



**ANALYSIS AND CONFIGURATION ALLUXIO SERVER FOR
MAPREDUCE PROCESS**

FIRDHA YULIANTI
UNIVERSITAS
41812010015
MERCU BUANA

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS ILMU KOMPUTER

UNIVERSITAS MERCU BUANA

JAKARTA

2016



**ANALYSIS AND CONFIGURATION ALLUXIO SERVER FOR
MAPREDUCE PROCESS**

Laporan Tugas Akhir

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh :
UNIVERSITAS
MERCU BUANA
FIRDHA YULIANTI

41812010015

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA**

2016

LEMBAR PERNYATAAN

Yang bertanda tangan dibawah ini:

NIM : 41812010015
Nama : Firdha Yulianti
Judul Skripsi : ANALYSIS AND CONFIGURATION ALLUXIO SERVER FOR
MAPREDUCE PROCESS

Menyatakan bahwa Laporan Tugas Akhir saya adalah hasil karya sendiri dan bukan plagiat.
Apabila ternyata ditemukan didalam laporan Tugas Akhir saya terdapat unsur plagiat, maka saya
siap untuk mendapatkan sanksi akademik yang terkait dengan hal tersebut.

Jakarta, 21 July 2016



(Firdha Yulianti)

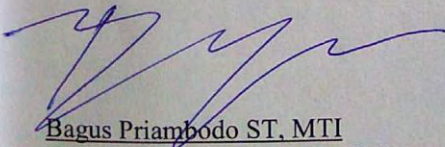
UNIVERSITAS
MERCU BUANA

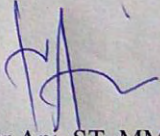
LEMBAR PENGESAHAN

NIM : 41812010015
Nama : Firdha Yulianti
Judul Skripsi : ANALYSIS AND CONFIGURATION ALLUXIO SERVER FOR
MAPREDUCE PROCESS

SKRIPSI INI TELAH DIPERIKSA DAN DISIDANGKAN
JAKARTA, 21 Juli 2016


UNIVERSITAS
MERCU BUANA
Nur Ani, ST, MMSI
Pembimbing


Bagus Priambodo ST, MTI
Koord. Tugas Akhir Sistem Informasi


Nur Ani, ST, MMSI
KaProdi Sistem Informasi

ACKNOWLEDGMENT

Alhamdulillah, blessing from Allah almighty for the mercy and guidance, the author can complete the thesis entitled Design and Performance Analysis of Big Data Processing System Based on Distributed Memory.

This thesis is a milestone represented by my own work and also by many other specialists and researchers of Computer Science department of BIT conducted during more than one decade. I have been studying at BIT since August 14th, 2014. Since my first day it has been an amazing experience. I started my Thesis during spring 2016. During my study I have learned that in Computer science there are many different fields and the application analysis, which is one of them, became my passion. I am going to present some facts which I have learnt from this field.

I would like to thank to all researches which published works I used for my research and also to all teachers of BIT who helped me with all my questions. My thesis represents the knowledge acquired during my studies at BIT and also during my previous study in Indonesia.

I would like to deeply thank to some of remarkable individuals who I have encountered at BIT. Let me acknowledge them. First and foremost I wish to thank to my advisor, professor Yanlong Zhai, supervisor of my thesis. He has been supporting me since the day I began working on my thesis as an undergraduate. He has been encouraging me during all the period of writing.

I wish to acknowledge thanks to MERCUBUANA UNIVERSITY which financially supports me through a scholarship since the first year I arrived at BIT. Finally I want to thank to my beloved family who has been supporting me during all those years in China. Surely I will leave Beijing with assurance that the basic source of my life energy resides my family. Thank all of my family members that you never stop pray for my studies. I know my family is unique in many ways but in the same time it is just perfect family as it should be. Their support has been unconditional all these years; they have given up many things for me to be at BIT; they have cherished with me every great moment and supported me whenever I needed it.

Jakarta, 21 July 2016

Firdha Yulianti

LIST OF CONTENTS

LEMBAR PERNYATAAN	i
LEMBAR PERSETUJUAN SIDANG	ii
LEMBAR PENGESAHAN	iii
ACKNOWLEDGMENT	iv
ABSTRACT	v
LIST OF CONTENTS	vi
LIST OF FIGURE	viii
LIST OF TABEL	ix
I. INTRODUCTION	1
1.1 Background	1
1.2 Formulation of The Problem	1
1.3 Limitation Problem	1
1.4 Research Objectives	2
1.5 Benefits of Research	2
1.6 Research Methodology	2
1.7 Systematics of Writing	3
II. BASIC THEORY	4
2.1 Big Data	4
2.2 Distributed File System	6
2.3 Alluxio	7
2.3.1 Design Overview	8
2.3.2 Advantages of Alluxio with Case Issue	21
2.4 MapReduce	23
III. ANALYSIS AND DESIGN	28
3.1 Design Single Node	28
3.2 Requirement Analysis	28
3.2.1 Specification Hardware	28
3.2.2 Specification Software	28
3.2.3 Specification File	28
3.2.4 Installation Java	28
3.2.5 Installation Alluxio	28
IV. IMPLEMENTATION AND TESTING	33
4.1 Implementation	33
4.1.1 Running MapReduce on Alluxio	35
4.2 Results	36
4.2.1 Computation Time	36
4.2.2 Memory	37
4.2.3 CPU	38
4.2.4 Stream Read File Input	39
4.2.5 Stream Write File Output	40
4.3 Analysis Result	40
V. CONCLUSION AND SUGGESTION	41
5.1 Conclusion	41
5.2 Suggestion	41
REFERENCES	42

LIST OF FIGURE

Figure 1.1. Workflow	2
Figure 2.1. Illustration by IBM	5
Figure 2.2. Big Data Ecosystem.....	8
Figure 2.3. Alluxio Architecture	8
Figure 2.4. A lineage graph example of multiple frameworks	9
Figure 2.5. Home Page.....	14
Figure 2.6. Configuration Page	15
Figure 2.7. Browse File System Page.....	15
Figure 2.8. Browse In-Memory Files Page	16
Figure 2.9. Workers Page.....	17
Figure 2.10. BlockInfo Page	18
Figure 2.11. Transparent Naming.....	19
Figure 2.12. Unified Namespace.....	20
Figure 2.13. Case 1.....	21
Figure 2.14. Case 2.....	22
Figure 2.15. Case 3.....	23
Figure 2.16. A client submitting a job to MapReduce.....	25
Figure 2.17. JobTracker and TaskTracker Interaction [13].....	27
Figure 3.1. Web Interface Alluxio.....	30
Figure 4.1. Input Directory Alluxio.....	35
Figure 4.1. Input Directory Alluxio.....	35
Figure 4.2. Output Directory Alluxio.....	36
Figure 4.2. Output Directory Alluxio.....	36
Figure 4.3. Computation Time	37
Figure 4.4. Memory.....	38
Figure 4.5. CPU.....	38
Figure 4.5. CPU.....	38
Figure 4.6. Stream Read File Input	39
Figure 4.7. Stream Write File Output.....	40

LIST OF TABEL

Table 4.1 The Result of Computation Time Mapreduce	36
Table 4.2 The Result for Size of Memory	37
Table 4.3 Result for the Speed of Processor to Process the Data	38
Table 4.4 This Result for the Stream Read File Input	39
Table 4.5 This Result for the Stream Write File Output	40

