Abstract

Generally, software projects’ outcomes will give us various aspects of quality parameters. In such cases, empirical studies with prototyping exercises are well suited to analyze/understand the system. ETL (Extraction-Transformation-Loading) is the software responsible for extracting data, cleaning, transforming and loading the data into a data warehouse. ETL is a large software system. The performance of the decision support system depends on the data warehouse that it uses. ETL tools play a major role in building the data warehouse; these tools need to have good performance in order to improve the performance of the whole system. An experimental study is conducted to analyze the performance of the ETL tool. Two ETL tools are considered; one with integrated security and another without integrated security. The time for data extraction in different environments is recorded. Further, regression analysis is done on the experimental data and observed the behavior of the tools and developed the empirical models. Both tools have shown the same behavior in performance for different extraction data sizes.
Empirical Models for the Performance of ETL Processes

- Stefano Rizzi, Alberto Abellón, Jens Lechtenborger and Juan Trujillo, 2006 “Research in Data Warehouse Modeling and Design: Dead or Alive?”, Proc. DOLAP, pp. 6, 3-10.
Empirical Models for the Performance of ETL Processes


Index Terms

Computer Science
Software Engineering

Keywords
Regression Analysis
Integrated Security
Secure ETL
Performance Analysis
Experimental Study.