



Volume - VIII

Advanced Engineering Research and Applications

Editor-in-Chief:
Hongseok Choi

RIP Research India Publications

CHAPTER - 3

A TECHNICAL COMPREHENSIVE SURVEY OF ETL TOOLS

Vaishali A. Kherdekar

*Assistant Professor, MIT Art's Commerce & Science College (MIT ACSC), Alandi,
Affiliated to Savitribai Phule Pune University Pune, Maharashtra State, India.*

Pravin S. Metkewar

*Associate Professor, Symbiosis Institute of Computer Studies and Research(SICSR),
Affiliated to Symbiosis International University(SIU), PUNE-411016, Maharashtra State,
India.*

Abstract

In modern days ETL tools are very useful in data integration and data warehousing. Input is given to the data warehouse through ETL. ETL means Extraction, Transformation and Loading. ETL tools transfer data from one source system to another source system.

As these tools are mainly used in Business Intelligence and Data Warehousing, there is lot of space for their progress. There are lots of ETL tools available in the market varying from version to version to stay proficient against other tools. Each and every tool has its own features and limitations. In this paper we have carried out technical survey of existing ETL tools and benchmarking of these tools has been performed by considering certain parameters including scalability, reusability, interoperability, support to big data, parallelism, usability, flexibility etc. Finally, problems and challenges of ETL tools have been discussed thoroughly and its state of the art is summarized.

Keywords : ETL tools, Data warehouse

INTRODUCTION

Now a day's data warehouse is used in industry to maintain optimize model of data for further mining and usage and also for report generation. By using data warehouse one can maintain historical data and used it in decision support system. To construct data warehouse model, ETL tools is being used. ETL tools act as basis for construction of data warehouse. Input is given to the data warehouse through ETL. ETL stands for Extraction, Transformation and Loading. In extraction phase, data is extracted from various