

Abstrak

Data warehouse hadir dan berkembang sebagai *repository* data yang saat ini sangat dibutuhkan oleh perusahaan. Pengelolaan data yang sangat banyak dan tersebar serta informasi yang dapat digali dari *historical* data menjadi tujuan dari pembuatan *data warehouse*. Sebagai sebuah *repository* data, *data warehouse* tentunya akan selalu berinteraksi dengan data dalam jumlah yang besar, menangani transaksi data dalam jumlah yang besar dan intensitas yang tinggi. Oleh karena itu, *data warehouse* dituntut untuk mampu menangani setiap transaksi yang terjadi di operasional *database* dan menangkap perubahan data dari setiap transaksi dan menyimpannya sebagai sebuah *historical* data dengan *delay* sekecil mungkin. Untuk menangani masalah tersebut, ada sebuah proses dalam *data warehouse* yang disebut *change data capture* (CDC). Proses CDC akan mengidentifikasi dan memproses setiap perubahan yang terjadi di operasional *database* kemudian dikirimkan dan disimpan ditempat penyimpanan sementara yang disebut *staging database*. Dengan banyaknya data dan transaksi yang terjadi setiap harinya, diperlukan performansi yang bagus dari proses CDC. Untuk mendapatkan performansi yang bagus, tentunya diperlukan kinerja yang bagus dari setiap komponen yang membangun proses CDC. Untuk itu, dalam tugas akhir ini, dilakukan implementasi dan pengujian terhadap proses CDC untuk melihat performansi yang dihasilkan berdasarkan waktu eksekusi CDC. Lalu kemudian di analisis komponen apa saja yang berperan dalam CDC dan bagaimana kinerja dari setiap komponen tersebut. Dengan analisis performansi dan mengetahui faktor-faktor yang mempengaruhi performansi, maka proses *waiting* dalam CDC dapat dihilangkan atau dikurangi intensitas dan waktunya. Sehingga proses CDC dapat menghasilkan *delay* sekecil mungkin dan menjadi *real time data warehouse*.

Kata kunci: *data warehouse, change data capture, historical data, performansi, delay*