

1. PENDAHULUAN

Pada bab 1 ini terdiri dari 6 sub bab diantaranya latar belakang, perumusan masalah, tujuan, batasan masalah, metode penelitian, sistematika penulisan.

1.1 Latar Belakang

Kualitas data pada proses *data warehouse* menjadi hal penting untuk diperhatikan. Karena data harus tersimpan dengan baik dan dijaga kualitasnya terlebih ketika proses *Extract, Transform, dan Loading* pada *data warehouse*. Seperti pada Dinas Kependudukan dan Catatan Sipil yaitu adanya NIK ganda. NIK ganda terjadi di beberapa wilayah Indonesia seperti Serang, Cirebon, Kabupaten Tangerang, Kota Tangerang menjadi masalah karena tidak sesuai dengan isi dari Undang-Undang nomor 24 tahun 2013. Jika terdapat masalah dalam pembuatan dokumen, akan terbentuk NIK baru pada orang yang sama sedangkan 1 orang hanya memiliki 1 NIK.

Permasalahannya terjadi karena proses *capture data* saat perubahan data itu terjadi tidak sempurna dan tidak tersimpan secara *real time* menyebabkan data menjadi hilang. Proses *capture data* terjadi pada tahap *extract*. Jika terdapat data yang hilang dari proses *extract*, maka kualitas data akan rendah. Selain itu, permasalahan lain mengenai akses bersamaan dari banyaknya permintaan user terhadap *data warehouse*. Permintaan user berupa *selection query* dan *update query*. Akibat dari permintaan secara terus menerus, mengakibatkan *resource competition* dan *high load* pada *real time data warehouse*. Sedangkan user menginginkan *response time* yang singkat dan *low result staleness* saat melakukan permintaan *query*. Sehingga, perlu dilakukan penanganan lebih khusus dengan mekanisme *change data capture* dan *scheduling mechanism*.

Salah satu proses penting dari *Real Time Data Warehouse* adalah *change data capture* pada *data extraction* [9]. *Change Data Capture* merupakan metode yang mendukung *real time captured data*. Terdapat 2 jenis *change data capture* yaitu yang *trigger based* dan *log based*. *Trigger based change data capture* mempengaruhi struktur *query* dari *source system* [9]. Sedangkan *log based change data capture* merupakan metode yang menjamin semua perubahan dari sumber data akan direfleksikan ke *data warehouse* tanpa mempengaruhi struktur dari *source system* secara *real time*.

Sedangkan *scheduling mechanism* yang digunakan yaitu *requirement-based query and update scheduling*. *requirement-based query and update scheduling* yaitu metode penjadwalan yang memenuhi kebutuhan *query* dari permintaan user. Metode ini memperkenalkan *acceptable response time delay* dan *acceptable result staleness* saat *query* dimasukkan. Untuk melakukan penjadwalan yang dapat diterima saat *query* dan *update* dilakukan secara bersamaan, penentuan nilai *acceptable response time delay* dan *acceptable result staleness* harus ditentukan dengan tepat sehingga kebutuhan user akan permintaannya dapat terpenuhi.

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

1.2 Perumusan Masalah

Dari latar belakang yang terjadi, permasalahan yang akan dirumuskan yaitu sebagai berikut :

1. Bagaimana cara dan hasil analisis penggunaan *log based change data capture* di *data extraction* pada *real time data warehouse* ?
2. Bagaimana cara dan hasil analisis saat melakukan penjadwalan yang sesuai dengan kebutuhan user yang dapat diterima dengan *response time* yang singkat dan *low result staleness* pada *real time data warehouse* ?

1.3 Tujuan

Tujuan dari pengerjaan tugas akhir ini adalah sebagai berikut :

1. mengimplementasikan *log based change data capture* dan *scheduling mechanism* menggunakan *requirement-based query and update scheduling* dari dinas kependudukan dan catatan sipil pada *real time data warehouse*.
2. Menganalisis performansi dari *log based change data capture* dan *scheduling mechanism* menggunakan *requirement-based query and update scheduling*.

1.4 Batasan Masalah

Adapun batasan masalah pada tugas akhir ini adalah sebagai berikut :

1. Data yang digunakan yaitu data dinas kependudukan dan catatan sipil republik Indonesia di wilayah kabupaten Tangerang tahun 2014-2015. Data tersebut akan digunakan pada proses *extraction* pada *Real Time Data Warehouse*.
2. Jenis file yang digunakan dari data yang digunakan yaitu berbentuk excel.
3. Sistem menggunakan 3 redo log file pada saat menjalankan aplikasi.
4. Permintaan user untuk pengujian *scheduling mechanism* yaitu berupa text file berisikan beberapa *query* yang dapat menghasilkan atau mengubah nilai tertentu di database. Jumlah *query* yang diujikan hanya sebatas pada text file tersebut.
5. Data warehouse yang dibuat diimplementasi berdasarkan produk yang dibuat oleh Dinas Kependudukan dan Catatan Sipil yaitu produk KTP el, produk Akte Kelahiran, produk Kartu Keluarga.

1.5 Metode Penelitian

Metode yang digunakan dalam penulisan ini adalah sebagai berikut :

1. Identifikasi masalah
Pada tahap ini, mempelajari masalah-masalah yang terdapat pada karya ilmiah maupun dalam kehidupan sehari-hari akan dikumpulkan menjadi poin-poin permasalahan.
2. Studi Literatur
Pada tahap ini dilakukan untuk memahami metode-metode yang digunakan dalam karya tulis dalam menyelesaikan permasalahan yang ada. Pencarian informasi dilakukan dengan membaca berbagai paper, jurnal, buku maupun studi tidak formal

dengan berdiskusi dengan orang-orang yang mempunyai minat dan latar belakang yang sama.

3. **Pemodelan sistem**

Pada tahap ini, akan dirancang model atau alur dari sistem yang akan dibuat. Hal ini bertujuan untuk memudahkan dan mencegah perubahan-perubahan yang signifikan pada saat tahap selanjutnya.

4. **Implementasi sistem**

Pada tahap ini pembuatan sistem akan dilakukan secara nyata dengan menggunakan perangkat lunak *oracle database* dan *netbeans* untuk membangun sistemnya.

5. **Pengujian sistem**

Pada tahap ini sistem yang sudah dibuat tadi diuji kemampuan dan keakuratannya. Pengujian dilakukan dengan menjalankan seluruh alur sistem yang telah dibuat dengan database yang akan diolah atau diuji.

6. **Pembuatan laporan**

Setelah hasil pengujian diperoleh, analisis dilakukan untuk menyimpulkan apakah sistem telah sesuai dengan kriteria yang ditentukan dan apakah tujuannya terealisasi atau tidak. Selain itu, sistem yang dibuat harus dievaluasi kelemahan dan kelebihan guna pengembangan penelitian lebih lanjut atau menjadi referensi untuk karya tulis lain.

1.6 Sistematika Penulisan

Buku laporan tugas akhir ini disusun menjadi lima bab, yaitu :

1. **Pendahuluan**

Pada bab ini terdiri dari latar belakang, rumusan masalah, tujuan, batasan masalah, metodologi penelitian, sistematika penulisan

2. **Landasan Teori**

Pada bab ini berisi uraian terkait teori-teori yang digunakan untuk menyelesaikan permasalahan yang ada.

3. **Analisis Perancangan dan Implementasi**

Pada bab ini menjelaskan tentang bagaimana rancangan sistem yang akan dibuat berdasarkan landasan teori yang ada untuk menyelesaikan masalah.

4. **Pengujian dan Analisis**

Pada bab ini menjelaskan bagaimana hasil pengujian terhadap sistem yang telah diimplementasikan. Kemudian hasilnya akan dilakukan analisis guna mendapatkan informasi.

5. **Kesimpulan dan Saran**

Pada bab ini menerangkan kesimpulan dari hasil penelitian yang dilakukan dan saran untuk penelitian selanjutnya.