



Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk menggubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

BAB 3

PELAKSANAAN KERJA MAGANG

3.1 Kedudukan, Koordinasi, dan Divisi

3.1.1 Divisi

Pada praktik kerja magang yang dilaksanakan penulis di PT. Putera Handal Indotama, penulis berada pada divisi *data warehouse* yang merupakan bagian dari divisi *human resource*. Divisi ini dibentuk khusus untuk menangani proyek yang berkaitan dengan *data warehouse*. Di dalam divisi ini ada dua orang dari tim PHI-Integration dan juga tiga anggota IBICC lainnya yang juga menangani proyek yang sama dengan penulis.

3.1.2 Kedudukan dan Koordinasi

Pada proyek implementasi *data warehouse* yang penulis kerjakan di PHI-Integration, penulis berkedudukan sebagai *Lead ETL Developer*. Sebagai *Lead ETL Developer* penulis memiliki beberapa tanggung jawab diantaranya adalah sebagai berikut :

1. Menerima koordinasi dan menyampaikannya kepada tim member lain
2. Melaporkan perkembangan harian kepada *Project Manager* melalui email
3. Membuat dan melakukan *test script ETL* Pentaho berdasarkan dokumen *Functional Specification Document (FSD)* yang diberikan oleh tim PHI

4. Melakukan validasi data dan *quality checking*
5. Melakukan *performance tuning*

3.2 Tugas – Tugas yang Dikerjakan

Tugas yang dikerjakan selama penulis melaksanakan kerja magang pada PT. Putera Handal Indotama pada bagian *data warehouse* adalah melakukan *parsing* data klien yang berbentuk *excel* (*.xlsx dan *.csv) ke dalam *database MySQL*. Data yang di *parsing* harus sesuai dengan *template* yang sudah diberikan. Proses *parsing* data terdiri dari beberapa tahapan yaitu *landing zone*, *staging*, MC (Microstrategy), pembuatan *job*, dan pembuatan *kitchen*. Microstrategy adalah *software* yang digunakan oleh perusahaan klien untuk membangun *business intelligence*-nya.

Tahap *landing zone* adalah tahap awal melakukan *parsing* data dari *excel* ke *MySQL*. Pada tahap ini data mentah dari *excel* yang di-*parsing* ke *database* tidak dirubah. Tabel yang menampung data *parsing* ini akan ditambahkan beberapa *field* sebagai keterangan dari data tersebut. Keterangan tersebut berupa *file name*, *file path*, dan *buss date*. *Filename* adalah nama file *excel* yang di-*parsing* ke *database*. *File path* adalah lokasi dari *file excel* yang di-*parsing*. *Buss date* adalah tanggal yang berasal dari nama *file excel*, misalkan *file excel* bernama 'cost_adjustment_20140120.xlsx' maka *buss date*-nya adalah '2014-01-20'. Fungsi dari *field buss date* adalah untuk mengetahui periode dari *file excel* yang diproses. *File name*, *file path*, dan *buss date* akan masuk ke setiap baris data yang di-*parsing*

ke *database*. Tujuan dari penambahan *field* keterangan ini adalah untuk memudahkan mencari informasi mengenai *file* sumber ketika terjadi *error*.

Tahap *staging* adalah tahap dimana dilakukan pembersihan data, perhitungan, dan beberapa proses lainnya untuk meningkatkan kualitas data. Data yang di-*parsing* dari *file excel* adalah bukan data yang bersih. Ada data yang terduplikasi, terdapat spasi di awal dan di akhir, data yang harus dipecah dan lain-lain. Pentaho merupakan salah satu *software* yang dapat digunakan untuk melakukan *data cleansing*. Tidak hanya melakukan *data cleansing* perhitungan juga dilakukan dalam *staging* untuk dapat digunakan di tahap selanjutnya. Dalam tahap *staging* terkadang penulis harus menggunakan *regex* untuk dapat memisahkan suatu *string*. Dikarenakan banyak pekerjaan yang dilakukan dalam tahap ini, *staging* merupakan tahap yang memakan waktu cukup lama.

Setelah tahap *staging* selesai akan dilanjutkan dengan tahap MC. MC adalah kode yang menunjukkan tahap tersebut sudah menghasilkan output sesuai dengan *template* yang diberikan. Pada tahap ini dilakukan penggabungan beberapa tabel dari *database* yang sudah melalui tahap *staging*. Tahap ini juga cukup memakan waktu karena untuk dapat menggabungkan tabel yang satu dengan tabel yang lainnya dibutuhkan suatu *field* yang berhubungan, namun sering kali tidak ditemukan *field* yang berhubungan saat akan menggabungkan tabel yang satu dengan yang lain. Hal ini terjadi karena sebagian besar data dari klien hanya berasal dari *excel*. Hanya beberapa data dari klien yang berasal dari *database* sistem yang digunakan oleh klien. Hal ini menimbulkan kesulitan saat ingin menggabungkan

tabel dan harus menunggu respon mengenai keterangan atau perubahan yang harus dilakukan dari klien.

Setelah membuat *file-file transformation* untuk *parsing* data pada Pentaho Data Integration maka akan dilanjutkan dengan membuat *job*. *Job* adalah *script* yang akan menjalankan *transformation* yang sebelumnya telah dibuat. *Job* akan berjalan seperti rantai mulai dari menjalankan tahap *landing zone*, dilanjutkan ke tahap *staging* dan terakhir tahap MC. Apabila terjadi *error* pada salah satu tahap maka tidak akan bisa melanjutkan ke tahap berikutnya. Dengan menggunakan satu *job* ini maka akan dapat menjalankan lebih dari satu *transformation*.

Tahap terakhir dalam pekerjaan yang penulis lakukan adalah membuat *kitchen*. *Kitchen* adalah sebuah *batch file* yang nantinya akan dijalankan di *server* dengan *scheduler*. *Scheduler* akan membuat *file kitchen* dijalankan sesuai penjadwalan yang dilakukan. Pada *file kitchen* akan di-*set job* mana yang akan dijalankan, maka dari itu semua *file transformation*, *job*, dan *kitchen* diletakkan dalam satu *folder* untuk setiap modul.

3.3 Hasil Kerja

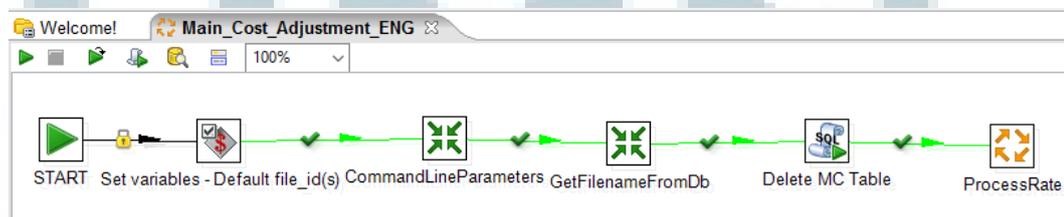
Untuk setiap tahapan yang penulis kerjakan selama praktik kerja magang ada hasil yang harus diserahkan. Pada tahap *landing zone*, *staging*, dan MC hasil keluarannya berupa *file kettle* (*.ktr). Pada tahap pembuatan *job* hasil keluarannya berupa *file* *.kjb. Sementara untuk *kitchen* bentuk hasil keluarannya adalah *file* *.bat. Semua hasil keluaran dari setiap tahapan yang dikerjakan akan diletakkan dalam satu *folder* untuk setiap modulnya.

Seluruh hasil keluaran berupa *screenshot* dari tahap *landing zone*, *staging*, *MC*, *Job* akan dilampirkan pada bagian lampiran. Salah satu contoh hasil keluaran yang sudah disatukan dalam *folder* akan seperti berikut :

Name	Date modified	Type	Size
logs	06/08/2015 20.43	File folder	
Kitchen_	10/03/2015 12.12	Windows Batch File	1 KB
Main_Cost_Adjustment_ENG.kjb	10/03/2015 12.11	KJB File	22 KB
MC_Cost_Adjustment_ENG_201401.ktr	13/03/2015 09.50	KTR File	34 KB
Parsing_Cost_Adjustment_ENG_201401.ktr	25/02/2015 13.32	KTR File	31 KB
Process_Cost_Adjustment_ENG.kjb	13/03/2015 10.06	KJB File	22 KB
Stg_Cost_Adjustment_ENG_201401.ktr	06/03/2015 15.43	KTR File	26 KB

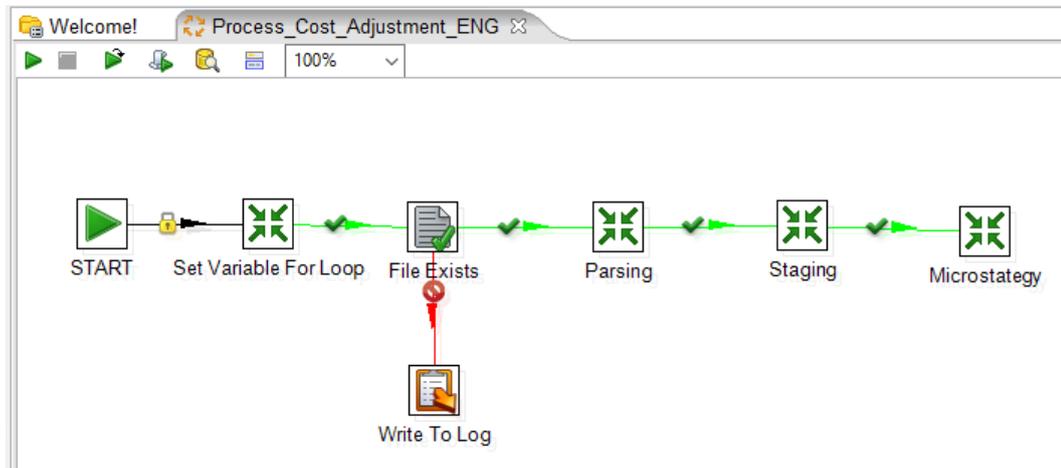
Gambar 3 : Contoh hasil pekerjaan

Untuk *file transformation* yang diawali dengan kata *Parsing*, *Stg*, dan *MC* adalah hasil pekerjaan yang dilakukan penulis yang dikerjakan dari awal. Sedangkan untuk *file job* yang diawali dengan kata *Main* dan *Process* adalah *file* yang telah disediakan untuk dilanjutkan oleh penulis. *File process* digunakan untuk dapat menjalankan *file-file transformation* yang telah dibuat dan *file Main* digunakan untuk menentukan *file-file* yang akan diproses berdasarkan *file_id*. Contoh dari *Main Job* adalah seperti berikut :



Gambar 4 : Contoh *Main Job*

Contoh dari *Process Job* adalah seperti berikut :



Gambar 5 : Contoh *Process Job*

Seluruh tahap pada *Main Job* dan *Process Job* akan sama pada setiap *folder* hanya saja *parameter* yang ada di dalamnya yang akan berbeda. Selain *Main Job* dan *Process Job* ada satu *file* lagi yang telah disediakan yaitu sebuah *batch file* (*.bat) yang disebut *kitchen*. *File* ini yang nantinya akan dijalankan oleh *scheduler* pada server. Penulis akan meng-*input* parameter berupa nama modul. Saat *file* ini dijalankan maka proses yang berlangsung akan terekam dalam sebuah *log file* yang tersimpan dalam sebuah *folder log* yang otomatis terbentuk jika belum ada.

3.4 Pekerjaan Tambahan Yang Dikerjakan

Pada saat berlangsungnya praktik kerja magang, penulis mendapatkan tambahan sebuah modul baru. Modul ini tidak tertera sesuai kontrak penulis. Namun dikarenakan modul ini adalah tambahan dari klien perusahaan tempat penulis melangsungkan praktik kerja magang, modul ini masih kurang jelas

mengenai sumber data dan *template* akhirnya. Sehingga penulis hanya melakukan tahap *landing zone* dan *staging* pada modul ini.

Penulis juga diberikan tugas untuk membuat artikel mengenai Pentaho Data Integration. Artikel ini berfungsi untuk memberikan pengetahuan teknis bagi pengguna Pentaho. Penulis membuat artikel ini berdasarkan pengalaman yang penulis dapatkan saat menjalankan praktik kerja magang. Artikel ini nantinya akan dipublikasikan pada situs IBICC.

3.5 Masalah Yang Dihadapi

Selama melaksanakan praktik kerja magang selama dua bulan, penulis menghadapi beberapa kendala yang baru pertama kali penulis alami di dunia kerja. Namun kendala-kendala tersebut dapat penulis atasi berkat pembelajaran yang penulis peroleh dari pembimbing dan tim dari tempat penulis melaksanakan kerja magang.

Kendala pertama yang dihadapi penulis adalah sulit memahami proses bisnis dari perusahaan yang menjadi klien dari *phi-integration*, hal ini dikarenakan klien dari *phi-integration* adalah perusahaan yang bergerak di bidang tambang dimana istilah-istilah yang digunakan terdengar asing bagi penulis. Selain istilah yang digunakan alur proses bisnisnya juga cukup membingungkan sehingga membuat penulis salah pengertian di awal, sebelum dijelaskan yang penulis bayangkan adalah klien ini adalah sebuah perusahaan yang memiliki sebuah pertambangan, namun setelah diberikan *briefing* yang cukup jelas oleh tim dari phi

penulis akhirnya mengerti kalau perusahaan yang menjadi klien ini adalah perusahaan yang menjadi kontraktor untuk pertambangan.

Kendala kedua yang dihadapi penulis adalah kesulitan dalam menggunakan *software* Pentaho Data Integration. Meskipun penulis telah mengikuti satu periode pada *Indonesian Business Intelligence Community on Campus* (IBICC) yang mengajarkan tentang *data warehouse* dan cara menggunakan Pentaho, namun kasusnya sangat berbeda dengan yang penulis alami dalam praktik kerja magang. Pada praktik kerja magang ini banyak sekali *tools* dalam Pentaho yang harus penulis gunakan dimana *tools* tersebut belum pernah digunakan selama belajar di IBICC. Selain dari *tools* yang belum sepenuhnya dikuasai oleh penulis, sumber *file* yang berasal dari klien juga cukup rumit seperti *header*-nya ada yang di-*merge*, terdapat baris yang kosong, ada yang memiliki nomor urut dan juga ada yang tidak, ada yang hanya berupa *template* dan lain-lain.

Kendala selanjutnya yang penulis alami adalah tidak adanya data *real* yang diberikan. Sumber data yang diberikan dari klien hanya berupa *template excel* dan data buatan. Hal ini membuat penulis kesulitan untuk mengecek apakah pekerjaan yang telah dilakukan sudah benar atau belum. Penulis diarahkan untuk membuat contoh data sendiri oleh tim dari PHI namun untuk melakukan hal tersebut penulis juga mengalami kesulitan karena kurang mengerti maksud dari kolom-kolom yang ada sehingga bingung untuk membuat data yang seperti apa. Selain dari data yang digunakan untuk mengecek pekerjaan, penulis juga sering terhenti di tengah pekerjaan karena ada kolom yang tidak ditemukan untuk dapat menggabungkan

tabel yang satu dengan tabel lainnya. Hal ini karena klien belum menggunakan sistem sepenuhnya sehingga beberapa data tidak memiliki kode atau id.

Kendala terakhir yang dialami penulis adalah *requirement* yang sering kali berubah dari klien. *Requirement* yang sering kali berubah ini membuat penulis harus merombak ulang ETL yang telah dibuat. Hal ini cukup memakan waktu karena penulis sering kali membuat ulang ETL karena perubahan yang cukup banyak. Contoh perubahan *requirement* yang dialami penulis misalnya berubahnya *template* awal, perubahan struktur tabel target, perubahan formula perhitungan.

Selain kesulitan yang penulis alami di tempat kerja magang, ada beberapa kesulitan lainnya yang penulis hadapi seperti kesulitan mengatur waktu pengerjaan karena penulis menjalani kerja magang di waktu kuliah, sehingga terkadang penulis juga harus mengerjakan pekerjaan di hari libur atau diluar jam kerja. Selain itu terdapat juga kesulitan dalam pengerjaan laporan seperti bagaimana format yang benar untuk laporan magang, apa judul yang tepat serta bagaimana tata bahasa yang baik untuk menulis laporan. Solusi dari masalah-masalah yang dihadapi oleh penulis akan dijelaskan pada bagian berikutnya.

3.6 Solusi Atas Masalah

Untuk masalah pemahaman proses bisnis klien, penulis mengikuti *briefing* yang dipimpin oleh tim dari *phi-integration* dan penulis mengajukan pertanyaan apabila ada yang tidak penulis mengerti. Pembimbing lapangan dan tim dari *phi-integration* juga membimbing penulis dalam menggunakan *Pentaho Data Integration* sehingga penulis dapat mengerti cara menggunakan *Pentaho* dengan

baik. Selain itu penulis juga diajarkan bagaimana cara membuat ETL yang baik sehingga memakan waktu dan memori yang seminimal mungkin.

Untuk kendala yang perlu berhubungan dengan klien seperti ada kolom yang kurang, adanya perubahan sumber data, tabel target dan *requirement* lainnya penulis harus berkomunikasi dengan tim dari *phi-integration* karena mereka yang berhubungan langsung dengan klien. Penulis akan diberitahu apabila sudah ada balasan atau *update* terbaru dari klien untuk pekerjaan penulis. Untuk kesulitan dalam waktu pengerjaan, penulis berusaha mengatur waktu sebaik mungkin dengan mengikuti perkuliahan terlebih dahulu setelah itu langsung melanjutkan kerja magang karena lokasi kerja magang berada di Universitas Multimedia Nusantara. Apabila ada pekerjaan yang belum terselesaikan penulis bisa melanjutkan pekerjaan tersebut di luar jam kerja.

Kendala yang berhubungan dengan laporan magang yang disusun, penulis konsultasikan dengan pembimbing laporan magang. Selain konsultasi dengan pembimbing, penulis juga melihat referensi dari laporan-laporan yang telah diselesaikan oleh kakak kelas yang terdapat di perpustakaan Universitas Multimedia Nusantara mengenai tata bahasa serta format laporan magang.