Abstract

A conventional data warehouse use to produce summary from an organization information system in a long time period. This condition will make the management unable to get the most up to date data every time it needed. Therefore a nearly real time data warehouse which will manage the ETL process with a more compact data and a shorter period is needed. The design of nearly real time data warehouse in this research is implemented in two steps. The first step is done by data collection technique modeling to make a more compact ETL data managed. This step is done by putting the staging area on an Online Transactional Processing (OLTP). It can minimize the failure of data movement process from the OLTP to the staging area. Besides that, the CDC method is also had applied on the OLTP. This method will be implemented with a trigger active database. The trigger will capture of the data changing on the
OLTP, transform it and then load it to the staging area in one time. The second step is the synchronization process of the data movement from the staging area to the nearly real time data warehouse. This process is done by mapping the movement which is ran by the SQL Yog. The mapping result will accomplished by the windows task scheduler.

References

- Jie Song; Yubin Bao; Jingang Shi; 2010, A Triggering and Scheduling Approach for ETL. Computer and Information Technology (CIT), 2010 IEEE 10th International Conference on , Page(s): 91 – 98.

Index Terms
Keywords
Nearly real time data warehouse  Change Data Capture  Surrogate key  Trigger