

Master Data and SOA in a distributed organisation

Thesis proposal

Jonas Eyob

Department of Computer Science

Royal Institute of Technology

jeyob@kth.se

February 19, 2014

1 Introduction

In today's rapidly changing and fast paced business environment enterprise agility has become an increasingly important factor. To achieve this, enterprises implement and maintain information systems that will make them agile enough to take part in emerging business opportunities and/or help them facilitate changing business requirements. Maintaining key organisational data has always been an important, e.g., knowing who the customers and suppliers are and which projects they are involved in are essential to be able to strengthen the customer relationship or rationalise decisions affecting the organisation, just to name a few. A commonly used term in this context is master data. Several criteria and guidelines for deciding what characterises master data exist, but in a general sense master data encompass data that more or less defines the enterprises i.e., those data that are critical to the organisation such as projects, customers, suppliers, products etc. Thus, it's not hard to understand why master data is the most valuable data the organisation owns, and the subject for many acquisitions.

Having a consistent and enterprise-wide understanding of what defines, for example, a customer or a project is essential as this both helps prevent bad things from happening - such as customers receiving bills from the different business units instead of a consolidated bill that both saves the issuing company money and improves customer satisfaction - and to seize beneficial business opportunities by analysing the customers involved in the projects and with that anticipating latent needs of the customer.

In many organisations today, however, it is the case that master data resides in a number of overlapping systems across the enterprise. This quickly becomes cumbersome as increasing complexity (due to change, growth, mergers & acquisitions or technology transformations) has made it hard for enterprises maintain consistent master data across the enterprise.

1.1 NCC

NCC is one of the leading construction and property development companies in the Nordic region with offices located in Sweden, Norway, Finland, Russia and Denmark. The NCC group had sales worth of SEK 57 Billion in 2013, with approximately 18 000 employees. Recent initiatives by NCC group include the establishment of an group IT function on corporate level to enforce a more holistic and unified view of how IT supports the enterprise. Each subsidiary NCC office around the Nordic region had up until recently had their own, more or less independently run, IT function which have led to locally optimal IT solution architectures with the drawbacks of having inconsistent definition master data across the different business units. For example, NCC which is a company highly centred around projects do not have an enterprise-wide understanding of what constitutes a project.

1.2 Problem formulation

With the absence of enterprise-wide set-up to handle master data at NCC this has led to an inconsistent definition of what, for example, constitutes a project or even which customers are shared across the various business units. Imagine, if a large customer of NCC would be in business with several of NCC's business units, this is indeed an important customer which should be treated so accordingly. However, since business lacks the necessary mechanisms to realise this, this customer

is likely to be viewed as a small customer. Not only can this lead to missed business opportunities but also the chance of permanently losing an important customer.

Consequently, the aim of this master thesis project is to look at how to best approach the problem of establishing a master data management solution at NCC. Furthermore, as the NCC group is a fairly decentralized organisation, with its master data residing in different subsidiary organisations and with a organisational structure likely to remain we need to investigate how we, despite this fact, can establish a master data management solution that addresses aspects such as data inconsistency and quality.

Moreover, we will also look at how we can assist access to master data through services using a service-oriented architecture (SOA) approach.

1.3 Delimitations

Due to the limited time frame for this project we will solely focus on a subset of master data - project data. Nevertheless, we believe that the learning outcomes from this project will greatly be of assistance stretching this work to remaining master data.

2 Methodology

Conceptually this project can be divided up into three parts *theory, current state analysis, future state analysis and implementation roadmap*.

2.1 Current state analysis

One of the very first things that we need to address for this project is to concretise the scope of the analysis. For example, there are many dimensions, with respect to master data management architecture, that must first be agreed upon as they may greatly impact the project delivery. In short, this means understanding the areas to focus our analysis around; some of these include, but are not limited to, data quality, access rights/security, data accuracy, data synchronization, timeliness, interoperability and governance (both SOA and data). Perhaps the primary objective of this project is to compare different possible technical architectures and evaluate them based on one or several of the aforementioned attributes.

Another important part during this phase will be to gather and build the necessary theoretical foundation for the area of MDM and SOA. This could both be achieved through literature studies but also by meeting with consultancy firms to understand how this is solved in the industry. And with the limited time for executing this project, its preferred that we also start to identify where the master data resides and which business processes actually accesses the master data. Furthermore, to understand how the master data differs between the business units we will need to identify their respective information models; focusing primarily on project data. By doing this we hope to understand how the master data is altered and how their information models (or master data definition) differs.

2.2 Future state analysis

Based on what we learned from the previous phases we will look at some different target architectures and analyse them with respect to areas of consideration (e.g., data accuracy, consistency, timeliness, synchronization, quality, security etc). The biggest challenge during this stage will be to compare the different information models and decide on which how an improved model could look like.