Integrated Data Architecture for Business

Richard Kumaradjaja
Renaissance Consulting, Indonesia

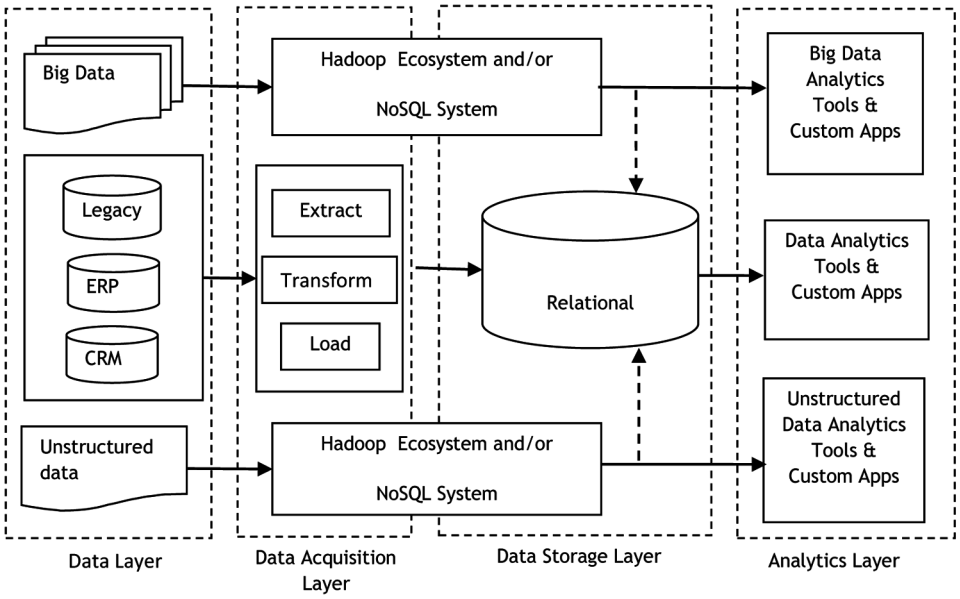
INTRODUCTION

According to (Kosala & Kumaradjaja, 2014), to achieve a better analytics capability business organizations must utilize three Business Analytics components altogether so that a 360 degree analysis or a more comprehensive view of data can be implemented. The three Business Analytics components are: Big Data Analytics Tools and custom applications, Data Analytics Tools and custom applications, Unstructured Data Analytics Tools and custom applications form the Analytics Layer of the Big Data Analytics within the Business Intelligence and Business Analytics framework proposed by (Kosala & Kumaradjaja, 2014). For convenience, the structure of the Big Data Analytics framework is shown in Figure 1.

As stated by (Kosala & Kumaradjaja, 2014), the integration methodology of the three analytics components in Figure 1 and how they can be utilized together to get a 360 degree analysis view of data could represent a major research effort in the next few years. One of those research areas is about the data integration framework between unstructured Hadoop or NoSQL systems to create a structured OLAP data structure, and its real business case studies.

The author believes that the data integration framework must consider all layers of the Big Data Analytics framework shown in Figure 1, namely, Data Layer, Data Acquisition Layer, Data Storage Layer and Analytics Layer. Therefore, the proposed data integration framework must consider the entire Big Data Analytics ecosystem as one integrated data architecture.

Figure 1. Big data analytics within the business intelligence and business analytics framework



DOI: 10.4018/978-1-5225-2255-3.ch074

The purpose of this article is to clarify and show the concept of an integrated data architecture as a foundation of a modern data architecture which has the potential of addressing and solving critical issues related to big data analytics implementation in business organizations. Eventually, the goal of this article is to improve the likelihood of successful big data analytics implementation in businesses.

BACKGROUND

In their article (Kosala & Kumaradjaja, 2014) has pointed to data integration framework between unstructured Hadoop or NoSQL systems to create a structured OLAP data structure as one of the promising future research areas in big data analytics. The issue, however, is that there are still some confusions about data integration as mentioned by (Russom, 2008). In the next section, the concept of data integration architecture will be clarified as well as the detailed explanation of the proposed integrated data architecture for business organizations.

MAIN FOCUS OF THE ARTICLE

The main focus of this article is to address the issues of data integration from the architectural point of view. Addressing the isues of data integration from the architectural point of view will lead to a better understanding of the current situation and better able to construct proposed solutions to those issues since architectural approach can give us a holistic and comprehensive view of the problems.

In this section, the data integration issues related to Big Data Analytics adoption will be discussed, and an integrated data architecture for implementing Big Data Analytics will be proposed.

Issues, Controversies, Problems

There are some issues and misunderstandings when we talk about data integration. (Russom, 2008) has stated that to a lot of people the term data integration architecture sounds like an oxymoron because they do not think that data integration has its own architecture. For example, some data warehouse professionals still use the methods and practices of the 1990s when data integration was subsumed into the larger data warehouse architecture. He also observed that many data integration specialists still cling to poor practice that's inherently anti-architectural by building one independent interface at a time. These misunderstandings are further exacerbated by the belief that using a vendor product for data integration automatically assures architecture.

The problems with the above misconception are:

- We can not address how architecture affects data integration's scalability, staffing, cost, and ability to support real time, master data management, SOA.
- We can not address interoperability with related integration and quality tools.

In other words, with the absence of data integration architecture, we will not be able to construct an integrated data architecture which is strongly independent, future-facing, productive, scalable and interoperable (Russom, 2008).

Why have this misconception and confusion about integrated data architecture persisted? The complexity and perception about the role of enterprise data warehouse are the major reasons behind the misconception.

Today, in most business organizations, data flow from diverse source systems such as ERP, CRM and supply chain applications where most enterprise data originates as well as non-traditional