

Abstract

The utilizing of data warehouse became widespread in organizations or companies. However, on-time or near real-time data warehousing are also increasingly required to support rapid decision making in business process. CDC (Change Data Capture) technique can bring near real-time data warehousing into reality because the refreshment is done by capturing only the change data, differ from conventional techniques that the refreshment process requires expensive memory load, a very big delay, and has many disadvantages. CDC consists of several methods, one of which is Asynchronous HotLog.

Data warehouse takes the information from source database. Asynchronous Hot-Log CDC will capture the information or the change data in source database by utilizing redo log which is delivered by the oracle streams and then stored in staging database. Staging database on Asynchronous HotLog CDC is reside on the source database itself. The configuration of Asynchronous HotLog CDC involves some components setting. There are specific components can affect its performance. This final project will not only perform the step by step on implementing Asynchronous HotLog CDC configuration but also make the performance analysis and the factors analysis that can influence its performance so it could minimize the delay and increasingly getting to near real-time data warehouse.

Keywords: data warehouse, near real-time data warehouse, refreshment, change data capture, Asynchronous HotLog.