

# Key organizational factors in data warehouse architecture selection

Ravi Kumar Choudhary

---

## ABSTRACT

*Deciding the most suitable architecture is the most crucial activity in the Data warehouse life cycle. Architecture is the key factor in setting up the abilities and the limitations of a data warehouse. This article was conducted to (1) better understand the factors that influence the selection of data warehouse architecture and (2) the success of the various architectures. The academic and data warehousing literature and industry experts were used to identify architecture selection factors and success measures and then to create questions for a Web-based survey that was used to collect data from many companies about the respondents, their companies, their data warehouses, the architectures they use, and the success of their architectures. The study findings provide interesting and useful things about topics of long-standing importance to the data-warehousing field.*

---

**Keywords:** Datawarehouse, Architecture, Data Mining

---

## 1. Introduction

Over the past decade, companies have spent billions of dollars on data marts and warehouses. From their experiences, a substantial body of knowledge has been created. We know, for example, the importance of thoroughly understanding source systems before building, starting with only a few subject areas or business processes but having an enterprise-wide goal in mind, and giving end users data access tools and applications that are appropriate for their needs.

There is one area, however, that still causes considerable confusion and disagreement: *Which architecture to use?* There are multiple options. The most common is the hub and spoke architecture (i.e., centralized data warehouse with dependent data marts) that is advocated by Bill Inmon, who is commonly referred to as “the father of data warehousing.”<sup>[1]</sup> Inmon refers to this architecture as the Corporate Information Factory (CIF). Another prevalent choice is the data mart bus architecture with linked dimensional data marts (bus

architecture), advocated by Ralph Kimball, the other preeminent figure in data warehousing.<sup>[2]</sup> Each has strong proponents.

Considering the importance of the choice of architecture, there is surprisingly little articles on the topic. The literature tends to either discuss the architectures, provide case study examples, or present survey data about the popularity of the various options.

## 2. Studying the Architectures

A three-phase study was conducted to provide answers to two research questions:

1. What factors lead companies to select a particular architecture and
2. How successful are the various architectures?

The answers to these questions are important to companies, vendors, and consultants.

The study's first phase identified the factors that potentially affect the selection of a data warehouse architecture and metrics to use in assessing the