

**DISTRIBUTED SERVICE ORIENTED ARCHITECTURE
BASED ON K-NEAREST NEIGHBORHOOD ALGORITHM
(Studi Kasus : Pengolahan data Survey di BPS)**

PROPOSAL TESIS

**Disusun sebagai syarat kelulusan matakuliah
EL5090 Metodologi Penelitian / Tesis 1**

Oleh :

ARIS PRAWISUDATAMA

NIM : 23215131

(Program Studi Magister Teknik Elektro)



INSTITUT TEKNOLOGI BANDUNG

2016

LEMBAR PENGESAHAN

**DISTRIBUTED SERVICE ORIENTED ARCHITECTURE
BASED ON K-NEAREST NEIGHBORHOOD ALGORITHM
(Studi Kasus : Pengolahan data Survey di BPS)**

**Oleh :
ARIS PRAWISUDATAMA
NIM : 23215131**

**Program Studi Magister Teknik Elektro
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung**

Bandung, Maret 2016
Menyetujui,
Pembimbing

Dr. I Bagus Baskara Nugraha

DAFTAR ISI

LEMBAR PENGESAHAN.....	ii
DAFTAR ISI.....	iii
DAFTAR GAMBAR	iv
DAFTAR TABEL.....	v
1. Ringkasan Proposal.....	1
2. Latar Belakang	dst
3. Rumusan Masalah.....	
4. Tujuan	
5. Batasan Masalah	
6. Studi Literatur	
7. Metodologi	
8. Implikasi	
9. Sistematika Penulisan	
10. Penjadwalan	
DAFTAR REFERENSI	

DISTRIBUTED SERVICE ORIENTED ARCHITECTURE

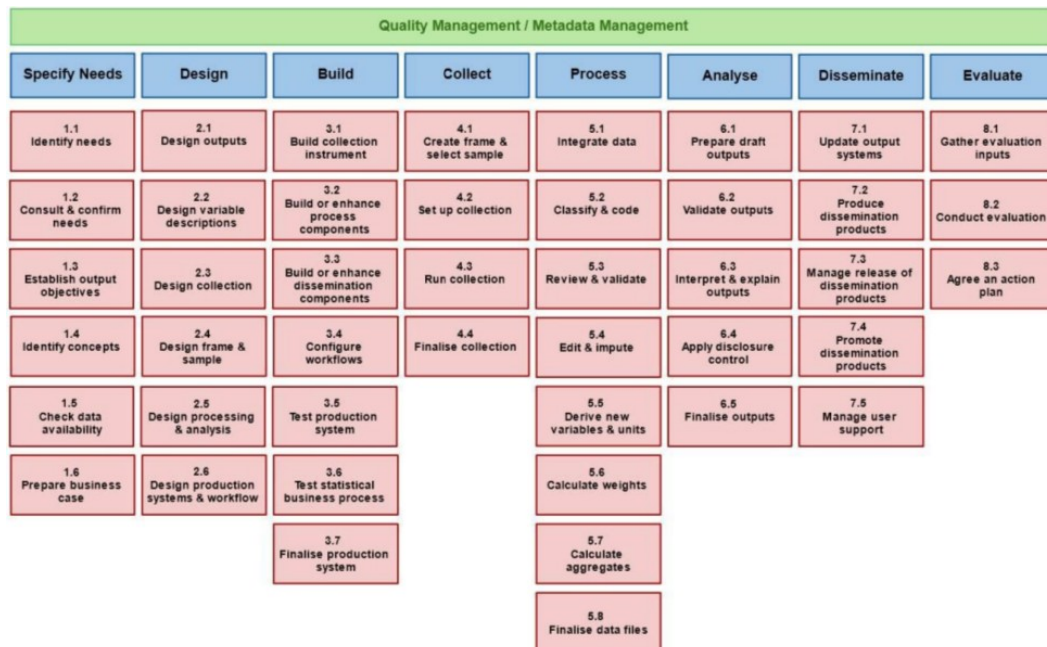
BASED ON K-NEAREST NEIGHBORHOOD

(Studi Kasus : Pengolahan data Survey di BPS)

1. PENDAHULUAN

Badan Pusat Statistik (BPS) merupakan suatu lembaga pemerintah non-departemen yang bertanggung jawab dalam penyediaan statistik dasar. Dalam peranannya sebagai penyedia data, BPS melakukan pengumpulan data dengan 2 (dua) metode : primer dan sekunder. Pengumpulan data primer berarti BPS secara mandiri mengumpulkan data dengan menggunakan metode wawancara langsung dengan responden, baik responden individu, rumah tangga, maupun perusahaan. Sementara pengumpulan data sekunder berarti BPS memperoleh data dari pihak lain.

Dalam melakukan kegiatan perstatistikan, yang selanjutnya merujuk kepada pengumpulan data primer, BPS merujuk kepada *General Statistical Business Process Model* (GSBPM)¹, yang merupakan suatu standard arsitektur bisnis kegiatan perstatistikan yang dirumuskan oleh *United Nations Economic Commission for Europe* (UNECE). Dalam GSBPM, *Business Process* Statistik dibagi menjadi 7 (tujuh) *phase* : *Specify Needs*, *Design*, *Build*, *Collect*, *Process*, *Analyse*, *Disseminate*, *Evaluate*, dimana masing-masing *phase* dipecah menjadi beberapa sub-proses.



Gambar 1. *Statistical Business Process Phases* dalam GSBPM

1 UNECE, 2013b. Generic Statistical Business Process Model, Version 5.0.

2. LATAR BELAKANG MASALAH

Pengumpulan dan pengolahan data dalam GSBPM tercakup dalam 3 (tiga) fase, yaitu : *Collect Phase*, *Process Phase*, dan *Analyze Phase*. Ketiga fase tersebut memerlukan peranan Teknologi Informasi dan Komunikasi (TIK) yang sangat besar. Berdasarkan keterangan dari narasumber di bidang Sistem Informasi Statistik², kedepannya BPS akan mengakomodasi pengumpulan data dengan metode *Computer Assisted Personal Interviewing* (CAPI), walaupun dari sisi *feasibility*-nya belum pernah diujicobakan³. Dengan menggunakan metode CAPI, berarti pengumpulan data akan langsung dipandu dengan menggunakan *device* yang akan terkoneksi langsung dengan server pengolahan.

Berdasarkan keterangan dari narasumber yang sama, server pengolahan data akan ditempatkan secara terpusat di BPS-RI. Dengan hanya menggunakan *single site data center*, maka sistem pengolahan hanya akan mempunyai *single point of failure*, dan jika mengalami masalah maka tidak ada *backup system* yang akan menggantikan peran sistem utama. Pendekatan yang dapat diadopsi adalah dengan menerapkan arsitektur sistem terdistribusi. Dengan arsitektur sistem terdistribusi, selain mengatasi *single point of failure*, juga dapat lebih mendekati pengguna yang juga terdistribusi.

Service Oriented Architecture (SOA) merupakan sebuah pendekatan yang sekarang banyak diadopsi dalam perancangan software. SOA memungkinkan penggunaan kembali komponen software untuk mengurangi effort dalam pengembangan. Teknologi *web service* juga memungkinkan komponen *software* yang terdistribusi dapat saling berinteraksi (*interoperable*) satu dengan yang lain.

Saat ini, *web service* hanya berfokus pada aspek fungsional komponen *software* yang terdistribusi, dimana data hanya diperlakukan sebatas sebagai *input*/parameter dan *output*/nilai balikan (*return value*)⁴. Sementara itu, pada sistem yang bersifat *data-intensive*, penerapan aplikasi berbasis SOA dengan menerapkan *web service* secara konvensional, menghasilkan proses yang tidak efisien. Penerapan pola terdistribusi, berdasarkan penelitian Takdir dkk, baik pada data (*database*) maupun proses (*web service*) terbukti dapat meningkatkan kinerja sistem yang bersifat *data-intensive*⁵. Tetapi, Takdir dkk hanya mengujicobakan penelitiannya dengan menggunakan

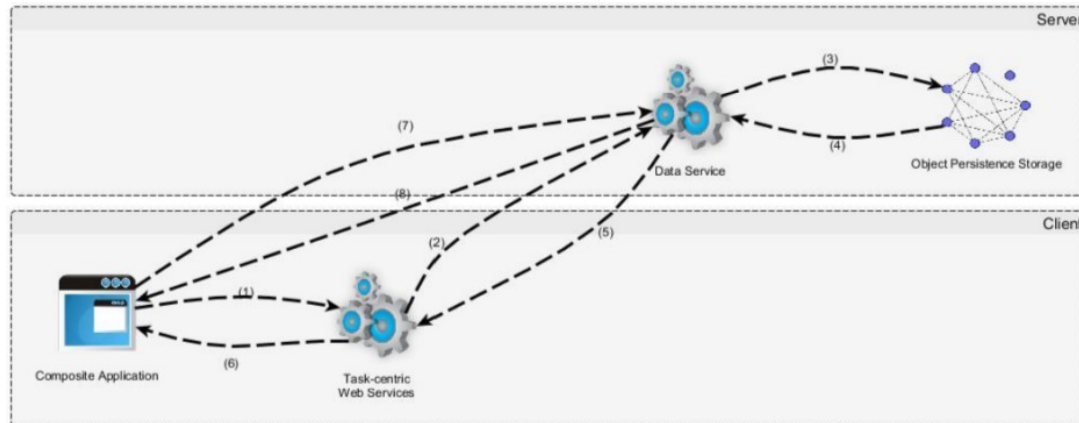
2 Keterangan Dr. Said Mirza Pahlevi, M.Eng., Kepala Sub Direktorat Pengembangan Basis Data, tanggal 24 Februari 2016 di ITB

3 Keterangan Dr. Muhammad Romzi, tanggal 4 Maret 2016 di ITB

4 R. Vaculín, T. Heath, and R. Hull, "Data-Centric Web Services Based on Business Artifacts," in *2012 IEEE 19th International Conference on Web Services (ICWS)*, 2012, 42–49, doi:10.1109/ICWS.2012.101.

5 Takdir and A. I. Kistijantoro, "Multi-Layer SOA Implementation Pattern with Service and Data Proxies for Distributed Data-Intensive Application System," in *2014 International Conference on ICT For Smart Society (ICISS)*, 2014, 37–41, doi:10.1109/ICTSS.2014.7013149.

sebuah *client* yang berupa *virtual machine*. Dalam skenario nyata, sistem (*web service*) ini akan diakses oleh ribuan atau lebih *device* yang terhubung dengannya. Kegagalan dalam pemrosesan dapat disebabkan oleh berbagai hal : *network error*, *network latency*, *external connection latency*, *input data corruption*, *hardware failure*, atau berbagai *runtime exceptions*⁶. Dengan mengakomodir sistem berbasis SOA terdistribusi, maka kegagalan pemrosesan pada satu sistem dapat diakomodir dengan me-*redirect* ke sistem terdekat.



Gambar 3. Skema Usulan, Takdir

3. RUMUSAN MASALAH

Berdasarkan uraian latar belakang permasalahan diatas, maka dapat dirumuskan suatu permasalahan penelitian yaitu bagaimana mendesain sistem berbasis SOA secara terdistribusi yang dapat mengakomodir *processing failure aware*.

4. TUJUAN

Tujuan utama dari penelitian ini adalah untuk mengembangkan sebuah desain implementasi sistem berbasis SOA secara terdistribusi untuk meminimalisir efek dari *processing failure*. Adapun tujuan khusus penelitian ini adalah :

- a) Mengembangkan desain implementasi sistem berbasis SOA terdistribusi,
- b) Melakukan ujicoba atas desain yang dikembangkan dengan studi kasus salah satu pengolahan survei di BPS, dan membandingkan dengan Sistem data-intensive terdistribusi yang dirancang Takdir serta metode SOA konvensional,

6 Tyler Akidau et al., "MillWheel: Fault-Tolerant Stream Processing at Internet Scale," *Proc. VLDB Endow.* 6, no. 11 (August 2013): 1033–44, doi:10.14778/2536222.2536229.

- c) Menganalisis dan mengevaluasi hasil ujicoba penerapan desain implementasi sistem berbasis SOA terdistribusi.

5. STUDI LITERATUR

Pembimbing,

Dr. I Bagus Baskara Nugraha