Using Agile Methodology in Data Warehouse

Rahul Gupta

Master of Business Administration (Computer Information Systems), University of Rochester, NY

Abstract- In recent years, Agile Methodology for software development has been embraced by many organizations where traditional waterfall approach has failed. Agile method promises higher software quality, reduced risk and more cost savings as compared to waterfall. Many organizations using Agile have realized these benefits for online transactional processing (OLTP) applications and have recently tried to apply this method and gain benefits for online analytical processing (OLAP) or data warehouse applications. In this paper, we aim to explain the challenges faced during development in datwarehousing projects using waterfall approach and how best the agile methodology can be applied for these types of projects.

Index Terms- OLTP, OLAP, Agile, Waterfall,Data Warehouse

I. INTRODUCTION

Data warehouse forms an integrated environment where data from disparate systems is bought together and presented in a consistent matter. Data Warehouse has several components such as data in different format from disparate sources, database, Extract Transform and Load (ETL) processes, Dashboards etc. Data integration using ETL from disparate sources is often the most time consuming and expensive tasks in developing a data warehouse and accounts for more than half of the project budget. Also the requirements in data warehouse project are subject to constant change by the business users making it harder to follow the traditional waterfall approach.

In waterfall method, the business value is not realized till the major phases of the project Requirement Analysis, Design and development are completed which makes the project more risky and likely to fail. As compared to Agile where the business benefit is realized on regular basis using 4-6 week of iterative development cycles called Sprints. At the end of each Sprint, development team is required to produce business value in terms of tangible work product. In case of data warehousing projects, these sprints can be useful in design and development of complex ETL integration

II. RESEARCH ELABORATIONS

For organization, which are starting to implement agile methodology and have only used waterfall methodology in the past, it is recommended to use agile first for online transactional processing projects. Once the organization gets used to agile environment, it is recommended to deploy agile on Data Warehousing projects which are generally more complex in nature.

Data warehousing projects tend to be longer and often the business value is not realized as by the time when the requirements are delivered business has moved on to a different set of requirements due to changing market conditions. To address problems like these, it is often recommended to take agile approach where business value can be realized sooner and is visible to business on recurring basis.

III. RESULT FINDINGS

The result findings from different data warehousing projects using Agile methodology have shown to deliver better value to business especially when the scope of project is uncertain.

IV. CONCLUSION

In conclusion using agile methodology for data warehousing projects, IT teams can not only achieve increased flexibility to address new requirements but also improve the quality and response time to new requirements. Agile offers value to business on regular basis and helps in staying on track especially for data warehouse projects which are mostly large and complex.

AUTHORS

First Author –Rahul Gupta, Master of Business Administration (MBA), Simon School, University of Rochester,mitarahul79@gmail.com.

Correspondence Author –Rahul Gupta, Master of Business Administration (MBA), Simon School, University of Rochester,mitarahul79@gmail.com,617-416-9502.