetWeaver Business Intelligence (BI)** gives organizations the ability to identify, integrate, report on and analyze disparate information from heterogeneous sources. SAP NetWeaver BI includes a data integration toolset that supports both SAP and non-SAP data in addition to standard file formats such as XML. **SAP NetWeaver Master Data Management (MDM)** extends these data integration capabilities by providing a facility for consolidating, storing, and augmenting master data such as customer information. SAP NetWeaver BI is tightly integrated with other SAP NetWeaver components, including the SAP NetWeaver Portal, and the Collaboration and Knowledge Management components
- The **BI Integrated Planning** capability extends the SAP NetWeaver BI environment with a tactical and operational business planning, budgeting and forecasting system. This capability brings planning, forecasting and budgeting together with BI monitoring, reporting and analysis.
- Process and application integration is provided in SAP NetWeaver by a **Business Process Management** capability and the **SAP NetWeaver Exchange Infrastructure (SAP XI)**, which include facilities for describing software components, interfaces, mappings and routing rules, and for executing business processes. SAP XI supports Web services, and, like other SAP NetWeaver components, runs on an SAP-supplied J2EE-compliant Application Server.

SAP SUPPORT FOR MASTER DATA MANAGEMENT

SAP NetWeaver MDM provides an MDM capability that enables organizations to evolve to a full enterprise MDM environment using the four approaches outlined in this paper.

SAP MDM supports all four approaches to MDM outlined in this paper

Using SAP-supplied and user-developed data models, SAP MDM can **consolidate** master data in an MDS from both SAP and non-SAP data sources. The consolidation process supports de-duplication and normalization, ID mapping, matching and merging, staging, change tracking, and interactive data quality analysis.

In addition to master data consolidation, SAP MDM also provides master data **propagation** via a distribution model that can update and synchronize master data in SAP and non-SAP systems. It also incorporates data workflow and multi-tiered role-based models for managing master data updates.

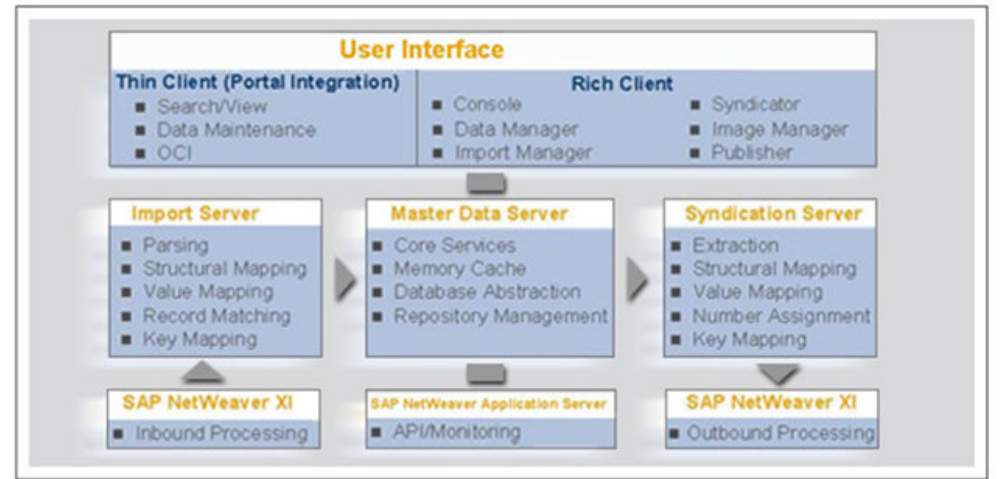
SAP MDM can also be used to **centralize** both the system of record and the system of entry for enterprise-wide master data. Users can manage master data through a supplied rich-client application. Master data can be propagated to other systems in XML format as required. It can also be extracted to a data warehouse within the SAP NetWeaver platform for use in SAP analytical applications.

SAP MDM also supports master data embedded in unstructured content

One important feature of SAP MDM is that it supports not only master data stored in structured files and databases, but also master data embedded in unstructured business content (PDF, audio, video, for example). For master data output, SAP MDM has integrated publishing capabilities that enable both electronic and printed output for the Web environment, CD-ROM, and paper publishing. These features could be used to publish product catalogs, for example. The output features can be used in conjunction with the Adobe InDesign and QuarkXpress publishing tools.

Figure 5 illustrates the main architecture of SAP NetWeaver MDM.

Figure 5.
SAP NetWeaver MDM



SAP MDM integrates with SAP NetWeaver BI, SAP NetWeaver XI, and SAP NetWeaver Portal, while also offering a modular approach for use as a stand-alone engine. SAP NetWeaver MDM can be used either in a full administrative environment (rich front-end clients), and/or through the SAP NetWeaver Portal (thin client).

Scenarios provide pre-packaged MDM solutions

In addition to providing the infrastructure for deploying an enterprise MDM environment, SAP MDM also provides pre-packaged IT and business scenarios. These scenarios contain pre-built data models, mapping, reports, workflows, and portal interfaces. Scenarios available with use with SAP MDM support both product and customer master data management. At the time of writing this paper, SAP had over 230 customers of its MDM solution in 27 vertical industries.

USING SAP NETWEAVER MDM

As already mentioned above, SAP MDM can support any combination of the four approaches outlined in this paper for deploying MDM solutions. A customer can, for example, begin with a consolidation approach, add master data propagation, and then finally move to a full centralized enterprise MDM environment.

An example to how to evolve to an SAP MDM environment was presented by Nortel at last year's SAP Sapphire conference in Orlando. Nortel is a leading communications provider that operates in some 150 countries and has approximately 30,000 employees worldwide.

Nortel implemented MDM as a part of its ERP and Web services project

Nortel implemented SAP MDM as a part of its project to move towards the use of mySAP ERP. The ERP project in turn was a component in Nortel's migration to a Web Services orientated architecture involving products from both Microsoft (BizTalk, for example) and SAP. The first step in the project in 2004 and 2005 was to eliminate some 150+ interfaces and applications. The ongoing project in 2006 and 2007 is to eliminate a further 300+ interfaces and applications.

Nortel initially developed its own MDM applications, but with the move to SAP MDM that effort has been abandoned. SAP implementation was carried out in three phases.

Phases 1 and 2 synchronize master data in multiple systems of entry

Phase 1 of the project used SAP MDM to propagate master data (customer and products, for example) between SAP applications, legacy applications, SAP BW, and business partner systems. Propagation was done using SAP XI and FTP. Phase 2 of the project extends the amount of master data propagated and the number of target master data applications.

Phase 3 deploys a full enterprise MDM system

The third phase of the project, which is currently under development, adds additional master data (suppliers and people) and deploys a full enterprise MDM environment where, except for three custom-built and packaged legacy applications, all master data will be maintained directly in the central master data store. This solution will enable master data management in some eight systems of entry to be retired. The enterprise MDM system will still be used to propagate master data to SAP BW and partner applications.

The scenario used by Nortel demonstrates the correct way to implement an MDM system, which is to develop a long-term strategy to move toward enterprise MDM and implement the strategy in small incremental steps.

About BI Research

BI Research is a research and consulting company whose goal is to help companies understand and exploit new developments in business intelligence and business integration. When combined, business intelligence and business integration enable an organization to become a smart and agile business.

BI Research

Post Office Box 398

Ashland, OR 97520

Telephone: (541)-552-9126

Internet URL: www.bi-research.com

E-mail: info@bi-research.com

Using Master Data in Business Intelligence

Version 1, March 2007

Copyright © 2007 by BI Research

All rights reserved