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**

**MEI, 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: LCS, OpenMPI, Dynamic Programming, Genom.***

***[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.

## Big Data

Istilah “*Big Data*” muncul seiring dengan berkembangnya proses industri, personil dan teknologi untuk mendorong bidang baru apa yang tampaknya akan meledak. Perusahaan besar seperti Amazon dan Wal-Mart serta badan lain seperti pemerintahan Amerika Serikat dan NASA menggunakan *big data* untuk memenuhi tujuan bisnis dan strategi mereka. *Big data* dapat memainkan peran bagi perusahaan skala menengah keatas dan organisasi yang dimanfaatkan agar mendapatkan keuntungan.

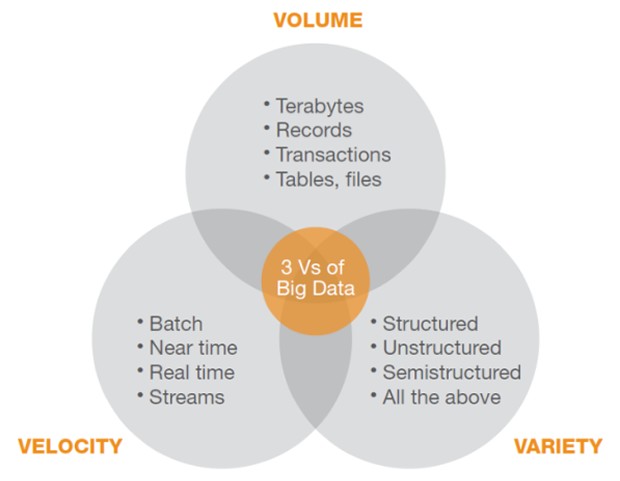
Tidak ada hal yang baru terkait dengan gagasan *big data*, yang mana telah ada sejak tahun 2001. Big data adalah informasi yang dimiliki perusahaan, didapatkan dan diproses dengan teknik baru untuk menghasilkan nilai dengan cara terbaik. [4]

Pada tahun 2001, Gartner’s Doug Laney pertama kalinya memperkenalkan istilah yang dikenal dengan “Tiga V” untuk menggambarkan beberapa karakteristik yang membuat *big data* berbeda dengan pemrosesan data yang lain. ‘Tiga’ V itu ialah *Volume*, *Velocity*, dan *Variety*. [5]

*Volume*, atau kapasitas, mengacu pada jumlah data yang tersedia untuk dianalisis. Namun ketika data terkumpulkan dari pertumbuhan jumlah perangkat, seperti sensor, telepon seluler dan komputer, jumlah data yang terakumulasikan menjadi sangat banyak. Namun sudah ada beberapa teknologi, seperti Hadoop, yang memudahkan beban ini.

*Velocity*, atau kecepatan, mengacu pada kecepatan pengolahan data. Karena data datang dari berbagai sumber, hal itu bisa terkumpulkan dengan cepat. Tantangannya adalah mengumpulkannya secepat dan seefisien mungkin.

*Variety*, atau keberagaman, mengacu pada banyaknya tipe format data. Contohnya seperti tipe data terstruktur, numerik, dokumen, suara, video, *e-mail* dan data transaksi penjualan. Tantangannya adalah membandingkan dan membedakan banyak data sedemikian rupa sehingga menjadi berpola dan dapat dimanfaatkan. [6]



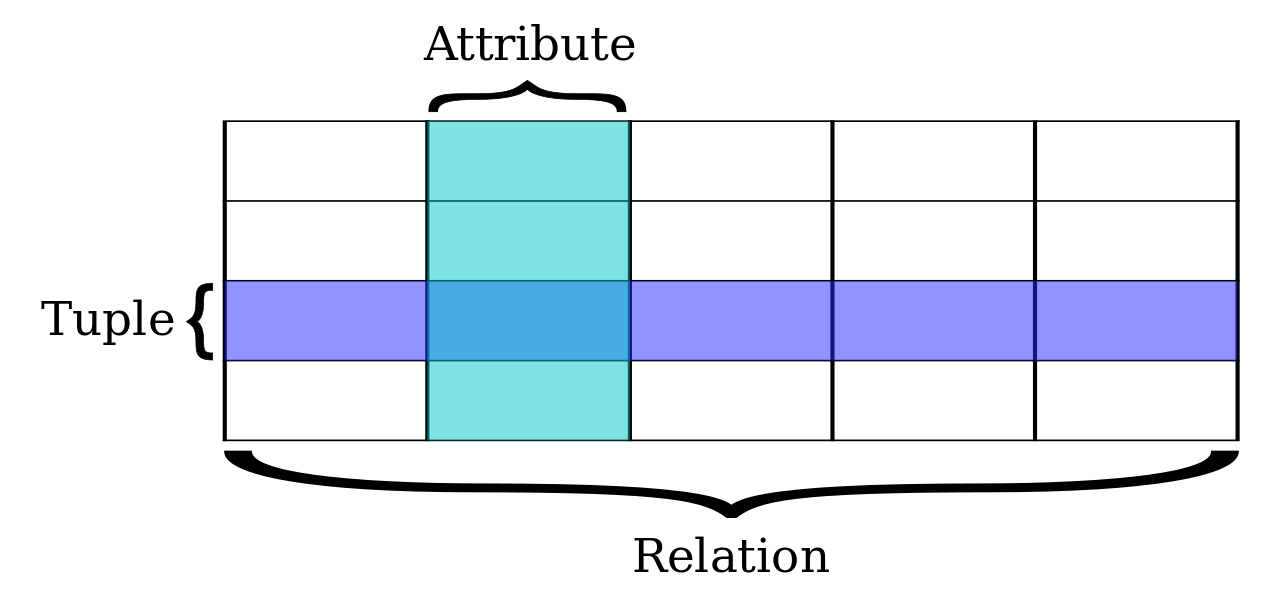
Gambar . Diagram Tiga ‘V’ Big Data

## 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 kelas, atribut, dan relasi antar basis data. Data adalah fakta. Informasi menggambarkan data. Informasi adalah data dengan konteks yang berarti. [7]

## Basis Data Relasional



Gambar . Terminologi Basis Data Relasional

*Relational database*, atau disingkat RDB, menggambarkan suatu kumpulan dari banyak relasi. Konsep basis data relasional, bahasa indonesia dari *relational database,* 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*. Basis data relasional 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 “*attribute*”. 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. Gambar 2.2 menunjukan terminologi basis data relasional.

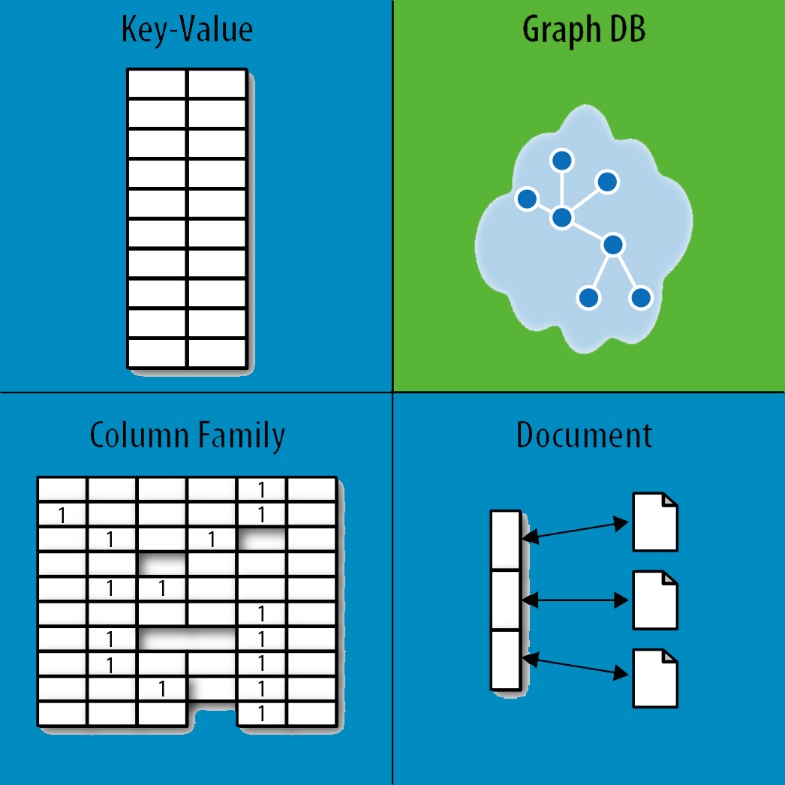
Tabel juga sering terdapat *primary key*, yang menyediakan tanda pengenal unik untuk setiap baris dalam tabel. Kunci ini bisa ditunjuk ke kolom, atau dapat terdiri dari beberapa kolom yang bersama-sama membentuk kombinasi unik dari beberapa nilai. Disisi lain, *primary key* menyediakan cara yang efisien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

Cara yang digunakan untuk mengakses data pada basis data 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. [4]

## Basis Data NoSQL

Basis data NoSQL, kependekan dari *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. [9]

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



Gambar . Beberapa Tipe Format Basis Data NoSQL

* 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* adalah dua sampai delapan belas bulan. Dijaman sekarang, tim bekerja dengan secepat-cepatnya, mengiterasi dan mem-*push* kode 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 global 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. [10]

Basis data NoSQL adalah spektrum teknologi penyimpanan data yang lebih bervariasi sehingga sulit untuk membuat karakteristik mereka menjadi general. NoSQL memberikan performa dan skalabilitas yang lebih baik jika dibandingkan dengan basis data relasional. Terdapat beberapa jenis tipe basis data NoSQL mencakup kelompok data secara kolektif seperti *key-value*, *column family*, dan dokumen serta berbagai jenis basis data grafik seperti grafik, *hypergraf* dan penyimpanan RDF triple. [11]

## 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.

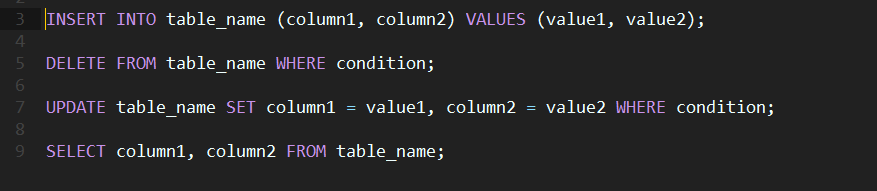
## MySQL

MySQL adalah aplikasi manajemen basis data, terjemahan dari *Database Management System* (DBMS), *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. [5]

Sebagai basis data yang relasional, MySQL menyimpan data dalam tabel yang terpisah dibandingkan dengan menyimpannya kedalam satu tabel. Model logika, dengan objek seperti basis data, tabel, *view*, baris dan kolom, menawarkan lingkungan pemrograman yang fleksibel. Basis data relasional memiliki pengaturan dasar diantara antara tabel yang berbeda seperti *one-to-one*, *one-to-many*, unik, diperlukan atau opsional, dan ‘pointer’ di antara tabel yang berbeda. Basis data dengan pengaturan seperti ini dirancang agar aplikasi tidak mengakses data yang tidak konsisten, duplikat, orphan, *out-of-date*, atau data yang hilang. [5]

MySQL memiliki beberapa terminologi dasar. MySQL membuat, mengkonfigurasikan dan berkomunikasi dengan basis data. Sebuah basis data adalah kumpulan data yang teratur. Tabel terdiri dari beberapa *records* (biasa disebut dengan baris), dan record mengandung *field* (biasa disebut dengan kolom). [6] Untuk mengakses dan mengubah data yang tersimpan, MySQL menggunakan sintaks *Structured Query Language* (SQL) sebagai standar bahasanya. Selain MySQL, ada beberapa aplikasi DBMS yang menggunakan sintaks SQL sebagai bahasa standar untuk memanipulasi data didalamnya, seperti SQL Server, MS Access, Sybase, Informix, dan Postgres. Beberapa sintaks sederhana dapat dilihat pada Gambar 2.1. Tabel 2. 1 adalah contoh gambaran sederhana bentuk tabel pada basis data MySQL.

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.



Gambar . Contoh Sintaks SQL

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

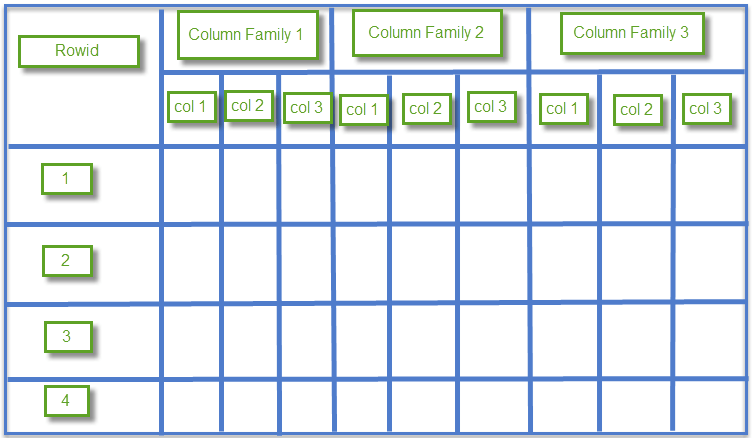
Apache Hbase dibuat pada tahun 2007 di Powerset dan awalnya merupakan bagian dari Hadoop. Sejak saat itu, proyek ini menjadi proyek tingkat atas dibawah Apache Software Foundation. HBase tersedia dibawah lisensi Apache Software License, versi 2.0. [7]

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.

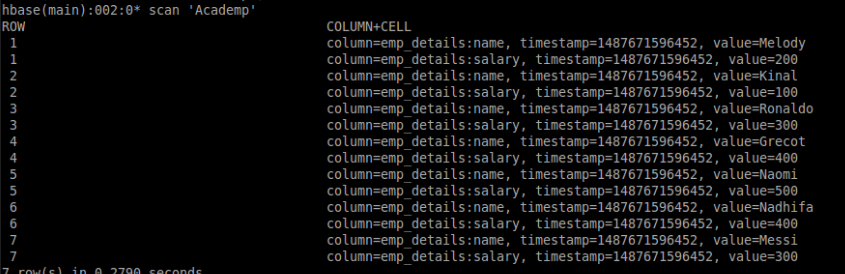
Apache HBase memberikan beberapa fitur seperti secara acak dan *real-time* mengakses data pada Hadoop. HBase dibuat untuk menampung tabel yang berukuran sangat besar, menjadikannya pilihan yang tepat untuk menyimpan *multi-structure* atau data yang jarang. Pengguna dapat melakukan kueri HBase dalam waktu 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. [8]

HBase adalah basis data yang berorientasi kolum dan data disimpan dalam tabel. Tabel diurutkan berdasarkakn *row id*. HBase memiliki *row id* yang merupakan kumpulan dari beberapa *column family* yang ditampilkan dalam bentuk tabel. *Column family* yang ditampilkan dalam skema adalah pasangan *key-value*. Jika dilihat secara rinci, setiap *column family* memiliki beberapa kolom. Nilai kolom ini disimpan dalam memori disk. Setiap sel dalam tabel memiliki data meta sendiri, seperti stempel waktu dan data lainnya. [9] Skema penyimpan di Hbase dapat dilihat pada Gambar 2.5.

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 yang disimpan pada Apache HBase jika dilihat melalui terminal ditunjukan pada Gambar 2.6.



Gambar . Skema penyimpanan pada Apache Hbase



Gambar . Contoh tampilan data yang disimpan dalam Hbase jika diakses melalui terminal

## 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. [10] Beberapa perusahaan besar telah menggunakan Apache Phoenix seperti salesforce.com, Bloomberg, Sogou, Hortonworks, Alibaba.com, CertusNet, ebay.com, NGDATA, eHarmony dan PubMatic. [11]

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. [12]

Apache Phoenix menyediakan bahasa kueri SQL yang sudah dikenal dan jauh lebih mudah dibandingkan dengan bahasa shell HBase atau HBase Java API. Manfaat utamanya adalah menjalankan semua kueri secara pararel si semua server wilayah. Akibatnya, HBase bekerja lebih cepat dan lebih efisien. Lapisan ini kompatibel dengan driver JDBC sehingga migrasi dari sistem basis data relasional biasa tidak memerlukan perubahan kode utama. Phoenix memerlukan beberapa perubahan sintaks, beberapa sintaks yang mengalami perubahan dapat dilihat pada Tabel 2.1.

Dalam pengerjaan tugas akhir ini, Apache Phoenix digunakan pada data adapter sebagai data *converter* untuk merubah data dari MySQL ke HBase dengan menjalankan sintaks kueri MySQL. Selain itu, Apache Phoenix juga digunakan oleh aplikasi untuk mengambil data dari HBase melalui data adapter.

Tabel . Beberapa Perbedaan Sintaks Antara Apache Phoenix dengan MySQL

|  |  |
| --- | --- |
| **Phoenix** | **MySQL** |
| **UPSERT** INTO t (col) VALUES (val); | **INSERT** INTO t (col) VALUES (val); |
| **UPSERT** INTO t (PK, col1) **SELECT** PK, ‘val1’ FROM t WHERE col2 = ‘val2’; | **UPDATE** t SET col1 = ‘val1’ WHERE col2 = ‘val2’; |

## 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). [13]

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. [14]

## Flask

Flask adalah sebuah kerangka kerja berbasis Python yang dipelopori oleh Armin Ronacer. Flask berada dibawah lisensi BSD. Flask menyediakan alat, beberapa library, dan teknologi yang membantu dalam membuat aplikasi web. Aplikasi web ini bisa terdiri dari beberapa halaman, blog, wiki atau e-commerce.

Flask termasuk kedalam kategori micro-framework. Micro-framework adalah kerangka kerja biasa dengan sedikit atau tanpa ketergantungan dengan libraries eksternal. Hal ini menimbulkan pro dan kontra. Kelebihannya adalah framework menjadi ringan, memiliki ketergantungan yang sedikit untuk mengupdate dan mengamati bug keamanan. [15]

Untuk dapat menggunakan Flask, yang perlu dilakukan adalah menginstallnya dengan menggunakan pip. Pip akan secara otomatis melakukan pemasangan berdasarkan versi Python yang ada di komputer. Secara sederhana, contoh penggunaan Flask dapat dilihat pada Gambar 2.3. [16]

Pada pengerjaan tugas akhir ini, Flask dipasang di sistem data adapter pada bagian DB Adapter. DB Adapter berfungsi untuk melakukan komunikasi dengan aplikasi dengan menerima beberapa permintaan seperti perubahan data, pengambilan data dan memasukan data baru. Flask menghasilkan antar muka dengan bentuk data JSON yang digunakan oleh aplikas untuk mengakses data di basis data.



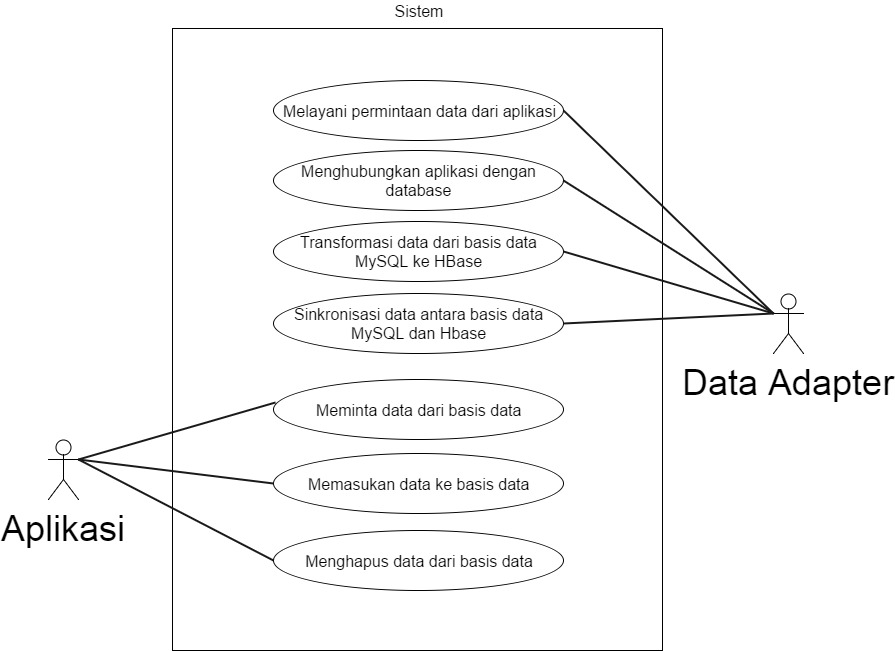
Gambar . Contoh Penggunaan Flask

# BAB III ANALISIS DAN PERANCANGAN PERANGKAT LUNAK

Pada bab ini akan dijelasakan mengenai analisis dan perancangan sistem. Perancangan di bagi menjadi perancangan arsitektur sistem dan perancangan proses utama sistem menggunakan data adapter.

## Kasus Pengguna

Terdapat dua aktor dalam diagram kasus pengguna yaitu aplikasi dan data adapter. Pada sistem, aplikasi memiliki tiga aktifitas dan data adapter memiliki 4 aktifitas yang di gambarkan pada Gambar 3.1.



Gambar . Diagram Kasus Pengguna

Diagram kasus pengguna pada Gambar 3.1 dijelaskan secara rinci pada tabel Tabel 3.1.

Tabel . Penjelasan Diagram Kasus Pengguna

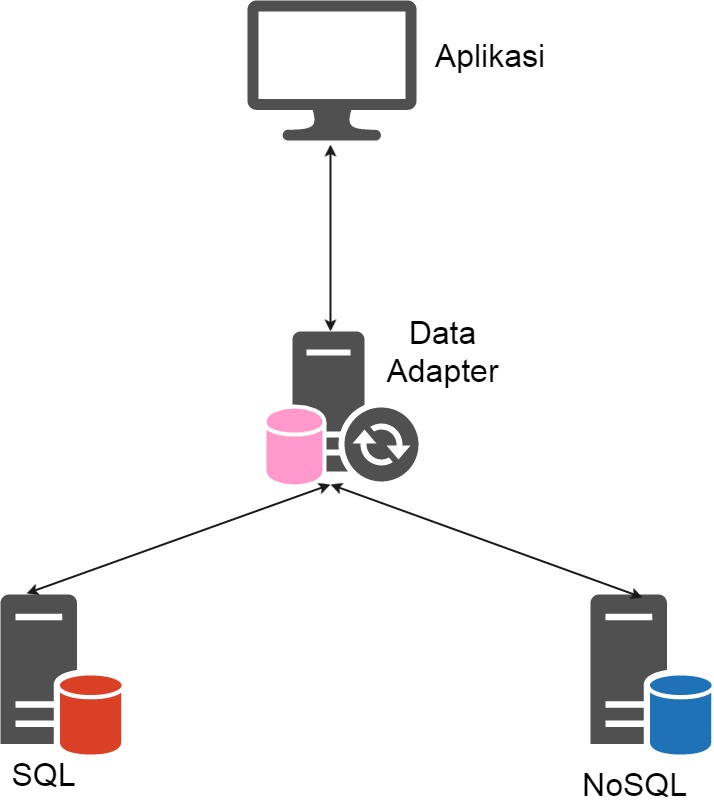
|  |  |  |  |
| --- | --- | --- | --- |
| **Kode** | **Nama Kasus Penggunaan** | **Aktor** | **Deskripsi** |
| UC01 | Melayani permintaan dari aplikasi | Data Adapter | Melayani permintaan data dari aplikasi, kemudian data adapter akan meneruskannya ke basis data |
| UC02 | Menghubungkan aplikasi dengan basis data | Data Adapter | Data adapter meneruskan permintaan aplikasi ke basis data. Data adapter menentukan dari basis data mana permintaan akan dijalankan. |
| UC03 | Mentransformasi data dari basis data MySQL ke HBase | Data Adapter | Data adapter bertugas melakukan transformasi data dari MySQL ke HBase. |
| UC04 | Melakukan sinkronisasi data antara basis data MySQL dan HBase | Data Adapter | Ketika aplikasi dijalankan oleh pengguna, data adapter akan melakukan proses sinkronisasi berdasarkan perubahan yang terjadi di basis data MySQL. |
| UC05 | Meminta data ke basis data | Aplikasi | Pengguna melalui aplikasi meminta data ke basis data. |
| UC06 | Memasukan data ke basis data | Aplikasi | Pengguna melalui aplikasi memasukan data baru ke basis data. |
| UC07 | Menghapus data dari basis data | Aplikasi | Pengguna melalui aplikasi menghapus data pada basis data. |

## Arsitektur Sistem

Sub-bab ini akan membahas mengenai analisis kebutuhan dan desain dari sistem yang akan diimplementasikan.

#### Desain Umum Sistem

Sistem yang dibangun adalah untuk mensinkronisasi dua database SQL dan No SQL. Untuk basis data tipe SQL, Database Management System (DBMS) yang digunakan adalah MySQL. Sedangkan untuk basis data tipe NoSQL, DBMS yang digunakan adalah Apache HBase. Jumlan total server yang digunakan pada penelitian ini adalah berjumlah tiga server, satu server untuk MySQL, satu server untuk Apache HBase dan satu server untuk Data Adapter. Secara umum, visualisasi arsitektur yang digunakan sistem data adapter dapat dilihat pada Gambar 3.2.

Data adapter adalah penghubung antara aplikasi dengan basis data. Data adapter berperan untuk menerima permintaan dari aplikasi dan melakukan transformasi data dari MySQL ke Apache Hbase. Komponen utama dari sistem data adapter ini terdiri dari empat, yaitu basis data relasional, basis data NoSQL, DB Adapter dan DB Converter. DB Adapter berfungsi untuk menerima permintaan dari aplikasi seperti insert data, pembaruan data, menghapus data dan mengambil data. Untuk dapat melakukan perubahan dan pengaksesan data, data adapter menyediakan *interface* yang dapat diakses oleh aplikasi yang terhubung. *Interface* yang disediakan oleh data adapter adalah dalam bentuk *Application Programming Interface* (API). *Interface*, atau antar muka, ini menghasilkan bentuk tipe data JSON yang dibuat dengan menggunakan Flask. Flask adalah kerangka kerja mikro yang berjalan menggunakan bahasa pemrograman Python. Antar muka ini dapat diakses melalui port 5000 oleh aplikasi. Dengan antar muka ini, aplikasi dapat melakukan perubahan dan pengaksesan data pada basis data.

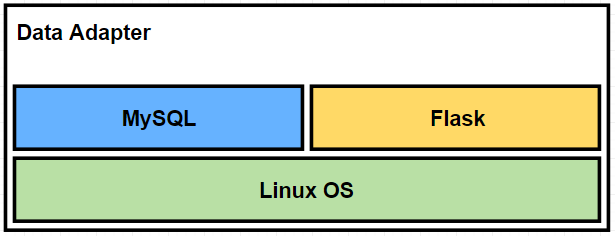
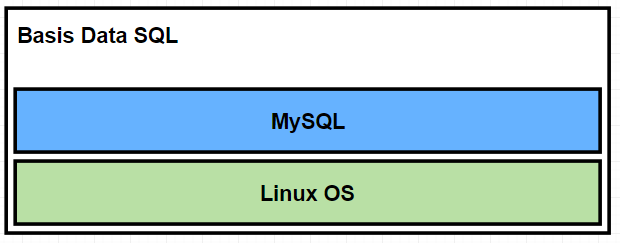
Gambar . Desain Sistem Data Adapter Secara Umum

Sementara itu, DB Converter bertanggung jawab dalam transformasi data dan pelaporan hasil transformasi yang dicatat di basis data. Transformasi data dilakukan dari basis data RDB ke basis data NoSQL. Untuk menerima permintaan dari DB Adapter, DB Converter berjalan pada port 5001. Proses transformasi ini dilakukan dengan bantuan Apache Phoenix. Pada DB Converter ini juga dipasang sebuah basis data menggunakan MySQL untuk menyimpan log sinkronisasi. Komponen utama dari data adapter ini dibangun dengan menggunakan bahasa pemrograman Python.

### Desain Data Adapter

Data adapter memiliki peran paling penting di dalam sistem. Setiap permintaan yang dilakukan oleh aplikasi akan diolah terlebih dahulu di data adapter, yang kemudian akan diteruskan ke basis data. Ada beberapa contoh antarmuka yang digunakan oleh aplikasi, dalam pengerjaan tugas akhir ini, antarmuka yang digunakan adalah API dalam bentuk JSON. Diagram arsitektur data adapter dapat dilihat pada

Data adapter terdiri dari dua komponen utama, yaitu DB Adapter dan DB Converter. DB Adapter berfungsi untuk menerima semua permintaan dari aplikasi seperti pengajuan kueri, pengambilan data, pembaruan data, dan penghapusan data. Untuk dapat melakukan hal ini, data adapter menyediakan antar muka dalam bentuk API. API menghasilkan keluaran data denga tipe data JSON. Semua komunikasi data yang dilakukan oleh aplikasi adalah melalui antar muka ini. Flask adalah kerangka kerja mikro yang digunakan untuk membuat fitur antar muka.

Sementara, DB Converter berfungsi untuk melakukan transformasi data dari basis data SQL ke basis data NoSQL. Dalam hal ini, transformasi yang dilakukan adalah dari MySQL ke Apache HBase. Secara *default*, permintaan perubahan data dari aplikasi akan diarahkan ke MySQL, sedangkan untuk permintaan pengaksesan data, permintaan dari aplikasi akan di arahkan ke HBase. Proses transformasi data ini dilakukan dengan bantuan Apache Phoenix. Apache Phoenix memungkinkan untuk menerjemahkan kueri SQL, mengkompilasinya menjadi beberapa rangkaian perintah HBase, kemudian mengeksekusinya ke HBase. Python menjadi bahasa pemrograman yang digunakan untuk mengimplementasikan semua proses yang ada pada data adapter. Gambar 3.3 menampilkan diagram arsitektur data adapter.

Gambar . Diagram Arsitektur Server Basis Data SQL

Gambar . Diagram Arsitektur Data Adapter

### Desain Basis Data SQL

Gambar . Diagram Arsitektur Basis Data NoSQL

Basis data SQL digunakan untuk menyimpan semua data aplikasi. Dalam perancangan dan pengembangan suatu aplikasi menuntut kemampuan untuk dapat diakses oleh banyak pengguna. Banyak pengembang yang memisahkan server aplikasi dengan server basis data. Pada tugas akhir ini, satu server yang terpisah dari server aplikasi digunakan sebagai basis data SQL. Aplikasi yang digunakan adalah MySQL. Semua permintaan dari aplikasi yang berkaitan dengan perubahan data, seperti *Insert, Update* dan *Delete,* dan pengambilan data dapat dieksekusi di MySQL. Diagram arsiterkur basis data SQL tertera pada Gambar 3.4.

### Desain Basis Data NoSQL

Basis Data NoSQL digunakan untuk menyimpan data dengan format yang berbeda dengan basis data NoSQL. Basis data ini adalah basis data tambahan karena jumlah data yang semakin bertambah jumlahnya. Aplikasi yang digunakan sebagai basis data NoSQL adalah Apache HBase. Untuk dapat menggunakan Apache HBase, terlebih dahulu harus memasang Apache Hadoop, sistem berkas milik Apache yang khusus menangani *big data*, pada server karena Apache Hbase berjalan diatas Apache Hadoop.

Pemasangan basis data NoSQL dilakukan pada satu server lain yang terpisah dari aplikasi dan basis data SQL. HBase akan menerima hasil transformasi data dari MySQL yang dikontrol melalui data adapter. Proses transformasi ini membutuhkan bantuan dari Apache Phoenix sebagai penerjemah kueri SQL. Desain arsitektur basis data NoSQL dapat dilihat pada Gambar 3.5.

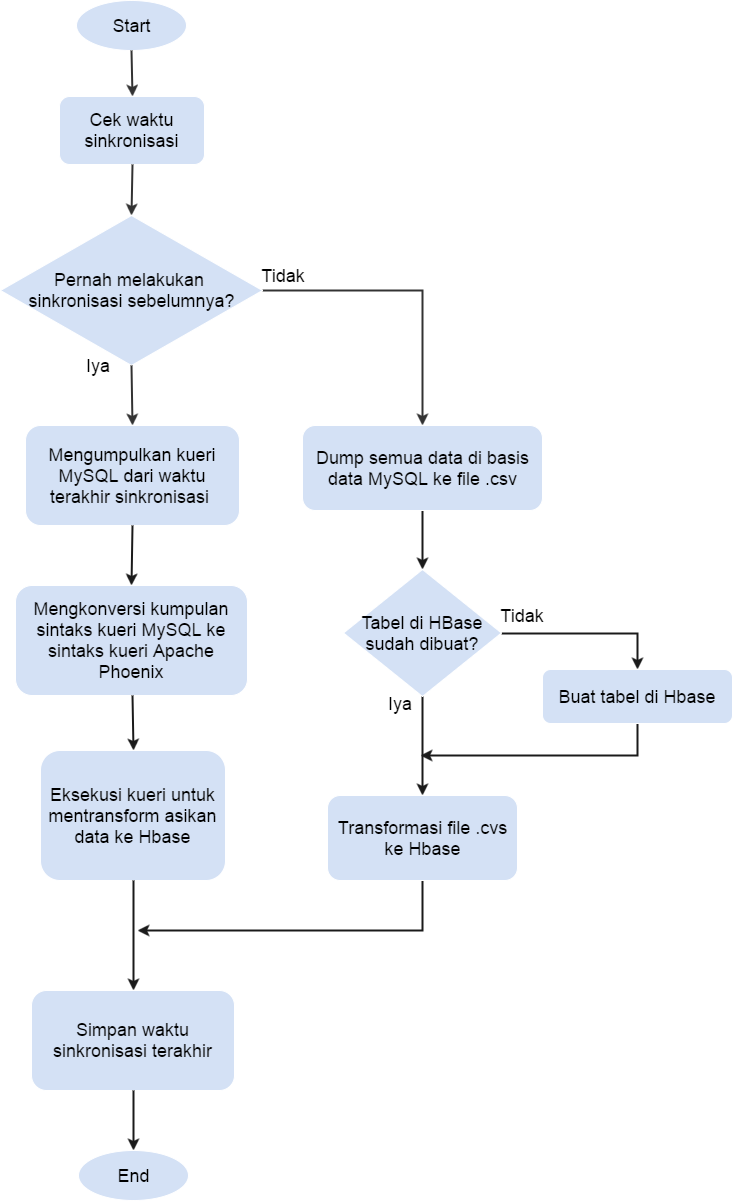
## Proses Transformasi dan Sinkronisasi

Proses transformasi dilakukan secara manual ketika user mengeksekusi aplikasi sinkronisasi. Hal pertama yang dilakukan ketika melakukan sinkronisasi adalah mengecek di basis data log sinkronisasi, apakah pernah melakukan sinkronisasi sebelumnya atau belum. Jika belum maka proses yang dilakukan adalah proses inisialisasi.

Proses inisialisasi adalah proses menyalin semua data di basis data SQL ke basis data NoSQL. Proses ini diawali dengan menge-dump atau mengekspor semua data pada tabel di MySQL menjadi file .csv. Kemudian sistem akan membuat tabel baru yang sama dengan tabel yang ada pada basis data di MySQL. Setelah selesai, proses selanjutnya adalah melakukan transformasi data ke HBase. Proses transformasi data ini ditangani dengan menggunakan bantuan Apache Phoenix.

Jika pernah melakukan sinkronisasi sebelumnya, maka inisialisasi tidak akan dilakukan. Proses yang akan dilakukan adalah sistem akan mengumpulkan semua kueri dari log pada basis data MySQL. Kueri yang diambil adalah kueri-kueri yang melakukan perubahan pada basis data yaitu perintah *insert*, *update* dan *delete*. Semua log kemudian disimpan didalam satu file .sql. Dengan bantuan Apache phoenix, semua kueri pada berkas akan ditranslasikan ke HBase satu satu, sehingga semua perubahan data yang terjadi pada basis data SQL akan terjadi juga di basis data NoSQL.

Jika proses sinkronisasi berhasil, sistem akan mencatat log dan status sinkronisasi tersebut ke basis data log sinkronisasi. Basis data yang digunakan untuk mencatat log ini menggunakan MySQL dan terpisah dari server basis data SQL, server basis data NoSQL maupun server aplikasi. Diagram alur proses sinkronisasi dapat dilihat pada Gambar 3.6.



***Gambar 3.6 Diagram Alur Proses Sinkronisasi***

# BAB IV IMPLEMENTASI

Pada bab ini akan dibahas mengenai implementasi sistem data adapter berdasarkan rancangan yang telah dijabarkan pada bab sebelumnya. Pembahasan dilakukan secara rinci untuk setiap komponen yang ada yaitu basis data SQL, basis data NoSQL dan data adapter.

## Lingkungan Implementasi

Lingkungan implementasi dan pengembangan dilakukan menggunakan komputer dengan spesifikasi Intel(R) Core(TM) i3-3240 CPU @ 3.40GHz dengan memory 6 GB di Laboratorium Arsitektur dan Jaringan Komputer, Jurusan Teknik Informatika ITS. Perangkat lunak yang digunakan dalam pengembangan adalah sebagai berikut :

* Sistem Operasi Linux Ubuntu 14.04 LTS
* Editor teks Sublime Text 2
* Python 2.7.6 untuk bahasa pemrograman utama
* Flask 0.12.1 untuk kerangka kerja pemrograman
* Virtual Environment 15.1.0 untuk lingkungan kerja pemrograman
* Apache Hadoop 2.7.3 untuk sistem berkas Apache HBase
* Apache HBase 1.2.4 untuk basis data NoSQL
* Apache Phoenix 4.10.0 untuk adapter antara MySQL dan Apache HBase
* MySQL versi 14.14 untuk basis data SQL
* Git 1.9.1 untuk pengelolaan versi program
* Postman 4.10.7 untuk uji coba API program
* Hedidi SQL 9.3.0.4984
* Microsoft Word 2013

## Rincian Implementasi Server Basis Data

Basis data digunakan untuk menyimpan semua data yang digunakan aplikasi. Pada pengerjaan Tugas Akhir ini, basis data yang akan disinkronisasikan adalah basis data SQL dan NoSQL. Sistem manajemen basis data SQL yang digunakan adalah MySQL Server dan untuk NoSQL menggunakan Apache HBase. Kedua basis data ini dipasang pada server yang terpisah. Pada sub bab ini akan dijelaskan secara rinci mengenai implementasi masing-masing basis data.

### Instalasi Server Basis Data SQL

Pada pengerjaan tugas akhir ini hanya menggunakan satu server untuk Basis data SQL. Perlu dilakukan konfigurasi lebih lanjut setelah melakukan pemasangan MySQL pada server agar data adapter dapat terhubung dengan server.

Alamat IP server diatur sesuai dengan subnet di Laboratorium Arsitektur dan Jaringan agar server data adapter dan server basis data NoSQL dapat saling terhubung dan bekomunikasi. Konfigurasi basis data agar berjalan dengan menggunakan IP yang tertera pada Kode Sumber 4.1.

Secara default, server MySQL hanya bisa diakses oleh localhost saja. Oleh karena itu perlu ditambahakn konfigurasi agar data adapter dapat mengakses server MySQL. Dalam pengerjaan tugas akhir ini, karena tingkat keamanan saat ini tidak terlalu dibutuhkan, maka perintah untuk menambahkan akses ke server diperbolehkan untuk semua alamat IP. Perintah ini dijalankan seperti dengan mengeksekusi kueri di MySQL. Terlebih dahulu harus masuk sebagai user root. Langkah-langkah menambahkan akses user lebih detailnya adalah sebagai berikut:

* 1. Masuk ke MySQL dengan menggunakan username dan password default. Jalankan perintah mysql –u root –p, kemudian tekan enter dua kali.
  2. Buat pengguna baru dari semua alamat dengan nama ‘wicak’ dan password ‘w’ dengan menjalankan perintah CREATE USER ‘wicak’@’%’ IDENTIFIED BY ‘w’;
  3. Berikan akses kepada pengguna baru yang kita buat agar pengguna yang dimaksud dapat mengakses basis data dari alamat IP lain. Jalankan perintah GRANT ALL PRIVILEGES ON \*.\* TO ‘wicak’@’%’ IDENTIFIED BY ‘w’;
  4. Muat ulang konfigurasi dengan menjalankan perintah FLUSH PRIVILEGES;

Kode Sumber . Konfigurasi MySQL pada Alamat IP AJK

|  |
| --- |
| bind-address = 10.151.36.129 |

### Instalasi Server Basis Data NoSQL

Basis data yang dipasang sebagai basis data NoSQL adalah Apache HBase. HBase dipasang pada satu server yang terpisah dengan server basis data MySQL. Untuk memasang HBase, terlebih dahulu harus melakukan instalasi Hadoop karena HBase berjalan diatas sistem berkas Hadoop.

#### Instalasi Hadoop

Pada bagian ini akan dipaparkan tahapan pemasangan sistem berkas Hadoop. Langkah-langkahnya adalah sebagai berikut :

Pastikan pada server yang digunakan telah terpasang java. Kemudian install OpenJDK dengan mengetikan perintah *apt-get install default-jdk*.

Kemudian buat pengguna untuk Hadoop. Namun terlebih dahulu buat group dengan dengan nama hadoop. Ketikan *sudo addgroup hadoop*.

Tambahkan pengguna baru pada group hadoop dengan menjalankan perintah *sudo adduser --ingroup hadoop hduser*.

Tambahkan pengguna hduser ke akses sudo dengan perintah *sudo adduser hduser sudo*.

Hadoop memerlukan SSH untuk memanajemen node-nya secara remote dan lokal. Lakukan instalasi ssh dengan menjalankan perintah *sudo apt-get install ssh*.

Hadoop menggunakan SSH (untuk mengakses node-nya) biasanya mengharuskan penggunanya untuk memasukan kata sandi. Namun, persyaratan ini bisa ditiadakan dengan membuat dan mengatur sertifikat SSH dengan menjalankan perintah *ssh-keygen –t rsa –P “.*

Selanjutnya adalah menambahkan kunci baru ke daftar kunci yang berwenang sehingga Hadoop dapat menggunakan ssh tanpa meminta password. Ketikan perintah *cat $HOME/ .ssh/id\_rsa.pub >> $HOME/ .ssh/authorized\_keys*.

Unduh berkas instalasi Hadoop dengan perintah *wget http://mirrors.sonic.net/apache/hadoop/common/stable/hadoop-2.7.3.tar.gz.*

Untuk mengekstrak berkas Hadoop jalankan tar xvzf hadoop-2.7.3.tar.gz.

Kemudian pindahkan semua berkas instalasi ke direktori */usr/local/hadoop*.

Selanjutnya adalah melakukan konfigurasi pada Hadoop. Yang pertama adalah melakukan konfigurasi pada berkas bashrc dengan menjalankan perintah *vim ~/.bashrc*. Konfigurasi yang ditambahkan pada berkas .bashrc dapat dilihat pada Kode Sumber 4.1

Tambahkan variabel JAVA\_HOME kedalam berkas hadoop-env.sh agar ketika Hadoop dijalankan maka nilai JAVA\_HOME tersedia. Ketikkan perintah *vim /usr/local/hadoop/etc/hadoop/hadoop-env.sh* untuk merubahnya. Isi konfigurasi variable ini dapat dilihat pada Kode Sumber 4.3.

Berkas /usr/local/hadoop/etc/hadoop/core-site.xml berisi konfigurasi yang digunakan Hadoop ketika mengawali proses. Buka dan sunting berkas tersebut dengan perintah *vi /usr/local/hadoop/etc/hadoop/core-site.xml*. Isi konfigurasi dapat dilihat pada Kode Sumber 4.4.

Salin dan *rename* /usr/local/hadoop/etc/hadoop/mapred-site.xml.template menjadi mapred-site.xml. Kemudian tambahkan konfigurasi seperti pada Kode Sumber 4.5.

Berkas /usr/local/hadoop/etc/hadoop/hdfs-site.xml diatur untuk setiap *host* kluster yang digunakan. Ubah berkas hdfs-site.xml menjadi seperti pada Kode Sumber 4.6. Sebelum menyuntingnya, buat dua direktori baru yang akan terdiri *namenode* dan *datanode* Hadoop. Buat direktori dengan menjalankan perintah *mkdir –p /usr/local/hadoop-store/hdfs/namenode* dan *mkdir –p /usr/local/hadoop-store/hdfs/datanode.*

Sistem berkas Hadoop perlu di format untuk dapat digunakan. Jalankan perintah *hadoop namenode –format*.

Hadoop telah siap digunakan. Masuk ke direktori /usr/local/hadoop/sbin. Kemudian gunakan *start-all.sh* untuk memulai Hadoop. Tampilan Web UI Hadoop dapat dilihat dengan mengakses *http://localhost:50070* pada browser.

Kode Sumber . Konfigurasi pada Berkas .bashrc

|  |
| --- |
| # HADOOP VARIABLES START  export JAVA\_HOME=/usr/lib/jvm/java-1.7.0-openjdk-amd64  export HADOOP\_INSTALL=/usr/local/hadoop  export PATH=$PATH:$HADOOP\_INSTALL/bin  export PATH=$PATH:$HADOOP\_INSTALL/sbin  export HADOOP\_MAPRED\_HOME=$HADOOP\_INSTALL  export HADOOP\_COMMON\_HOME=$HADOOP\_INSTALL  export HADOOP\_HDFS\_HOME=$HADOOP\_INSTALL  export YARN\_HOME=$HADOOP\_INSTALL  export HADOOP\_COMMON\_LIB\_NATIVE\_DIR=$HADOOP\_INSTALL/lib/native  export HADOOP\_OPTS="-Djava.library.path=$HADOOP\_INSTALL/lib" |

Kode Sumber . Variable JAVA\_HOME pada hadoop-env.sh

|  |
| --- |
| export JAVA\_HOME=${JAVA\_HOME}  export JAVA\_HOME=/usr/lib/jvm/java-7-openjdk-amd64 |

Kode Sumber . Konfigurasi pada Berkas core-site.xml

|  |
| --- |
| <configuration>  <property>  <name>hadoop.tmp.dir</name>  <value>/app/hadoop/tmp</value>  <description>A base for other temporary directories.</description>  </property>  <property>  <name>fs.default.name</name>  <value>hdfs://localhost:54310</value>  <description>The name of the default file system. A URI whose scheme and authority determine the FileSystem implementation. The uri's scheme determines the config property (fs.SCHEME.impl) naming the FileSystem implementation class. The uri's authority is used to determine the host, port, etc. for a filesystem.  </description>  </property>  </configuration> |

Kode Sumber . Konfigurasi pada Berkas mapred-site.xml

|  |
| --- |
| <configuration>  <property>  <name>mapred.job.tracker</name>  <value>localhost:54311</value>  <description>The host and port that the MapReduce job tracker runs at. If "local", then jobs are run in-process as a single map and reduce task.  </description>  </property>  <property>  <name>mapreduce.framework.name</name>  <value>yarn</value>  </property>  </configuration> |

Kode Sumber . Konfigurasi pada Berkas hdfs-site.xml

|  |
| --- |
| <configuration>  <property>  <name>dfs.replication</name>  <value>1</value>  <description>Default block replication. The actual number of replications can be specified when the file is created. The default is used if replication is not specified in create time.  </description>  </property>  <property>  <name>dfs.namenode.name.dir</name>  <value>file:/usr/local/hadoop\_store/hdfs/namenode</value>  </property>  <property>  <name>dfs.datanode.data.dir</name>  <value>file:/usr/local/hadoop\_store/hdfs/datanode</value>  </property>  </configuration> |

#### Instalasi Apache HBase

Setelah melakukan instalasi sistem berkas Hadoop, dapat dilakukan instalasi Apache HBase. Apache HBase yang digunakna adalah versi 1.2.4. Pada sub bab ini akan dijelaskan mengenai cara pemasangan Apache HBase.

* + - 1. Unduh paket Apache Hbase dengan menjalankan perintah *wget https://archive.apache.org/dist/hbase/1.2.4/hbase-1.2.4-bin.tar.gz*
      2. Unzip berkas dengan menjalankan tar –xvf hbase-1.2.4-bin.tar.gz.
      3. Buat direktori hbase dengan menjalankan perintah *mkdir /usr/lib/hbase.* Pindahkan folder hasi ekstrasi kedalam folder ini.
      4. Atur path java yang ada pada server pada berkas hbase-env.sh. Ubah dengan menjalankan perintah vim /usr/lib/hbase/hbase-1.2.4/conf/hbase-env.sh dan tambahkan konfigurasi seperti pada
      5. Tambahkan juga path HBASE\_HOME pada berkas .bashrc. Tambahkan konfigurasi seperti pada dengan menjalan perintah *vim ~/.bashrc*.
      6. Selanjutnya adalah mengubah berkas konfigurasi pada HBase pada berkas hbase-site.xml untuk mengatur direktori tempat HBase akan menyimpan data. Menunjukan isi konfigurasi hbase-site.xml
      7. Untuk memulai HBase masuk ke direktori /usr/lib/hbase/hbase-1.2.4/bin kemudian jalankan start-hbase.sh.
      8. Untuk mengecek apakah HBase telah berjalan, jalankan perintah *hbase shell.*

## Rincian Implementasi Sistem Data Adapter

Komponen utama pada sistem data adapter yaitu DB Adapter dan DB Converter. Sistem data adapter dibangun dengan menggunakan bahasa pemrograman Python dan kerengka kerja pemrograman Flask. Selain itu, digunakan beberapa modul tambahan yang dipasang secara terpisah dengan menggunakan pip pada terminal. Pip adalah sebuah perkakas untuk melakukan pemasangan modul-modul paket Python. Pada sub bab ini akan dijelaskan secara rinci mengenai implementasi data adapter.

### Instalasi Paket

Seperti yang dijelaskan sebelumnya, pemasangan modul-modul Python ini dilakukan dengan menggunakan pip. Cara kerja pip ini adalah mengunduh modul untuk sistem yang dibutuhkan, kemudian dipasang pada server. Semua modul yang dibutuhkan sistem disimpan dalam berkas requirement\_interface.txt. Untuk menjalankan pemasangan modul di dalam file tersebut dapat dijalankan dengan menggunakan perintah pip install –r requirement\_interface.txt. Semua daftar modul di dalam berkas tersebut akan dipasang di server. Modul-modul tambahan yang dipasang yaitu :

* Flask 0.12.1
* Appdirs 1.4.3
* Click 6.7
* Itsdangerous 0.24
* Jinja2 2.9.6
* MarkupSafe 1.0
* MySQL-python 1.2.5
* Packaging 16.8
* Protobuf 3.2.0
* Pyparsing 2.2.0
* Six 1.10.0
* Werkzeug 0.12.1

### Implementasi DB Adapter

DB Adapter adalah komponen pada sistem data adapter yang bertugas sebagai penghubung antara aplikasi dengan basis data. DB Adapter menyediakan antar muka yang dapat diakses oleh aplikasi dimana antarmuka ini dapat menerima permintaan yang dilakukan oleh aplikasi seperti pengajuan kueri, pengambilan data, pembaruan data, dan penghapusan data. Semua permintaan ini diterima oleh DB Adapter dengan menggunakan API yang dibuat dengan menggunakan Flask. Flask yang digunakan adalah versi 0.12.1. Antar muka ini menghasilkan keluaran data dalam bentuk JSON.

Untuk melakukan permintaan atau proses data, DB Adapter menyediakan antar muka yang diakses oleh aplikasi dengan melakukan permintaan ke rute yang disediakan. Dalam pengerjaan Tugas Akhir ini, beberapa rute yang disediakan beserta dengan rincian penjelasannya dapat dilihat pada Tabel 4.1. Rute ini diimplementasikan dalam sebuah berkas python.

Tabel . Implementasi Rute Pada DB Adapter

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **No** | **Rute** | **Masukan** | **Luaran** | **Proses** |
| 1 | GET /sinkron |  | status, message | Melakukan sinkronisasi basis data MySQL ke HBase |
| 2 | POST /insert\_routes | data route | status, message | Memasukan data baru ke tabel routes |
| 3 | POST /delete\_routes\_by\_id | id\_route | status, message | Menghapus satu baris data data pada tabel routes |
| 4 | POST /update\_routes\_by\_id | id\_route | status, message | Melakukan perbaruan data pada tabel routes |
| 5 | GET /select\_all\_routes |  | host, database, rows, flight\_routes | Mengambil semua data pada tabel routes |
| 6 | GET /select\_route\_by\_id | id\_route | host, database, route | Mengambil satu baris data pada tabel route yang memiliki id route <id> |
| 7 | POST /insert\_airline | Data\_airline | Status, message | Memasukan data baru ke tabel airline |
| 8 | POST /delete\_airline\_by\_id |  |  |  |
| 9 | POST /update\_airline\_by\_id |  |  |  |
| 10 | GET /select\_all\_airline |  |  |  |
| 11 | GET /select\_airline\_by\_id |  |  |  |
| 12 | POST /insert\_airport |  |  |  |
| 13 | POST /delete\_airport\_by\_id |  |  |  |
| 14 | POST /update\_airport\_by\_id |  |  |  |
| 15 | GET /select\_all\_airport |  |  |  |
| 16 | GET /select\_airport\_by\_id |  |  |  |

Untuk melakukan pengolahan data, terdapat kontroler di tiap-tiap di rute yang berfungsi untuk memilih pada basis data mana kueri akan dieksekusi. Pseudocode kontroler untuk proses perubahan data ditunjukan pada Pseudocode 4.1. Proses pengambilan data ditunjukan pada

|  |  |
| --- | --- |
| 1  2  3  4 | **post** data from application  connect to MySQL DB  **excecute** query  **return** result |

Pseudocode . Eksekusi Kueri untuk Proses Perubahan Data

Penjelasan Pseudocode 4.1 adalah sebagai berikut :

Mengambil data post dari aplikasi

Jika aksi yang dilakukan aplikasi adalah perubahan pada basis data, seperti menambahkan, menghapus dan menyunting, maka DB Adapter akna mengarahkan kueri ke basis data MySQL.

Proses menjalankan sintaks kueri

Mengembalikan hasil eksekusi kueri

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7 | **get** status synchroniztion  **if** syncrhonization process still working  connect to MySQL DB  **else**  connect to HBase  result = **excecute** query  **return** result |

Pseudocode . Eksekusi Kueri untuk Proses Pengambilan Data

Berikut adalah penjelasan Pseudocode 4.2 :

1. Pengeksekusian dimulai dari mengecek status apakah sedang berlangsung proses sinkronisasi atau tidak
2. Cek jika proses sinkronisasi sedang berlangsung, maka proses pengambilan data diarahkan ke basis data MySQL. Jika tidak maka diarahakn ke HBase.
3. Proses menjalankan sintaks kueri.
4. Mengembalikan hasil kueri.

### Implementasi DB Converter

DB Converter adalah komponen lain data adapter yang berfungsi untuk mencatat log dan status sinkronisasi serta melakukan transformasi data dari basis data SQL ke basis data NoSQL. Komponen yang terpasang pada DB Converter ini diantaranya Flask dan Apache Phoenix.

Flask digunakan untuk membuat antar muka yang berfungsi untuk menerima permintaan proses sinkronisasi dari DB Adapter. Apache Phoenix berfungsi sebagai lapisan layer yang bekerja diatas HBase yang dapat menerjemahkan SQL dan mengeksekusinye ke HBase. Terdapat basis data yang menggunakan MySQL untuk mencatat log dan status setiap kali menjalankan proses sinkronisasi.

Proses sinkronisasi basis data SQL ke NoSQL dari MySQL ke Apache HBase ditunjukan pada Pseudocode 4.3. Penjelasan secara rinci dari Pseudocode 4.3 adalah sebagai berikut:

1. Proses sinkronisasi diawali dengan mengambil data-data server dari berkas konfigurasi. Data-data ini diambil dari berkas configuration.ini.
2. Atur inisialisasi

Ketika DB Adapter menerima permintaan sinkronisasi, DB Adapter kemudian meneruskan permintaan tersebut dengan meminta DB Converter untuk melakukan sinkronisasi. DB Converter akan mengambil daftar kueri dari log kueri dari server MySQL. Semua daftar kueri, yang akan dieksekusi ke HBase, disimpan dalam satu berkas dengan naman list\_all\_query.sql. Selanjutnya, Apache Phoenix akan menjalankan semua kueri yang ada dalam berkas tersebut dan dieksekusi ke basis data HBase.

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25  26  27  28  29  30  31 | **get** mysql\_db\_conf\_data  **get** hbase\_db\_conf\_data  **get** sync\_log\_db\_conf\_data  **get** ssh\_access\_data  **get** last\_sync\_log\_data  **set** initial\_sync\_log  **set** start\_time  query\_count = 0  do\_patching = false  **if** (last\_sync\_log\_data is empty)  dump mysql\_table to mysql\_table.csv  **execute** create\_table.sql with apache phoenix  **if** (create\_table is success)  **execute** mysql\_table.csv with apache phoenix  **else**  ssh to mysql\_db\_server  get log\_query\_file  **for** line **in** log\_query\_file  **get** query\_time  **if** (query\_time **>** last\_sync\_log\_data)  get query\_syntax  query\_tmp = convert query\_syntax to apache phoenix sintax  append query\_tmp to list\_all\_query  do\_patching = true  **if** (query\_count **is** 0)  update sync\_log  **if** (do\_patching **is** true)  **execute** list\_all\_query with apache phoenix  update sync\_log  duration = time\_now – start\_time |

Pseudocode . Proses sinkronisasi

# 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

***[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

Dari hasil uji coba yang telah dilakukan dapat diambil beberapa kesimpulan sebagai berikut:

1. .

## Saran

Saran yang diberikan untuk pengembangan aplikasi ini adalah:

1. .

# DAFTAR PUSTAKA

|  |  |
| --- | --- |
| [1] | 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. |
| [2] | A. Levitin, “Dynamic Programming,” dalam *Introduction to The Design and Analysis of Algorithms 3rd Edition*, United States of America, Pearson Education, Inc., 2012, p. 311. |
| [3] | D. Abraham dan N. S. Raj, “Approximate String Matching Algorithm for Phishing Detection,” dalam *International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, New Delhi, 2014. |
| [4] | M. Elhadi dan A. Al-Tobi, “Duplicate Detection in Documents and WebPages using Improved Longest Common Subsequence and Documents Syntactical Structures,” dalam *Fourth International Conference on Computer Sciences and Convergence Information Technology*, Seoul, 2009. |
| [5] | A. Grama, A. Gupta, G. Karypis dan V. Kumar, Introduction to Parallel Computing, Second Edition, Addison Wesley, 2003, p. 534. |
| [6] | 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. |
| [7] | “Genetics Home Reference,” National Library of Medicine (US), 14 June 2016. [Online]. Available: https://ghr.nlm.nih.gov/primer/hgp/genome. [Diakses 16 June 2016]. |
| [8] | “What is DNA?,” National Library Medicine (US), 14 June 2016. [Online]. Available: http://ghr.nlm.nih.gov/handbook/basics/dna. [Diakses 15 June 2016]. |
| [9] | “The Science Behind the Human Genome Project: Understanding the Basics,” U.S. Department of Energy Human Genome Project, [Online]. Available: http://web.ornl.gov/sci/techresources/Human\_Genome/project/info.shtml. [Diakses 15 June 2015]. |
| [10] | “What is a gene?,” National Library of Medicine (US), 14 June 2016. [Online]. Available: https://ghr.nlm.nih.gov/primer/basics/gene. [Diakses 15 June 2016]. |
| [11] | “What is Heredity?,” Genetic Science Learning Center, 22 June 2014. [Online]. Available: http://learn.genetics.utah.edu/content/inheritance/intro/. [Diakses 15 June 2016]. |
| [12] | “NCBI National Center for Biotechnology Information,” [Online]. Available: http://www.ncbi.nlm.nih.gov/. [Diakses 21 June 2015]. |
| [13] | “Open MPI: Open Source High Performance Computing,” [Online]. Available: http://www.open-mpi.org/. [Diakses 21 June 2015]. |
| [14] | “MPI: A Message-Passing Interface Standard,” University of Tennessee, Knoxville, 2015. |
| [15] | GeeksforGeeks, “Printing Longest Common Subsequence,” [Online]. Available: http://www.geeksforgeeks.org/printing-longest-common-subsequence/. [Diakses 11 May 2016]. |

**LAMPIRAN A**

# KODE SUMBER

Pada lampiran berikut dijelaskan mengenai kode sumber yang digunakan. Kode sumber menggunakan Bahasa Pemrograman C++. Berikut kode sumber yang dicantumkan:

1. Kode Sumber 1 Fungsi Perhitungan LCS Paralel merupakan kode sumber dari prosedur perhitungan *LCS* paralel.
2. Kode Sumber 2 Fungsi Mencetak LCS merupakan kode sumber dari prosedur mencetak karakter *LCS* dari banyak proses.
3. Kode Sumber 3 Pemilihan Indeks Perhitungan merupakan potongan kode sumber bagian dari Kode Sumber 1 untuk memilih indeks dalam perhitungan *LCS* agar tidak terjadi perulangan perhitungan.
4. Kode Sumber 4 Pemilihan Indeks Kolom Untuk Menerima Data merupakan potongan kode sumber bagian dari Kode Sumber 1 untuk memilih kolom dalam penerimaan data ketika *proses* > 0.
5. Kode Sumber 5 Pengambilkan Karakter LCS merupakan potongan kode sumber bagian dari Kode Sumber 2 dalam mengambil karakter *LCS*.
6. Kode Sumber 6 Perhitungan Nilai Indeks LCS merupakan potongan kode sumber bagian dari Kode Sumber 1 dalam menghitung nilai indeks *LCS*.

Kode Sumber . Fungsi Perhitungan LCS Paralel

|  |
| --- |
| 1. **for** (k = 1; k <= numberOfProcess / alpa; k++) { 2. **if** (process == 0) { 3. **int** rowStart = ((alpa \* (k - 1) \* lengthOfX) / numberOfProcess); 4. **int** rowEndPrevProcess = ((alpa \* (k - 2) \* lengthOfX) / numberOfProcess); 5. **int** rowEnd = (alpa \* k \* lengthOfX) / numberOfProcess; 7. **if** (rowStart != 0) rowStart = rowStart + 1; 9. **if** (k != 1 && rowEnd == 0 && rowStart == 0) { 10. rowStart = rowStart + 1; 11. } 12. **if** (k != 1 && rowEndPrevProcess == 0 && rowStart == 0) { 13. rowStart = rowStart + 1; 14. } 16. **int** coloumnStart = (process) \* lengthOfY / numberOfProcess; 17. **if** (coloumnStart != 0) coloumnStart = coloumnStart + 1; 19. **int** coloumnEnd = (process + 1) \* lengthOfY / numberOfProcess; 20. **for** (rowCalculate = rowStart; rowCalculate <= rowEnd; rowCalculate++) { 21. numberOf*LCS*Character[rowCalculate].resize(sizeForY); 22. **for** (coloumnCalculate = coloumnStart; coloumnCalculate <= coloumnEnd; coloumnCalculate++) { 24. **if** (rowCalculate == 0 || coloumnCalculate == 0) { 25. numberOf*LCS*Character[rowCalculate][coloumnCalculate] = 0; 26. } **else** **if** (X[rowCalculate - 1] == Y[coloumnCalculate - 1]) { 27. numberOf*LCS*Character[rowCalculate][coloumnCalculate] =  numberOf*LCS*Character[rowCalculate - 1][coloumnCalculate - 1] + 1; 28. } **else** { 29. numberOf*LCS*Character[rowCalculate][coloumnCalculate] =  max(numberOf*LCS*Character[rowCalculate -1][coloumnCalculate],  numberOf*LCS*Character[rowCalculate][coloumnCalculate - 1]); 30. } 31. } 33. } 34. dest = process + 1; 35. **int** rowSendProcess0, coloumnSendProcess0; 36. **for** (rowSendProcess0 = rowStart; rowSendProcess0 <= rowEnd; rowSendProcess0++) { 37. coloumnSendProcess0 = (dest \* lengthOfY / numberOfProcess); 39. MPI\_Send( & numberOf*LCS*Character[rowSendProcess0][coloumnSendProcess0], 1, MPI\_INT, dest,  0, MPI\_COMM\_WORLD); 40. } 42. } **else** **if** (process > 0) { 43. **int** finish = 0; 44. **int** source = process - 1; 45. dest = process + 1; 46. **int** coloumnReceive, rowReceive, coloumnSend, rowSend; 47. **int** finishReceived = 0; 49. **int** firstColoumn = ((process) \* lengthOfY / numberOfProcess); 50. **int** rowStart = ((alpa \* (k - 1) \* lengthOfX) / numberOfProcess); 51. **int** rowEndPrevProcess = ((alpa \* (k - 2) \* lengthOfX) / numberOfProcess); 53. **int** coloumnStart = (process) \* lengthOfY / numberOfProcess; 54. **int** coloumnEnd = (process + 1) \* lengthOfY / numberOfProcess; 56. **int** rowEnd = (alpa \* k \* lengthOfX) / numberOfProcess; 57. **int** coloumnPreviousProcess = (process - 1) \* lengthOfY / numberOfProcess; 58. **if** (rowStart != 0) { 59. rowStart = rowStart + 1; 60. } 61. **if** (k != 1 && rowEnd == 0 && rowStart == 0) rowStart = rowStart + 1; 62. **if** (k != 1 && rowStart == 0 && rowEndPrevProcess == 0) rowStart = rowStart + 1; 64. **if** (coloumnStart != 0) { 65. coloumnStart = coloumnStart + 1; 66. } 68. **int** colStartForProcess1; 69. colStartForProcess1 = (1 \* lengthOfY / numberOfProcess); 70. **if** ((process \* lengthOfY / numberOfProcess) != colStartForProcess1) { 71. coloumnReceive = firstColoumn - coloumnPreviousProcess; 72. } **else** { 73. coloumnReceive = coloumnStart - 1; 74. } 76. **for** (rowReceive = rowStart; rowReceive <= rowEnd; rowReceive++) { 77. numberOf*LCS*Character[rowReceive].resize(sizeForY); 78. MPI\_Recv( & numberOf*LCS*Character[rowReceive][coloumnReceive], 1, MPI\_INT, source, 0,  MPI\_COMM\_WORLD, & status); 79. numberOf*LCS*Character[rowReceive][coloumnReceive - coloumnReceive] =  numberOf*LCS*Character[rowReceive][coloumnReceive]; 80. finishReceived = 1; 81. } 82. **if** (finishReceived == 1) { 83. **for** (rowCalculate = rowStart; rowCalculate <= rowEnd; rowCalculate++) { 84. numberOf*LCS*Character[rowCalculate].resize(sizeForY); 85. **for** (coloumnCalculate = coloumnStart; coloumnCalculate <= coloumnEnd;  coloumnCalculate++) { 86. **if** (rowCalculate == 0 || coloumnCalculate == 0) { 87. numberOf*LCS*Character[rowCalculate][coloumnCalculate - firstColoumn] = 0; 88. } **else** **if** (X[rowCalculate - 1] == Y[coloumnCalculate - 1]) { 89. numberOf*LCS*Character[rowCalculate][coloumnCalculate - firstColoumn] =  numberOf*LCS*Character[rowCalculate - 1][coloumnCalculate - 1 - firstColoumn] + 1; 90. } **else** { 91. numberOf*LCS*Character[rowCalculate][coloumnCalculate - firstColoumn] =  max(numberOf*LCS*Character[rowCalculate - 1][coloumnCalculate - firstColoumn],  numberOf*LCS*Character[rowCalculate][coloumnCalculate - 1 - firstColoumn]); 92. } 93. rowIndex = rowCalculate; 94. coloumnIndex = coloumnCalculate - firstColoumn; 95. } 97. }; 98. coloumnSend = coloumnEnd - firstColoumn; 99. **if** (process != numberOfProcess - 1) { 100. **for** (rowSend = rowStart; rowSend <= rowEnd; rowSend++) { 101. **http://www.fileformat.info/info/unicode/font/arial_unicode_ms/u21AA.png**                    MPI\_Send( & numberOf*LCS*Character[rowSend][coloumnSend], 1, MPI\_INT, dest, 0,  MPI\_COMM\_WORLD); 102. } 103. } 104. } 105. } 106. finishedCalculating = 1; |

Kode Sumber . Fungsi Mencetak LCS

|  |
| --- |
| 1. **if** (finishedCalculating == 1) { 2. **if** (process == numberOfProcess - 1) { 3. **int** destBacktrack, coloumnStartBacktrack, coloumnEndBacktrack; 4. point valueForBacktrack; 5. **int** firstColoumn = ((process)\*lengthOfY / numberOfProcess); 7. destBacktrack = process - 1; 9. coloumnStartBacktrack =   (((process)\*lengthOfY / numberOfProcess) + 1) - firstColoumn; 10. coloumnEndBacktrack =   ((process + 1) \* lengthOfY / numberOfProcess) - firstColoumn; 11. **int** index = numberOf*LCS*Character[rowIndex][coloumnIndex]; 12. **int** initialIndex*LCS* = index - 1; 13. **int** lengthOf*LCS* = index; 15. vector<**char**> character*LCS*; 16. character*LCS*.resize(index); 18. **int** rowBacktrack = rowIndex, coloumnBacktrack = coloumnIndex; 19. **int** originalColoumnIndexY = coloumnBacktrack + firstColoumn; 20. **while** (coloumnBacktrack >= coloumnStartBacktrack  && coloumnBacktrack <= coloumnEndBacktrack) { 21. **if** (X[rowBacktrack - 1] == Y[originalColoumnIndexY - 1]) { 22. character*LCS*[index - 1] = X[rowBacktrack - 1]; 23. originalColoumnIndexY--; 24. rowBacktrack--; 25. coloumnBacktrack--; 26. index--; 27. } 28. **else** **if** (numberOf*LCS*Character[rowBacktrack - 1][coloumnBacktrack] > numberOf*LCS*Character[rowBacktrack][coloumnBacktrack - 1]) { 29. rowBacktrack--; 30. } 31. **else** { 32. coloumnBacktrack--; 33. originalColoumnIndexY--; 34. } 35. } 36. valueForBacktrack.rowForBacktrack = rowBacktrack; 37. valueForBacktrack.coloumnForBacktrack = coloumnBacktrack; 38. valueForBacktrack.index*LCS*ForBacktrack = index; 39. valueForBacktrack.originalColoumnBacktrack = originalColoumnIndexY; 40. **int** endIndex*LCS* = index; 42. MPI\_Send(&valueForBacktrack, 4, MPI\_INT, destBacktrack, 3, MPI\_COMM\_WORLD); 44. MPI\_Send(&initialIndex*LCS*, 1, MPI\_INT, 0, 4, MPI\_COMM\_WORLD); 45. MPI\_Send(&endIndex*LCS*, 1, MPI\_INT, 0, 4, MPI\_COMM\_WORLD); 47. MPI\_Send(&lengthOf*LCS*, 1, MPI\_INT, 0, 5, MPI\_COMM\_WORLD); 48. **for** (**int** sendCharacter = endIndex*LCS*; sendCharacter <= initialIndex*LCS*;  sendCharacter++) { 49. MPI\_Send(&character*LCS*[sendCharacter], 1, MPI\_CHAR, 0, 4, MPI\_COMM\_WORLD); 50. } 51. } 52. **if** (process < numberOfProcess - 1) { 54. **int** sourceBacktrackProcess, destBacktrackProcess; 55. **int** coloumnStartBacktrack, coloumnEndBacktrack; 56. **int** firstColoumn = ((process)\*lengthOfY / numberOfProcess); 57. point valueForBacktrack; 59. sourceBacktrackProcess = process + 1; 60. destBacktrackProcess = process - 1; 62. MPI\_Recv(&valueForBacktrack, 4, MPI\_INT, sourceBacktrackProcess, 3,  MPI\_COMM\_WORLD, &status); 63. **if** (process != 0) { 64. coloumnStartBacktrack =   (((process)\*lengthOfY / numberOfProcess) + 1) - firstColoumn; 65. coloumnEndBacktrack =   ((process + 1) \* lengthOfY / numberOfProcess) - firstColoumn; 66. **int** rowBacktrack, coloumnBacktrack, indexOf*LCS*; 67. **int** initialIndex*LCS*, endIndex*LCS*, originalColoumnIndexY; 68. rowBacktrack = valueForBacktrack.rowForBacktrack; 69. coloumnBacktrack = coloumnEndBacktrack; 70. indexOf*LCS* = valueForBacktrack.index*LCS*ForBacktrack; 71. originalColoumnIndexY = valueForBacktrack.originalColoumnBacktrack; 73. vector<**char**> character*LCS*; 74. character*LCS*.resize(indexOf*LCS*); 75. initialIndex*LCS* = indexOf*LCS* - 1; 76. **int** backtrack = 0; 77. **while** (coloumnBacktrack >= coloumnStartBacktrack &&  coloumnBacktrack <= coloumnEndBacktrack) { 78. **if** (indexOf*LCS* <= 0) 79. **break**; 80. **if** (X[rowBacktrack - 1] == Y[originalColoumnIndexY - 1]) { 81. character*LCS*[indexOf*LCS* - 1] = X[rowBacktrack - 1]; 82. rowBacktrack--; 83. coloumnBacktrack--; 84. indexOf*LCS*--; 85. originalColoumnIndexY--; 86. } 87. **else** **if** (numberOf*LCS*Character[rowBacktrack - 1][coloumnBacktrack] >  numberOf*LCS*Character[rowBacktrack][coloumnBacktrack - 1]) { 88. rowBacktrack--; 89. } 90. **else** { 91. coloumnBacktrack--; 92. originalColoumnIndexY--; 93. } 94. valueForBacktrack.rowForBacktrack = rowBacktrack; 95. valueForBacktrack.coloumnForBacktrack = coloumnBacktrack; 96. valueForBacktrack.index*LCS*ForBacktrack = indexOf*LCS*; 97. valueForBacktrack.originalColoumnBacktrack = originalColoumnIndexY; 98. endIndex*LCS* = indexOf*LCS*; 99. backtrack = 1; 100. } 101. MPI\_Send(&valueForBacktrack, 4, MPI\_INT, destBacktrackProcess, 3,  MPI\_COMM\_WORLD); 103. **char** flag[] = { '\0' }; 104. **int** flag\_notcount = 1; 105. **if** (backtrack == 1) { 106. MPI\_Send(&initialIndex*LCS*, 1, MPI\_INT, 0, 4, MPI\_COMM\_WORLD); 107. MPI\_Send(&endIndex*LCS*, 1, MPI\_INT, 0, 4, MPI\_COMM\_WORLD); 108. **for** (**int** sendCharacter = endIndex*LCS*; sendCharacter <= initialIndex*LCS*;  sendCharacter++) { 109. MPI\_Send(&character*LCS*[sendCharacter], 1, MPI\_CHAR, 0, 4,  MPI\_COMM\_WORLD); 110. } 111. } 112. **if** (backtrack == 0) { 113. initialIndex*LCS* = -1; 114. endIndex*LCS* = -1; 115. MPI\_Send(&initialIndex*LCS*, 1, MPI\_INT, 0, 4, MPI\_COMM\_WORLD); 116. MPI\_Send(&endIndex*LCS*, 1, MPI\_INT, 0, 4, MPI\_COMM\_WORLD); 117. } 118. } 119. **if** (process == 0) { 120. **int** initialIndex*LCS*, endIndex*LCS*, lengthOf*LCS*; 121. **int** coloumnBacktrack, rowBacktrack; 123. MPI\_Recv(&lengthOf*LCS*, 1, MPI\_INT, numberOfProcess - 1,  5, MPI\_COMM\_WORLD, &status); 124. vector<**char**> characterOf*LCS*; 125. characterOf*LCS*.resize(lengthOf*LCS*); 127. characterOf*LCS*[lengthOf*LCS*] = '\0'; 128. **int** finishReceived*LCS* = 0, finishBacktrackAndReceived = 0,  finishBacktrackProcess0 = 0; 129. **int** firstColoumn = ((process)\*lengthOfY / numberOfProcess); 130. coloumnStartBacktrack =   (((process)\*lengthOfY / numberOfProcess)) - firstColoumn; 131. coloumnEndBacktrack =   ((process + 1) \* lengthOfY / numberOfProcess) - firstColoumn; 133. **int** index*LCS*, originalColoumnIndexY; 135. rowBacktrack = valueForBacktrack.rowForBacktrack; 136. coloumnBacktrack = coloumnEndBacktrack; 137. index*LCS* = valueForBacktrack.index*LCS*ForBacktrack; 138. originalColoumnIndexY = valueForBacktrack.originalColoumnBacktrack; 139. **while** (coloumnBacktrack >= coloumnStartBacktrack &&  coloumnBacktrack <= coloumnEndBacktrack) { 140. **if** (rowBacktrack - 1 < 0 || coloumnBacktrack - 1 < 0 ||  originalColoumnIndexY - 1 < 0 || index*LCS* <= 0) 141. **break**; 142. **else** **if** (X[rowBacktrack - 1] == Y[originalColoumnIndexY - 1]) { 143. characterOf*LCS*[index*LCS* - 1] = X[rowBacktrack - 1]; 144. rowBacktrack--; 145. coloumnBacktrack--; 146. index*LCS*--; 147. originalColoumnIndexY--; 148. } 149. **else** **if** (numberOf*LCS*Character[rowBacktrack - 1][coloumnBacktrack] >  numberOf*LCS*Character[rowBacktrack][coloumnBacktrack - 1]) { 150. rowBacktrack--; 151. } 152. **else** { 153. coloumnBacktrack--; 154. originalColoumnIndexY--; 155. } 156. } 157. finishBacktrackProcess0 = 1; 159. **if** (finishBacktrackProcess0 == 1) { 160. **for** (**int** fromProcess = numberOfProcess - 1;  fromProcess >= 1; fromProcess--) { 161. MPI\_Recv(&initialIndex*LCS*, 1, MPI\_INT, fromProcess, 4,  MPI\_COMM\_WORLD, &status); 162. MPI\_Recv(&endIndex*LCS*, 1, MPI\_INT, fromProcess, 4,  MPI\_COMM\_WORLD,  &status); 163. **if** (initialIndex*LCS* == -1 && endIndex*LCS* == -1) { 164. } 165. **else** { 166. **for** (**int** received*LCS* = endIndex*LCS*;  received*LCS* <= initialIndex*LCS*; received*LCS*++) { 167. MPI\_Recv(&characterOf*LCS*[received*LCS*], 1, MPI\_INT,  fromProcess, 4, MPI\_COMM\_WORLD, &status); 168. } 169. } 171. finishReceived*LCS* = 1; 172. } 173. } 175. finishBacktrackAndReceived = 1; 176. } 177. } 178. } |

Kode Sumber . Pemilihan Indeks Perhitungan

|  |
| --- |
| 1. **int** rowStart = ((alpa \* (k - 1) \* lengthOfX) / numberOfProcess); 2. **int** rowEndPrevProcess = ((alpa \* (k - 2) \* lengthOfX) / numberOfProcess); 3. **int** rowEnd = (alpa \* k \* lengthOfX) / numberOfProcess; 5. **if** (rowStart != 0) 6. rowStart = rowStart + 1; 8. **if** (k != 1 && rowEnd == 0 && rowStart == 0) { 9. rowStart = rowStart + 1; 10. } 11. **if** (k != 1 && rowEndPrevProcess == 0 && rowStart == 0) { 12. rowStart = rowStart + 1; 13. } 15. **int** coloumnStart = (process)\*lengthOfY / numberOfProcess; 16. **if** (coloumnStart != 0) 17. coloumnStart = coloumnStart + 1; 19. **int** coloumnEnd = (process + 1) \* lengthOfY / numberOfProcess; |

Kode Sumber . Pemilihan Indeks Kolom Untuk Menerima Data

|  |
| --- |
| 1. **int** colStartForProcess1; 2. colStartForProcess1 = (1 \* lengthOfY / numberOfProcess); 3. **if** ((process \* lengthOfY / numberOfProcess) != colStartForProcess1) { 4. coloumnReceive = firstColoumn - coloumnPreviousProcess; 5. } 6. **else** { 7. coloumnReceive = coloumnStart - 1; 8. } |

Kode Sumber . Pengambilkan Karakter LCS

|  |
| --- |
| 1. **if** (X[rowBacktrack - 1] == Y[originalColoumnIndexY - 1]) { 2. character*LCS*[indexOf*LCS* - 1] = X[rowBacktrack - 1]; 3. rowBacktrack--; 4. coloumnBacktrack--; 5. indexOf*LCS*--; 6. originalColoumnIndexY--; 7. } 8. **else** **if** (numberOf*LCS*Character[rowBacktrack - 1][coloumnBacktrack] >  numberOf*LCS*Character[rowBacktrack][coloumnBacktrack - 1]) { 9. rowBacktrack--; 10. } 11. **else** { 12. coloumnBacktrack--; 13. originalColoumnIndexY--; 14. } |

Kode Sumber . Perhitungan Nilai Indeks LCS

|  |
| --- |
| 1. **if** (rowCalculate == 0 || coloumnCalculate == 0) { 2. numberOf*LCS*Character[rowCalculate][coloumnCalculate - firstColoumn] = 0; 3. } 4. **else** **if** (X[rowCalculate - 1] == Y[coloumnCalculate - 1]) { 5. numberOf*LCS*Character[rowCalculate][coloumnCalculate - firstColoumn] =  numberOf*LCS*Character[rowCalculate - 1][coloumnCalculate - 1 - firstColoumn] + 1; 6. } 7. **else** { 8. numberOf*LCS*Character[rowCalculate][coloumnCalculate - firstColoumn] =  max(numberOf*LCS*Character[rowCalculate - 1][coloumnCalculate - firstColoumn],  numberOf*LCS*Character[rowCalculate][coloumnCalculate - 1 - firstColoumn]); 9. } |

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

BIODATA PENULIS

I Gusti Ngurah Adi Wicaksana, lahir pada 29 September 1995 di Denpasar. Penulis merupakan seorang mahasiswa yang tengah menempuh studi di Jurusan Teknik Informatika Institut Teknologi Sepuluh Nopember. Memiliki beberapa hobi antara lain mendengarkan dan bermain musik serta bermain futsal. Pernah menjadi asisten dosen di mata kulian Sistem Digital dan Kecerdasan Buatan, serta asisten praktikum pada mata kuliah Sistem Operasi dan Jaringan Komputer. Selama menempuh pendidikan di kampus, penulis juga aktif dalam organisasi kemahasiswaan, antara lain Staff Departemen Media Informasi Himpunan Mahasiswa Teknik Computer-Informatika pada tahun ke-2 dan Staff Ahli Departemen Media Informasi Himpunan Mahasiswa Teknik Computer-Informatika pada tahun ke-3.

Kritik dan saran sangat diharapkan guna peningkatan kualitas dan penulisan selanjutnya. Untuk itu, silahkan kirim kritik dan saran ke : [adiwicaksana29@gmail.com](mailto:adiwicaksana29@gmail.com).