



Master Data Management from a Technical Perspective

Writer: Tyler Graham

Technical Reviewer: Roger Wolter

Published: April 2010

Applies to: SQL Server 2008 R2 Master Data Services

Summary: This article provides a technology-centric view of master data management concepts, including possible implementation techniques. After reading this paper, information technology professionals should be able to accurately define the master data management problem for their organizations and have a vision for what the solution may be. They should be able to engage the business and recognize the proper role that IT must play in this process long term.

This article is part of a series called “Organizational Approaches to Master Data Management.” For more articles in this series, see <http://go.microsoft.com/fwlink/?LinkId=187888>.

This document is provided “as-is”. Information and views expressed in this document, including URL and other Internet Web site references, may change without notice. You bear the risk of using it.

This document does not provide you with any legal rights to any intellectual property in any Microsoft product. You may copy and use this document for your internal, reference purposes.

© 2010 Microsoft Corporation. All rights reserved.

Introduction

There are four techniques for implementing master data management within an organization. These methods differ in the amount of control they exert over the master data they manage. All of these techniques rely on reliable data integration solutions.

Common Master Data Management Terminology

When discussing master data management and specific architectures used in master data management, there are a number of common acronyms or system names that should be identified and defined.

Acronym	Long Name	Description
SOE	System of Entry	Primary point of data entry. This may be direct entry or through services that update the data in virtual real time.
SOR	System of Record	Most, if not all systems, receive their data from this source. When conflicts arise, this system is considered primary.
MDIS	Master Data Integration Services	Data cleansing and integration processes that provide automated methods for some of the following activities: segmentation, aggregation, transformations, match/merge, and grouping.
IM	Identity Mapping	To eliminate repeated integration, identity maps should be used to manage surrogate key relationships. These may be one-to-one or one-to-many.
BI	Business Intelligence	Technology and applications dedicated to the analysis and presentation of business information.
CRM	Customer Relationship Management System	Customer management software.
ERP	Enterprise Resource Planning System	Software system that serves all areas of a business enterprise.
DW	Data Warehouse	Centralized repository of electronically stored data, used to facilitate reporting.

Table 1: Master data management acronyms

Master data registry implementations

In registry implementations, each system remains in control of its own data. All system data records are mapped in the master data registry. Data maps show the relationships between the keys of two different systems. These keys can be mapped in two different ways:

- One to one: Every record in the main system will have only one corresponding record in the secondary system.
- One to many: Every record in the main system will have one or more corresponding records in the secondary system.

These mappings provide the data integration applications a reliable way to compare related items. At any time, different systems can be compared and cleansed. Although this technique provides important mapping information to the organization, any new items in any system will lead to data inconsistencies within the solution requiring a very complex data management story.

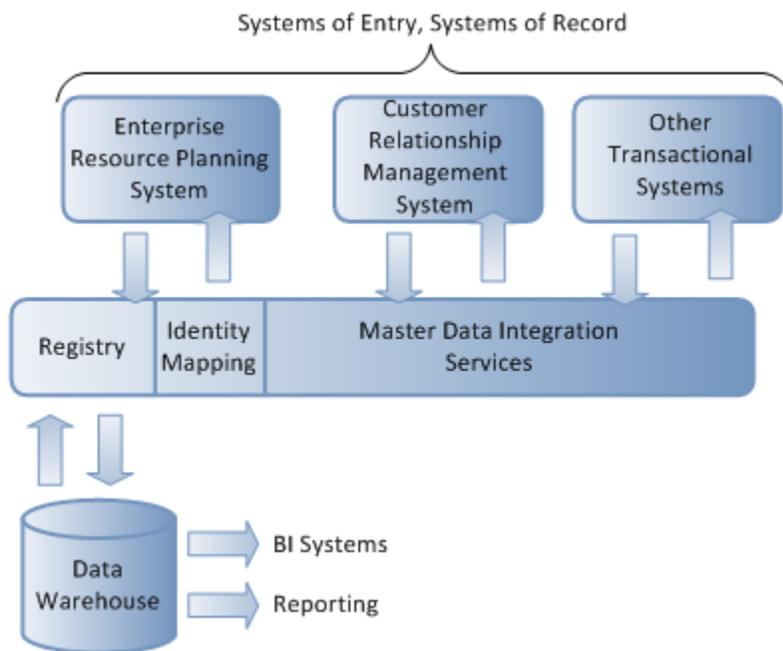


Figure 1: Master data registry implementation

Data aggregation implementations

Data aggregation implementations involve the creation of a new system that is neither the system of entry nor the system of record but a downstream system used to aggregate the attributes of multiple systems and pass them to lower level subscribing systems.

It is common for initial master data management implementations to implement this type of watered-down approach. In many cases the applications and systems used for this technique are identical to the more advanced techniques listed below. The major factor missing in the solution is control. It is very difficult for an initial master data management project to get all of the necessary stakeholders to

relinquish control of their data to a new system immediately. Another system, usually the most critical business transaction application, remains the system of record and system of entry. Integration processes transfer data from this initial source to the master data management application. This master data may be enhanced by the master data management application itself, but a majority of the important information is still imported from the more entrenched systems. These systems may be more entrenched due to a number of factors, including importance to the business, amount of time spent cleansing the data, current process, or even perception of value. Data will then be propagated to other systems. Data controls will be limited because the source system will not be designed to account for any other system's requirements.

Despite the limitations inherent with this method, this is actually a good method for bringing quick wins to an organization with master data management. Many stakeholders are reluctant to give up the security of their traditional data management processes. By showing early benefits and demonstrating application reliability, an organization can develop trust in the master data management application as the system of record and system of entry.

Using this technique for the initial implementation, risk to the mission-critical application can be mitigated effectively. Less critical applications can begin to source their master data from the master data management application, solving any integration issues that arise without major ramification to the organization. Integration processes can be tested and modified in an iterative fashion.

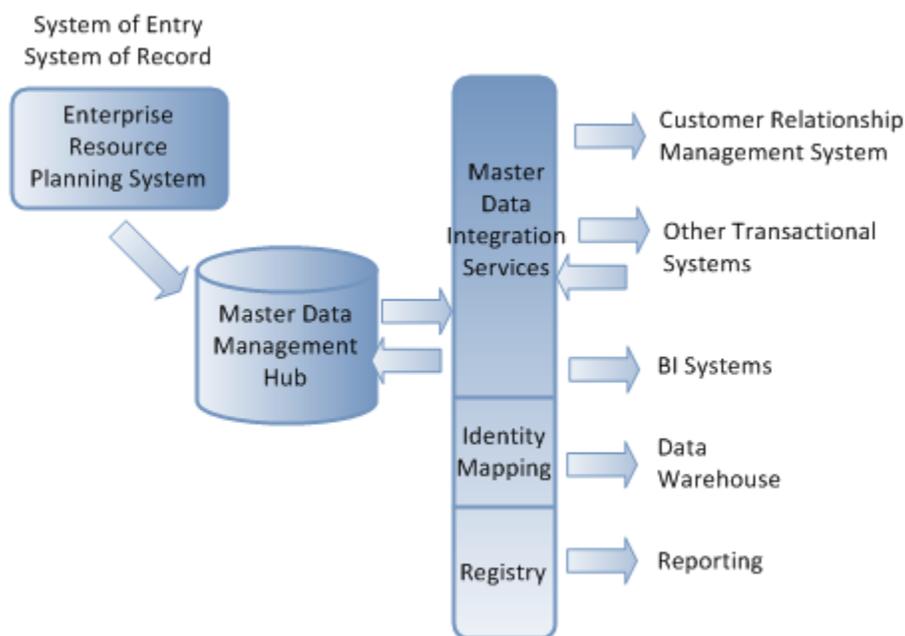


Figure 2: Data aggregation implementation

System-of-record-only implementations

These implementations give complete control of the master data sets to the master data management application. Other systems provide the initial data to be imported into the master data management

system, but, unlike the data aggregation solution, the flow of data from this System of Entry is bidirectional. New records are transferred into the master data management application for integration. Any discrepancies in the data defer to the master data management application, which is the system of record.

These implementations still require a degree of data integration and ongoing cleansing as elements may come from both the source system and the master data management application. Also, many times this system of entry only has the ability to detect data issues directly related to the initial use. For instance, any customer information that is not stored in the CRM solution will not be available to determine complete data quality.

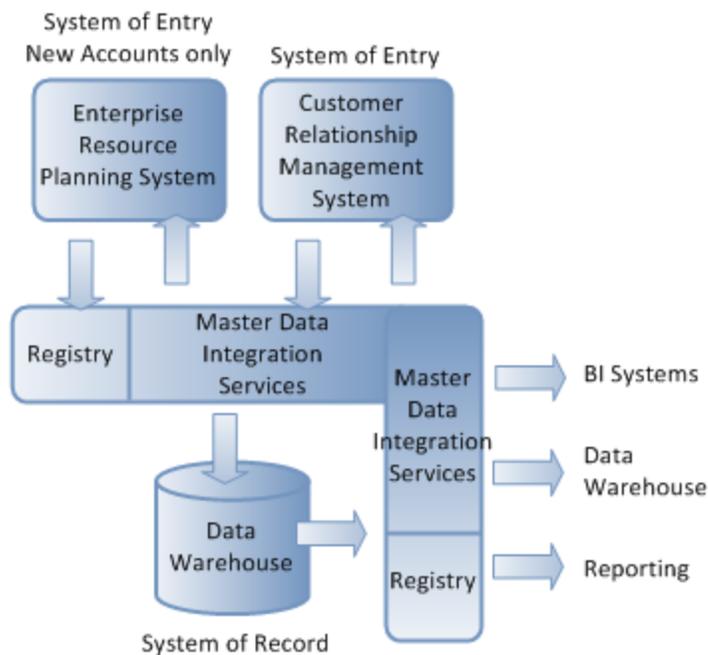


Figure 3: System-of-record only (hub) implementation

Complete enterprise master data management implementations

Complete master data management solutions require the entire lifecycle of the master entities to be managed from within the master data management solution. Controlling the entry of the master data allows the enterprise application to proactively manage the quality of the data. Although an enterprise implementation will be both the system of entry and system of record for all master data entities, it may still require mapping data to other applications.

It is not realistic for an organization to get all of their systems to use the exact same set of data. Some transformations will still be required to run the process systems. This does not mean that every defining characteristic of an entity is managed within the solution. Those defining attributes unique to an organization's system operation should be managed within the source system where they have relevance. The enterprise solution should provide a broad range of entry points to be a viable option as the system of entry.

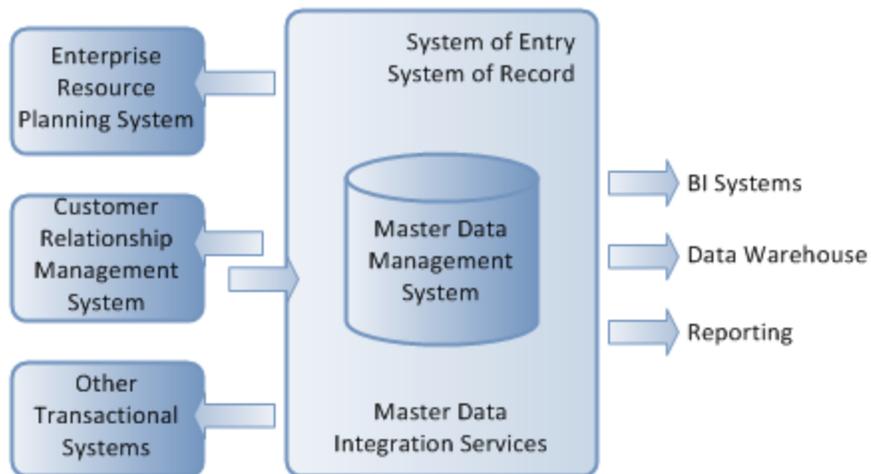


Figure 4: Enterprise implementation

Conclusion

There are a number of ways to architect a successful master data management solution. To determine the best architecture for your domain and organization, it is imperative to understand the benefits and drawbacks of each type of solution.

Feedback

Did this paper help you? Please give us your feedback. Tell us on a scale of 1 (poor) to 5 (excellent), how would you rate this paper and why have you given it this rating? For example:

- Are you rating it high due to having good examples, excellent screen shots, clear writing, or another reason?
- Are you rating it low due to poor examples, fuzzy screen shots, or unclear writing?

This feedback will help us improve the quality of white papers we release. [Send feedback](#).