

Abstrak

Gangguan terhadap kualitas data pada proses *data warehouse* menjadi hal penting untuk diperhatikan. Gangguan yang terjadi disebabkan proses *capture data* saat perubahan data yang terjadi tidak sempurna dan tidak tersimpan pada *data warehouse* secara *real time* menyebabkan adanya data yang hilang. Pada dinas kependudukan dan catatan sipil memiliki gangguan yang serupa. Gangguan tersebut menyebabkan NIK ganda terjadi di beberapa wilayah Indonesia. Selain itu, *Data warehouse* secara kontinu menerima permintaan *query* dari user secara bersamaan. permintaan *query* secara bersamaan memungkinkan terjadinya *resource competition* dan *high load*. Sedangkan, user membutuhkan *response time* yang singkat. Solusi dari permasalahannya yaitu dengan menggunakan *log based change data capture* dan *scheduling mechanism* menggunakan *requirement-based query and update scheduling*.

Penelitian ini akan mencari solusi dari kualitas data dan permintaan query dari user dengan studi kasus dinas kependudukan dan catatan sipil. Tahapan yang dilakukan pada penelitian ini yaitu membuat skema *data warehouse* menggunakan *star schema* berdasarkan proses bisnis yang terdapat di dinas kependudukan dan catatan sipil. Lalu, *data source* akan melalui proses *extract* menggunakan *log based change data capture*. Proses selanjutnya yaitu transform dimana data yang tidak memiliki nilai akan diinisialisasi dengan “null”. Kemudian, data akan dimasukkan ke dalam *data warehouse* melalui proses *loading*. Permintaan user yang dilakukan secara bersamaan akan dijadwalkan menggunakan *requirement-based query and update scheduling*, disingkat RQUS. RQUS memperkenalkan *acceptable response time delay* dan *acceptable result staleness* untuk memenuhi kebutuhan permintaan user. *Log based change data capture* dapat menyimpan semua perubahan data tanpa merubah struktur dari *source system* dan *requirement-based query and update scheduling* dapat menjalankan penjadwalannya sesuai dengan kebutuhan user.

Kata Kunci : *log based change data capture*, penjadwalan, *requirement query and update scheduling*, *acceptable response time delay*, *acceptable result staleness*, *real time*