W

;/--n

**TUGAS AKHIR – KI141502**

**SINKRONISASI BASIS DATA SQL DENGAN BASIS DATA NOSQL MENGGUNAKAN *DATA ADAPTER* DENGAN PENDEKATAN *QUERY* *DIRECT ACCESS***

**I GUSTI NGURAH ADI WICAKSANA**

**NRP 5113100110**

**Dosen Pembimbing I**

**Ir. Muchammad Husni, M.Kom.**

**Dosen Pembimbing II**

**Henning Titi Ciptaningtyas, S.Kom., M.Kom.**

**JURUSAN TEKNIK INFORMATIKA**

**Fakultas Teknologi Informasi**

**Institut Teknologi Sepuluh Nopember**

**Surabaya 2017**



**TUGAS AKHIR – KI141502**

**SINKRONISASI BASIS DATA SQL DENGAN BASIS DATA NOSQL MENGGUNAKAN *DATA ADAPTER* DENGAN PENDEKATAN *QUERY DIRECT ACCESS***

**I GUSTI NGURAH ADI WICAKSANA**

**NRP 5113100110**

**Dosen Pembimbing I**

**Ir. Muchammad Husni, M.Kom.**

**Dosen Pembimbing II**

**Henning Titi Ciptaningtyas, S.Kom., M.Kom.**

**JURUSAN TEKNIK INFORMATIKA**

**Fakultas Teknologi Informasi**

**Institut Teknologi Sepuluh Nopember**

**Surabaya 2017**

***[Halaman ini sengaja dikosongkan]***



**UNDERGRADUATE THESES – KI141502**

**SYNCRONIZATION BETWEEN SQL AND NOSQL DATABASES USING DATA ADAPTER WITH DIRECT ACCESS QUERY APPROACH**

**I GUSTI NGURAH ADI WICAKSANA**

**NRP 5113100110**

**Supervisor I**

**Ir. Muchammad Husni, M.Kom.**

**Supervisor II**

**Henning Titi Ciptaningtyas, S.Kom., M.Kom.**

**DEPARTMENT OF INFORMATICS**

**FACULTY OF INFORMATION TECHNOLOGY**

**INSTITUT TEKNOLOGI SEPULUH NOPEMBER**

**SURABAYA 2017**

***[Halaman ini sengaja dikosongkan]***

# LEMBAR PENGESAHAN

**SINKRONISASI BASIS DATA SQL DENGAN BASIS DATA NOSQL MENGGUNAKAN DATA ADAPTER DENGAN PENDEKATAN QUERY DIRECT ACCESS**

**TUGAS AKHIR**

Diajukan Untuk Memenuhi Salah Satu Syarat

Memperoleh Gelar Sarjana Komputer

pada

Rumpun Mata Kuliah Komputasi Berbasis Jaringan

Program Studi S-1 Jurusan Teknik Informatika

Fakultas Teknologi Informasi

Institut Teknologi Sepuluh Nopember

Oleh

**I GUSTI NGURAH ADI WICAKSANA**

**NRP : 5113 100 110**

Disetujui oleh Dosen Pembimbing Tugas Akhir:

1. Ir. Muchammad Husni, M.Kom .....................

NIP: 19600221 198403 1 001 (Pembimbing 1)

# Henning Titi Ciptaningtyas, S.Kom., M.Kom .....................

NIP: 19840708 201012 2 004 (Pembimbing 2)

**SURABAYA**

**MARET, 2017**

***[Halaman ini sengaja dikosongkan]***

**SINKRONISASI BASIS DATA SQL DENGAN BASIS DATA NOSQL MENGGUNAKAN DATA ADAPTER DENGAN PENDEKATAN QUERY DIRECT ACCESS**

|  |  |  |
| --- | --- | --- |
| **Nama Mahasiswa** | **:** | **I Gusti Ngurah Adi Wicaksana** |
| **NRP** | **:** | **5113100110** |
| **Jurusan** | **:** | **Teknik Informatika FTIF-ITS** |
| **Dosen Pembimbing 1** | **:** | **Ir. Muchammad Husni, M.Kom.** |
| **Dosen Pembimbing 2** | **:** | **Henning Titi Ciptaningtyas, S.Kom., M.Kom.** |

# Abstrak

*Basis data NoSQL, singkatan dari Not SQL, semakin banyak digunakan seiring dengan bertambahnya jumlah aplikasi big data. Kebanyakan sistem masih menggunakan relational databases (RDB), namun seiring dengan bertambahknya jumlah data tiap tahunnya, sistem menangani big data dengan basis data NoSQL untuk menganalisis dan mengakses data dengan lebih cepat.*

*Data dari genom sendiri sangatlah besar sehingga dibutuhkan suatu teknik yang dapat mengurangi waktu komputasi. Hal tersebut dapat tercapai dengan cara mengimplementasikan komputasi paralel terhadap algoritma LCS. Oleh karena itu Tugas Akhir ini mengimplementasikan Algoritma LCS yang diparalelkan.*

*Tugas Akhir ini memiliki tujuan untuk mengimplementasikan mekanisme data adapter dalam sinkronisasi basis data RDB dan NoSQL. Berdasarkan hasil uji coba, waktu eksekusi LCS secara paralel terbukti lebih cepat daripada sekuensial dan panjang karakter yang dapat diproses pada komputasi paralel lebih banyak daripada sekuensial.*

***Kata kunci: RDB, NoSQL, Data Adapter, Sinkronisasi.***

**SYNCHRONIZATION BETWEEN SQL AND NOSQL DATABASE USING DATA ADAPTER WITH DIRECT ACCESS QUERY APPROACH**

|  |  |  |
| --- | --- | --- |
| **Nama Mahasiswa** | **:** | **I Gusti Ngurah Adi Wicaksana** |
| **NRP** | **:** | **5113100110** |
| **Jurusan** | **:** | **Teknik Informatika FTIF-ITS** |
| **Dosen Pembimbing 1** | **:** | **Ir. Muchammad Husni, M.Kom.** |
| **Dosen Pembimbing 2** | **:** | **Henning Titi Ciptaningtyas, S.Kom., M.Kom.** |

# Abstract

*The process of finding the global alignment of two genomes by comparing the sequence is one of the problems frequently encountered in computational biology. One of the ways to solve this problem is to implement the algorithm Longest common subsequence (LCS) in finding the global alignment of both genomes.*

*Data from the genome itself is large so it takes a technique that can reduce the computation time. This can be achieved by implementing parallel computation of the LCS algorithm. Therefore, this undergraduate thesis implements the parallelization of LCS algorithm.*

*The purpose of this undergraduate thesis is to find the similarity of the two genomes and analyze and compare the performance of the sequential of LCS algorithms and the parallelization of LCS algorithm. LCS algorithm implemented in parallel using OpenMPI library. Based on trial results, the execution time LCS in parallel proved faster than sequential and length of characters that can be processed in parallel computing is longer than sequential.*

***Keywords: RDB, NoSQL, Data Adapter, Synchronization.***

***[Halaman ini sengaja dikosongkan]***

# KATA PENGANTAR

Alhamdulillahirabbil’alamin, segala puji bagi Allah SWT, yang telah melimpahkan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul ***“*RANCANG BANGUN KOMPUTASI PARALEL PADA ALGORITMA *LONGEST COMMON SUBSEQUENCE* UNTUK PERBANDINGAN GENOM MENGGUNAKAN *FRAMEWORK* OPENMPI*”*.** Dengan pengerjaan Tugas Akhir ini, penulis bisa belajar lebih banyak untuk memperdalam dan meningkatkan apa yang telah didapatkan penulis selama menempuh perkuliahan di Teknik Informatika ITS. Dengan Tugas Akhir ini penulis juga dapat menghasilkan suatu implementasi dari apa yang telah penulis pelajari.

Selesainya Tugas Akhir ini tidak lepas dari bantuan dan dukungan beberapa pihak. Sehingga pada kesempatan ini penulis mengucapkan syukur dan terima kasih kepada:

1. Allah SWT dan Nabi Muhammad SAW.
2. Keluarga di Pontianak khususnya Mama, Bapak, dan Ndut yang telah memberikan dukungan moral dan material serta do’a yang tak terhingga untuk penulis. Serta selalu memberikan semangat dan motivasi pada penulis dalam mengerjakan Tugas Akhir ini.
3. Keluarga di Surabaya, Tante Vitri, Om Dodo, Mba Dita, Mas Dimas, Ivan dan Iyuk yang telah memberikan kehangatan keluarga serta support untuk penulis selama kuliah di ITS.
4. Bapak Waskitho Wibisono, S.Kom., M.Eng., Ph.D. selaku pembimbing I yang telah membantu, membimbing, dan memotivasi penulis dalam menyelesaikan Tugas Akhir ini dengan sabar.
5. Bapak Hudan Studiawan, S.Kom., M.Kom**.** selaku pembimbing II yang juga telah membantu, membimbing, dan memotivasi kepada penulis dalam mengerjakan Tugas Akhir ini.
6. Bapak Dr. Darlis Herumurti, S.Kom., M.Kom. selaku Kepala Jurusan Teknik Informatika ITS.
7. Bapak Radityo Anggoro, S.Kom., M.Sc. selaku koordinator Tugas Akhir, dan segenap dosen Teknik Informatika yang telah memberikan ilmunya.
8. Teman-teman semangat TA Altea, Angga, Bimo, Dian, Fatih, Fifi, Lubna, Luqman, Oshi, Pinas, dan Puni.
9. Teman-teman administrator laboratorium Arsitektur dan Jaringan Komputer(AJK), Mas Sam, Mas Uyung, Mbak Vivi, Kak Harum, Mas Romen, Mas Dimas, Agus, Pur, Thiar, Uul, Wicak, Zaza, Risma, Nindy, Asbun, Daniel, Fatih, Oink, dan Syukron.
10. Teman-teman angkatan 2012 yang yang telah berbagi ilmu, dan memberi motivasi kepada penulis.
11. Serta semua pihak yang yang telah turut membantu penulis dalam menyelesaikan Tugas Akhir ini.

Penulis menyadari bahwa Tugas Akhir ini masih memiliki banyak kekurangan. Sehingga dengan kerendahan hati, penulis mengharapkan kritik dan saran dari pembaca untuk perbaikan ke depannya.

Surabaya, Juni 2016

# DAFTAR ISI

# DAFTAR GAMBAR

# DAFTAR TABEL

# DAFTAR KODE SUMBER

***[Halaman ini sengaja dikosongkan]***

# DAFTAR PSEUDOCODE

***[Halaman ini sengaja dikosongkan]***

# BAB I PENDAHULUAN

## Latar Belakang

*Big data* dan *hybrid database* menjadi semakin pupoler seiring dengan berkembangnya servis *cloud computing*. The United Nations Economic Commission for Europe, atau UNECE, memprediksi jika peningkatan jumlah data akan mencapai 350% pada tahun 2019 jika dibandinkan dengan tahun 2015. [1] Basis data NoSQL, singkatan dari *Not Only SQL*, semakin banyak digunakan seiring dengan bertambahnya jumlah aplikasi *big data*.

Kebanyakan sistem dewasa ini masih menggunakan *relational database* (RDB). MySQL adalah salah satu *Database Management System* (DBMS) populer untuk aplikasi berbasis web. Dengan performanya yang terpercaya, keandalan, dan penggunaan yang mudah, MySQL telah menjadi basis data pilihan utama yang digunakan oleh perusahaan dengan profil tinggi seperti Facebook, Twitter, dan Youtube. [2] Tetapi seiring dengan jumlah data yang terus bertambah tiap tahunnya, sistem menangani *big data* dengan basis data NoSQL untuk menganalisis dan mengakses data dengan lebih cepat. NoSQL muncul sebagai akibat dari pertumbuhan eksponensial dari internet dan perkembangan aplikasi web. NoSQL mendukung pengembangan yang gesit, karena NoSQL adalah *schema-less* dan tidak tidak perlu mendefinisikan secara statik bagaimana data harus dimodelkan. [3]

Sistem yang menggunakan basis data NoSQL, ketika mengkueri data tidak menggunakan sintaks yang sama seperti mengambil data pada basisdata SQL karena sintaks *query* yang digunakan berbeda. Dibandingkan merubah kode sumber atau merubah basis data RBD menjadi NoSQL, penulis melakukan riset untuk mengintegrasikan kedua basis data. Aplikasi terkoneksi dengan RDB untuk menangani jumlah data dengan skala kecil dan menengah, server basis data NoSQL sebagai *back-end* sistem untuk menganalisis data dan melakukan sejumlah operasi *read/write*, atau secara periodik melakukan *backup* data dari RDB.

Integrasi basis data dapat mempengaruhi desain sistem yang asli. Dalam sistem, aplikasi memperoleh data dari basis data relasional menggunakan kueri SQL, namun dalam NoSQL tidak dapat diakses dengan menggunakan SQL. Oleh karena itu aplikasi harus merubah desain untuk dapat mengakses basis data RBD dan NoSQL. Mekanisme transformasi data dari RDB ke NoSQL diperlukan ketika mengintegrasikan sistem dengan NoSQL. Selama proses transformasi, aplikasi dituntut untuk menunggu hingga proses sinkronisasi selesai. Proses transformasi ini dapat berlangsung sangat lama karena jumlah data yang diubah dalam skala besar. Hal ini menjadi permasalahan utama untuk beberapa kasus seperti *real-time* atau layanan *non-stop* seperti analisis ilmiah dan aplikasi web.

## Rumusan Masalah

Rumusan masalah yang diangkat dalam Tugas Akhir ini adalah sebagai berikut:

1. Bagaimana mengimplementasikan data adapter untuk mensinkronisasikan RDB dan NoSQL?
2. Bagaimana hasil implementasi data adapter dalam sinkronisasi basis data RDB dan NoSQL?

## Batasan Masalah

Permasalahan yang dibahas dalam Tugas Akhir ini memiliki beberapa batasan yaitu sebagai berikut:

* 1. Banyak node server untuk basis data yang digunakan adalah berjumlah dua komputer.

1. Mekanisme yang digunakan adalah menggunakan konsep data adapter
2. Basis data yang digunakan adalah MySQL sebagai basis data RDB dan Apache HBase sebagai basis data NoSQL.

## Tujuan

Tujuan Tugas Akhir ini adalah mengimplementasikan mekanisme data adapter dalam sinkronisasi basis data RDB dan NoSQL.

## Manfaat

Dengan dibuatnya Tugas Akhir ini yaitu dapat mensinkronisasikan basis data RDB dengan NoSQL dengan mekanisme data adapter.

## Metodologi

Tahapan-tahapan yang dilakukan dalam pengerjaan Tugas Akhir ini adalah sebagai berikut:

1. Penyusunan proposal Tugas Akhir.

Tahap awal untuk memulai pengerjaan Tugas Akhir adalah penyusunan proposal Tugas Akhir. Penyusunan proposal Tugas Akhir dilaksanakan untuk merumuskan masalah serta melakukan penetapan rancangan dasar dari sistem yang akan dikembangkan dalam pelaksanaan Tugas Akhir ini.

1. Studi literatur

Pada tahap ini dilakukan pencarian informasi dan literatur dari paper dan artikel di internet yang diperlukan untuk tahap implementasi program. Tahap ini diperlukan untuk membantu memahami konsep dan penggunaan komponen-komponen terkait dengan sistem yang akan dibangun, diantaranya data adapter, basis data MySQL dan Apache Hbase.

1. Analisis dan perancangan perangkat lunak

Tahap ini meliputi perancangan sistem berdasarkan studi literatur dan pembelajaran konsep teknologi dari perangkat lunak yang ada. Tahap ini mendefinisikan alur dari implementasi. Langkah-langkah yang dikerjakan juga didefinisikan pada tahap ini. Pada tahapan ini dibuat *prototype* sistem, yang merupakan rancangan dasar dari sistem yang akan dibuat. Serta dilakukan desain fungsi yang akan dibuat yang ditunjukkan melalui *pseudocode*.

1. Implementasi perangkat lunak

Implementasi perangkat lunak merupakan tahap membangun rancangan program yang telah dibuat. Pada tahap ini akan direalisasikan mengenai rancangan apa saja yang telah didefiniskan pada tahap sebelumnya. Fungsi yang ada pada tahap ini merupakan fungsi hasil implementasi dari tahap analisis dan perancangan perangkat lunak.

1. Pengujian dan evaluasi

Pada tahap ini dilakukan pengujian dan pencatatan hasil dari metode yang diimplementasikan. Pengujian akan dilakukan di Laboratorium Arsitektur dan Jaringan Komputer. Parameter yang akan diujikan adalah berapa waktu yang dibutuhkan untuk melakukan sinkronisasi basis data antara RDB dan HBase NoSQL database sehingga dapat diketahui efektifitas implementasi sistem data adapter. Banyak komputer server basis data yang digunakan adalah dua komputer.

1. Penyusunan buku Tugas Akhir

Pada tahap ini disusun buku yang memuat dokumentasi mengenai perancangan, pembuatan serta hasil dari implementasi perangkat lunak yang telah dibuat.

## Sistematika Penulisan Laporan Tugas Akhir

Buku Tugas Akhir ini bertujuan untuk mendapatkan gambaran dari pengerjaan Tugas Akhir ini. Secara garis besar, buku Tugas Akhir terdiri atas beberapa bagian seperti berikut ini:

**Bab I Pendahuluan**

Bab yang berisi mengenai latar belakang, tujuan, dan manfaat dari pembuatan Tugas Akhir. Selain itu perumusan masalah, batasan masalah, metodologi yang digunakan, dan sistematika penulisan juga merupakan bagian dari bab ini.

**Bab II Tinjauan Pustaka**

Bab ini berisi penjelasan secara detail mengenai dasar-dasar penunjang dan teori-teori yang digunakan untuk mendukung pembuatan Tugas Akhir ini.

**Bab III Analisis dan Perancangan Perangkat Lunak**

Bab ini berisi tentang dasar dari algoritma yang akan diimplementasikan pada Tugas Akhir ini.

**Bab IV Implementasi**

Bab ini membahas mengenai implementasi dari rancangan yang telah dibuat pada bab sebelumnya.

**Bab V Uji Coba Dan Evaluasi**

Bab ini menjelaskan mengenai kemampuan perangkat lunak dengan melakukan pengujian kebenaran dan pengujian kinerja dari perangkat lunak yang telah dibuat sesuai dengan data yang diujikan.

**Bab VI Kesimpulan Dan Saran**

Bab ini merupakan bab terakhir yang menyampaikan kesimpulan dari hasil uji coba yang telah dilakukan dan saran untuk pengembangan perangkat lunak ke depannya.

# BAB II TINJAUAN PUSTAKA

Bab ini berisi penjelasan teori-teori yang berkaitan dengan algoritma yang diajukan pada pengimplementasian program. Penjelasan ini bertujuan untuk memberikan gambaran secara umum terhadap program yang dibuat dan berguna sebagai penunjang dalam pengembangan perangkat lunak.

## Basis Data

Basis data atau database merupakan sebuah koleksi atau kumpulan dari data yang bersifat mekanis, terbagi, terdefinisi secara formal serta terkontrol. Pengontrolan dari sistem database tersebut adalah terpusat, yang biasanya dimiliki dan juga dipegang oleh suatu organisasi

Bentuk sebuah basis data, elektronik dan lainnya, harus di rencanakan. Proses dalam merancang basis data adalah aktivitas dalam merepresentasikan class, atribut, dan relasi antar basis data. Data adalah fakta. Informasi menggambarkan data. Informasi adalah data dengan konteks yang berarti

## Relational Database

*Relational database*, atau disingkat RDB, menggambarkan suatu kumpulan dari banyak relasi. Konsep RDB pertama kali diperkenalkan oleh Dr. Codd pada tahun 1970. Sebuah sistem yang mengatur hal ini disebut dengan Relational Database Management System (RDBMS). Model relasional terdiri dari beberapa komponen, diantaranya kumpulan objek berelasi, kumpulan operasi yang bekerja pada relasi, dan integritas data untuk.

Sebagian besar database yang digunakan pada aplikasi modern saat ini adalah relasional database. Relational databases adalah model database yang menyimpan data pada tabel. Setiap tabel terdiri dari baris (record) dan kolom (field). Dalam terminologi ilmu komputer, baris sering disebut dengan “tuples”, dan kolom dapat disebut dengan “atribut”. Sebuah tabel dapat divisualisasikan sebagai sebuah matriks baris dan kolom, dimana setiap persimpangan dari baris dan kolom berisi nilai tertentu. Hal ini adalah relasional selama semua *record* berbagi bidang yang sama.

Tabel juga sering terdapat *primary key*, yang menyediakan tanda pengenal unik untuk setiap baris dalam tabel. Kunci ini bisa ditunjuk ke kolom, atau dapat terdiri dari beberapa kolom yang bersama-sama membentuk kombinasi unik dari beberapa nilai. Disisi lain, primary key menyediakan cara yang efisien untuk pengindeksan dan dapat digunakan untuk berbagi nilai antar tabel dalam database. Sebagai contoh, nilai *primary key* dari satu tabel dapat digunakan pada field baris pada tabel lainnya. Nilai yang disisipkan ini tabel lain ini disebut dengan *foreign key*

Cara yang digunakan untuk mengakses data pada database relasional adalah dengan menggunakan kueri SQL (*Structured Query Language).* Kueri SQL dapat digunakan untuk membuat, memodifikasi, dan menghapus tabel, serta memilih, insert, dan menghapus data dari tabel yang ada [2].

## NoSQL Database

Database NoSQL, atau *Not Only SQL*, adalah sebuah pendekatan untuk manajemen data dan desain *database* yang berguna untuk kumpulan distribusi data yang sangat besar. NoSQL, meliputi berbagai teknologi dan arsitektur, berusaha untuk memecahkan masalah skalabilitas dan permasalahan performa big data pada kinerja database relasional. NoSQL sangat berguna ketika suatu perusahaan perlu untuk mengakses dan menganalisis data yang tidak terstruktur dalam jumlah yang sangat besar atau data yang disimpan dari jarak jauh pada beberapa server virtual.

NoSQL mencakup berbagai teknologi basis data yang berbeda dan dikembangkan untuk menangani permintaan dalam membangan aplikasi modern :

* Pengembang bekerja dengan aplikasi yang membuat volume baru yang sangat besar, dengan cepat melakukan perubahan tipe data (terstruktur, tidak terstruktur, semi terstruktur dan polimorfik)
* Lama pengerjaan siklus pengembangan waterfall dua sampai delapan belas bulan. Dijaman sekarang, tim bekerja dengan secepat-cepatnya, mengiterasi dan mem-push code hampir setiap minggu bahkan setiap hari.
* Aplikasi yang pernah dibuat untuk khalayak tertentu, kini dibuat menjadi servis yang harus selalu menyala, dapat diakses dari berbagai perangkat dalam skala globbal oleh jutaan pengguna.
* Perusahaan kini beralih ke *scale-out architecture* menggunakan perangkat lunak *open source*, server komoditas dan komputasi awan ketimbang monolotik server dan infrastruktur penyimpanan.

Basis data relasional tidak dirancang untuk mengatasi permasalahan dengan tantangan skala dan kegesitan yang dihadapi aplikasi modern. Basis data relasional juga tidak dibangun untuk mendapatkan keuntungan dari penyimpanan komoditas dan kemampuan pemrosesan yang tersedia saat ini.

## Data Adapater

Data adapter sangat termodulasi, yang mana terletak diantara aplikasi dengan basis data. Data adapter bertanggung jawab sebagai untuk melakukan query dari aplikasi dan transformasi data antar basis data dalam waktu. Sistem menyajikan antarmuka parsing query SQL untuk mengakses RDB dan NoSQL.

Terdapat tiga fitur utama pada data adapter. Yang pertama adalah data adapter menyediakan antarmuka SQL untuk basis data RDB dan NoSQL. Di dalamnya terdiri atas Apache Phoenix sebagai penerjemah SQL untuk terhubung dengan HBase, dan MySQL JDBC driver untuk menghubungkan dengan RDB yaitu MySQL. Dengan ini aplikasi tidak perlu mengubah kueri untuk menangani kueri NoSQL.

Berikutnya adalah *Database Converter* yang digunakan untuk menangani transformasi basis data. *Database Converter* merubah data dari MySQL ke HBase menggunaakn Apache Sqoop dan Apache Phoenix. Mekanisme ini mensinkronisasi data setelah transformasi selesai dilakukan dengan penambalan blok kueri.

Fitur terakhir adalah *Query Approach* atau pendekatan kueri. Pendekatan kueri yang digunakan adalah *direct access (DA mode)*. Pendekatan ini memperbolehkan aplikasi untuk melakukan kueri secara langsung ke basis data tanpa harus menunggu proses transformasi selesai dilakukan.

## MySQL

MySQL adalah basis data *open-source* yang paling terkenal di seluruh dunia. MySQL termasuk kedalam basis data relasional. Dengan performa yang terbukti, kehandalan dan penggunaanya yang mudah, membuat MySQL menjadi basis data dengan pilihan terdepan untuk aplikasi bebrbasis web. MySQL telah digunakan oleh beberapa situs bersar dunia seperti Facebook, Twitter, Youtube, Yahoo dan masih banyak lagi. Oracle menjadi penggerak inovasi MySQL, menyajikan kapabilitas baru untuk memperkuat web, *cloud*, *mobile* dan *embedded application* dimasa mendatang. [3]

MySQL memiliki beberapa terminologi dasar. MySQL membuat, mengkonfigurasikan dan berkomunikasi dengan database. Sebuah database adalah kumpulan data yang teratur. Tabel terdiri dari beberapa records (biasa disebut dengan baris), dan record mengandung field (biasa disebut dengan kolom). Tabel 2. 1 adalah contoh gambaran sederhana bentuk tabel pada basis data MySQL. [4]

Pada penilitian ini, MySQL berperan sebagai basis data relasional. Basis data MySQL menyimpan data asli sistem, dimana pada percobaan penelitian ini, basis data yang ditambahkan menggunakan basis data NoSQL yaitu Apache Hbase. Kedua basis data dinsinkronisasi menggunakan data adapter.

Tabel 2. . Contoh tabel pada basis data MySQL

|  |  |  |  |
| --- | --- | --- | --- |
|  | *Field 1* | *Field 2* | *Field 3* |
| *Field Names:* | **Nama** | **Umur** | **Warna Favorit** |
| *Record 1* | Bruce Callow | 13 | Tidak punya |
| *Record 2* | Frank Wright | 37 | Merah |
| *Record 3* | Seymour Hawthorne | 82 | Hitam, Putih |

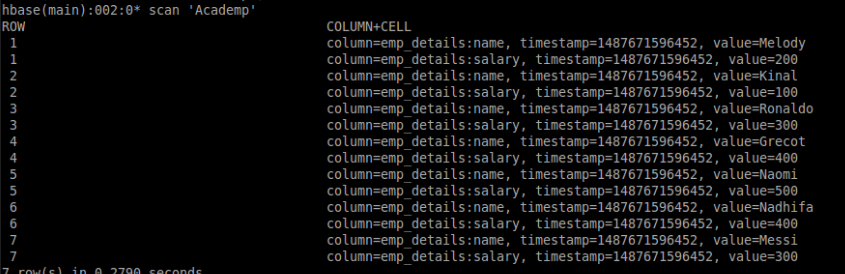
## Apache HBase

Apache HBase adalah salah satu basis data NoSQL yang open source dimana menyajikan akses *real-time read/write* untuk sebuah basis data yang besar. HBase berjalan dalam sebauh file sistem yang bernama Hadoop. Skala penggunaan Hbase adalah untuk menangani data set yang sangat besar dengan berjuta-juta baris dan kolom, dan basis data ini dengan mudah mengkombinasikan sumber data yang menggunakan struktur dan skema dengan variasi yang berbeda. HBase terintegrasi dengan Hadoop dan disamping itu bekerja dengan baik pada *data* *engine* YARN.

Apache HBase memberikan beberapa fitur seperti secara acak dan *real-time* mengakses data pada Hadoop. HBase dibuat untuk menampung tabel yang berukuran sangat besar, menjadikannya pilihan yag tepat untuk menyimpan *multi-structure* atau data yang jarang. Pengguna dapat melakukan query HBase dalam waktu tertentu, untuk melakukan “*flashback*” *query*. Karakteristik ini membuat HBase menjadi pilihan yang tepat untuk menyimpan data *semi-structured* seperti data log dan memberikan data yang sangat cepat untuk pengguna atau aplikasi yag terintegrasi dengan HBase. [5]

Pada pengerjaan tugas akhir ini, Apache HBase digunakan sebagai basis data NoSQL. Apache HBase disinkronisasikan dengan basis data SQL, dalam tugas akhir ini menggunakan MySQL, melalui data adapter. Contoh bentuk data pada Apache HBase ditunjukan pada Gambar 2.1.

Gambar . Contoh data di basis data Apache HBase



## Apache Phoenix

Banyak perkakas kecerdasan bisnis dan perkakas analisis data memiliki kemampuan yang kurang untuk bekerja dengan basis data HBase secara langsung. Apache Phoenix dapat membuat Anda untuk berinteraksi dengan HBase menggunakan SQL. Dengan menggunakan ODBC *drivers*, Anda dapat menghubungkan aplikasi ODBC dengan HBase melalui Apache Phoenix. [6] Beberapa perusahaan besar telah menggunakan Apache Phoenix seperti salesforce.com, Bloomberg, Sogou, Hortonworks, Alibaba.com, CertusNet, ebay.com, NGDATA, eHarmony dan PubMatic. [7]

Apache Phoenix adalah lapisan SQL yang efisien untuk Apache HBase. Apache Phoenix menambahkan SQL ke HBase, aplikasi *big data* yang terdistribusi dan *scalable* dan berjalan di atas file sistem Hadoop. Phoenix, bertujuan untuk mempermudah mengakses basis data HBase dengan mendukung sintaks SQL dan memperbolehkan *input* dan *output* dengan menggunakan JDBC API standar dibandingkan dengan menggunakan *HBase’s Java client APIs*. Hal ini membuat kita dapat melakukan operasi CRUD dan DDL seperti membuat tabel, memasukan data baru, dan kueri data. SQL dan JDBC mengurangi jumlah kode yang harus dituliskan oleh pengguna, dapat melakukan optimasi performa yang transparan untuk pengguna, dan membukakan perkakas lain yang ada untuk menggunakan dan mengintegrasikannya.

Lebih mendalam, Phoenix mengambil kueri SQL, meng-*compile*-nya menjadi rangkaian pemanggilan API HBase, dan mendorongnya kedalam kluster untuk eksekusi pararel. Phoenix secara otomatis membuat *metadata* yang menyediakan tipe akses untuk data yang disimpan ke tabel HBase. Phoenix secara langsung menggunakan HBase API, bersama dengan *coprocessor* dan penyaringan secara *custom*, yang menghasilkan performa seper seribu detik untuk kueri data kecil, atau beberapa detik untuk 10 juta baris. [8]

Dalam pengerjaan tugas akhir ini, Apache Phoenix digunakan pada data adapter sebagai data converter untuk merubah data dari MySQL ke HBase. Selain itu, Apache Phoenix juga diunakan sebagai digunakan untuk transformasi data dari HBase yang mana digunakan oleh aplikasi untuk menjalankan sintaks SQL untuk mengambil data melalui data adapter.

## Python

Python dikembangkan oleh Guido van Rossum di akhir tahun 80-an di *National Research Institute for Mathematics and Computer Science Netherlands*. Python berasal dari beberapa bahasa pemrogramman lain seperti ABC, Modula-3, C, C++, Algol-68, SmallTalk, Unix Shell dan bahasa pemrograman lainnya. Seperti Perl, Python berada dibawah GNU *General Public License* (GPL). [11]

Python adalah salah satu bahasa pemrograman tingkat tinggi yang *interpreted* dan *object-oriented* dengan semantik yang dinamis. Dibuat di dalam struktur data, dikombinasikan dengan tipe dan ikatan yang dinamis, membuatnya menjadi atraktif untuk mengembangkan aplikasi secara sepat. Bahasa python adalah bahasa pemrograman yang sederhana, mudah untuk dipelajari dan karena itu dapat mengurangi biaya perawatan. Python mendukung penggunaan modul dan paket, yang mendorong modularitas dan penggunaan kode kembali. Python *intepreter* dan *library* standar yang tersedia dapat digunakan secara gratis dan dapat didistribusikan dengan bebas. Banyak programmer yang suka dengan bahasa Python karena dapat meningkatkan produktivitas seperti tidak perlu adanya langkah kompilasi, dan siklus edit-tes-debug menjadi sangat cepat.[12]

# BAB III ANALISIS DAN PERANCANGAN PERANGKAT LUNAK

Pada bab ini akan dijelaskan perancangan perangkat lunak yang dibuat. Perancangan akan dibagi menjadi dua tahapan utama, yaitu perancangan program dengan OpenMPI dan peracancangan alur proses utama program. Pada bab ini juga akan dijelaskan mengenai gambaran umum setiap proses utama program dalam diagram alir beserta penjelasannya.

# BAB IV IMPLEMENTASI

Pada bab ini akan dibahas mengenai implementasi yang dilakukan berdasarkan rancangan yang telah dijabarkan pada bab sebelumnya. Sebelum penjelasan implementasi akan ditunjukkan terlebih dahulu lingkungan untuk melakukan implementasi.

## Lingkungan Implementasi

# BAB V UJI COBA DAN EVALUASI

Pada bab ini akan dijelaskan uji coba yang dilakukan pada aplikasi yang telah dikerjakan serta analisa dari uji coba yang telah dilakukan. Pembahasan pengujian meliputi lingkungan uji coba, skenario uji coba yang meliputi uji kebenaran dan uji kinerja serta analisa setiap pengujian.

## Lingkungan Uji Coba

## *Dataset* Uji Coba

## Skenario dan Evaluasi Pengujian

### Skenario Uji Coba 1

***[Halaman ini sengaja dikosongkan]***

# BAB VI KESIMPULAN DAN SARAN

Bab ini membahas mengenai kesimpulan yang dapat diambil dari hasil uji coba yang telah dilakukan pada bab sebelumnya, yaitu Bab Uji Coba dan Evaluasi. Bab ini juga digunakan sebagai jawaban dari rumusan masalah yang dikemukakan pada Bab Pendahuluan. Selain kesimpulan, juga terdapat saran yang ditujukan untuk pengembangan penelitian lebih lanjut.

## Kesimpulan

## Saran

# DAFTAR PUSTAKA

|  |  |
| --- | --- |
| [1] | G. Audin, “No Jitter,” UBM, 2017. [Online]. Available: http://www.nojitter.com/post/240170228/the-network-impact-of-big-data. [Diakses 03 April 2017]. |
| [2] | P. ". D. D. Christensson, “Techterms.com,” Sharpened Production , 16 July 2016. [Online]. Available: https://techterms.com/definition/relational\_database.. [Diakses 28 March 2017]. |
| [3] | MySQL, “MySQL,” Oracle Corporation, [Online]. Available: https://www.mysql.com/about/. [Diakses 20 10 2016]. |
| [4] | P. Lutus, “arachnoid.com,” 2012. [Online]. Available: https://arachnoid.com/MySQL/. [Diakses 28 Marh 2017]. |
| [5] | H. Inc., “HORTONWORKS,” Hortonworks Inc., 2016. [Online]. Available: http://hortonworks.com/apache/hbase/. [Diakses 20 11 2016]. |
| [6] | “HORTONWORKS,” Hortonworks Inc, [Online]. Available: https://hortonworks.com/hadoop-tutorial/bi-apache-phoenix-odbc/. [Diakses 6 April 2017]. |
| [7] | Apache Software Foundation, “Apache Phoenix,” Apache Software Foundation, [Online]. Available: https://phoenix.apache.org/who\_is\_using.html. [Diakses 6 April 2017]. |
| [8] | S. Srungarapu, “Cloudera Engineering Blog,” Cloudera, Inc, 6 Mei 2015. [Online]. Available: https://blog.cloudera.com/blog/2015/05/apache-phoenix-joins-cloudera-labs/. [Diakses 6 Maret 2017]. |
| [9] | N. F. Almeida Jr, E. N. Caceres, C. E. R. Alves dan S. W. Song, “Comparison of Genomes using High-Performance Parallel Computing,” dalam *Proceedings of the 15th Symposium on Computer Architecture and High Performance Computing*, 2003. |
| [10] | GeeksforGeeks, “Printing Longest Common Subsequence,” [Online]. Available: http://www.geeksforgeeks.org/printing-longest-common-subsequence/. [Diakses 11 May 2016]. |
| [11] | “NCBI National Center for Biotechnology Information,” [Online]. Available: http://www.ncbi.nlm.nih.gov/. [Diakses 21 June 2015]. |
| [12] | C. E. R. Alves, E. N. Cáceres, F. Dehne dan S. W. Song, “A Parallel Wavefront Algorithm for Efficient,” dalam *International Conference on Computational Science and Applications 2003*, Canada, 2003. |

**LAMPIRAN A**

# KODE SUMBER

***[Halaman ini sengaja dikosongkan]***

BIODATA PENULIS