Cairo University

NOVEL FAT TREE TOPOLOGIES AND ROUTING SCHEMES FOR DATA CENTER NETWORKS

By

MennaTallah Adel Ahmed Mazen Hussein

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
in
ELECTRONICS AND COMMUNICATIONS ENGINEERING

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2017

NOVEL FAT TREE TOPOLOGIES AND ROUTING SCHEMES FOR DATA CENTER NETWORKS

By

MennaTullah Adel Ahmed Mazen Hussein

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
in
ELECTRONICS AND COMMUNICATIONS ENGINEERING

Under the Supervision of

Prof. Dr.
Mohamed Hazim Tawfik

…….

Professor of Wireless Communications
Communications and Electronics
Department, Faculty of Engineering,
Cairo University

Prof. Dr.
Tamer Elbatt

…….

Associate Professor
Communications and Electronics
Department, Faculty of Engineering,
Cairo University

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2017

NOVEL FAT TREE TOPOLOGIES AND ROUTING SCHEMES FOR DATA CENTER NETWORKS

By

MennaTallah Adel Ahmed Mazen Hussein

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
in
ELECTRONICS AND COMMUNICATIONS ENGINEERING

Approved by the Examining Committee:

Prof. Dr. **Mohamed Hazim Tawfik**, Thesis Main Advisor

……....

Prof. Dr. **Tamer Elbatt**, Thesis Advisor

……....

Prof. Dr. **Magdi El-Soudani**, Internal Examiner, Professor of Communications

……....

Dr. **Fadel Digham**, External Examiner, Executive Director of R&D, National Telecom Regulatory Authority (NTRA)

…….

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2017

Engineer's Name: MennaTallah Adel Ahmed Mazen Hussein
Date of Birth: 15/08/1989
Nationality: Egyptian
E-mail: Menna_mzn@yahoo.com
Phone: +201002355199
Address: 45 Khedr elTony, Nasr City
Registration Date: 1/10/2012
Awarding Date: 14/8/2017
Degree: Master of Science
Department: Electronics and Communications Engineering

Supervisors:

Prof. Mohamed Hazim Tawfik
Prof. Tamer Elbatt

Examiners:

Prof. Fadel Digham (External examiner)
Prof. Magdi El-Soudani (Internal examiner)
Prof. Mohamed Hazim Tawfik (Thesis Main Advisor)
Prof. Tamer Elbatt (Thesis Advisor)

Title of Thesis:

Novel Fat Tree Topologies and Routing Schemes for Data Center Networks

Key Words:

Data Centers, Fat Trees, Extra Links, Load balancing, Aggregate bandwidth.

Summary:

This thesis explores optimized datacenter topologies for large scale networks. In particular, we study the Fat Tree topology [1] and propose novel topologies and routing schemes to enhance the bandwidth aggregated at the host end.

Network models of each topology and associated routing algorithms were developed for simulation. Performance evaluation is presented using different traffic patterns including extreme worst cases. The Use of 60 GHz wireless technology in the novel topologies is then studied. Results show that improved performance and interesting insights.

Acknowledgments

I would certainly thank my supervisors for putting up with all my ups and downs during the research period and for always pushing me forward and for their understanding. I really enjoyed working with you. Thank you.

I really appreciate dr. Nora Ahmed's efforts for trying to help as much as she could.

I would never forget to thank my Family, especially my parents, my sisters and brother. They are my backbone, they believed in me when everything was no working out. I love you.

Thanks to all my friends and colleagues, especially Mayy Magdy, for being there for me all the time.

Dedication

I will dedicate this work to Adel Mazen and Eman Shokry; my ever loving and caring Father and Mother .May Allah protect you always.

Table of Contents

ACKNOWLEDGMENTS ..... **I**

DEDICATION ..... **II**

TABLE OF CONTENTS ..... **III**

LIST OF TABLES ..... **V**

LIST OF FIGURES ..... **VI**

ABSTRACT ..... **VII**

CHAPTER 1 : INTRODUCTION **1**

1.1. THESIS ORGANIZATION 2

CHAPTER 2 : BACKGROUND **3**

2.1. INTRODUCTION TO DATACENTERS 3

2.1.1. Datacenter Topologies and Technologies 3

2.1.2. Datacenter Traffic Study 5

2.2. INTRODUCTION TO FAT TREES 6

2.2.1. Fat Trees History and Motivation: 6

2.2.2. Hierarchical tree 7

2.2.3. Fat Trees Main Features: 9

2.2.4. Addressing and Two-Level Table Routing 10

2.2.5. Two- Level table Routing Example 11

2.3. RELATED WORK 12

CHAPTER 3 : FAT TREE WITH ADDITIONAL LINKS **13**

3.1. ADDITIONAL INTER-POD LINKS TO THE FAT TREE 13

3.2. PROPOSED ROUTING ALGORITHMS 15

3.2.1. Preventive Routing (PR) 15

3.2.1.1.Traffic flow categories 15

1-One Hop Away Flows 16

2-Non Adjacent Pod Flows 16

3-Adjacent Pod Flows 16

3.2.1.2.Routing Algorithm at the Edge Switches 18

3.2.1.3.Routing Algorithm at the Aggregate Switches 23

3.2.1.4.Routing Algorithm at the Core Switches 23

3.2.2. Load Balanced Shortest Path Routing (LB-SPR) 24

3.2.2.1.Routing Algorithm at the Edge Switches 24

3.2.2.2.Routing Algorithm at the Aggregate Switches 28

3.2.2.3.Routing Algorithm at the Core Switches 29

3.2.3.Preventive Routing Algorithm Scalability 30

3.2.3.1.Routing Algorithm at the Edge Switches 32

3.2.3.2.Routing Algorithm at the aggregate and core Switches 35
3.2.3.3.Discussion 36

CHAPTER 4 : FIT FAT TOPOLOGY 37

4.1. REMOVING TWO CORE SWITCHES VS. EXTRA LINKS 38
4.2. REMOVED CORES PREVENTIVE ROUTING (RCPR) 39
4.2.1. Routing Algorithm at the Edge Switches 39
4.2.2. Routing Algorithm at the Aggregate Switches 41
4.2.3. Routing Algorithm at the Core Switches 41

CHAPTER 5 : 60 GHZ WIRELESS L INKS IN DATACENTERS 42

5.1. LITERATURE REVIEW OF STANDARD AND APPLICATIONS 42
5.2. FAT TREES AND 60 GHZ 44
5.2.1. Modeling of Hybrid Fat trees with Wireless Interpod Links 44
5.2.2. Simulation Scenarios 45
5.2.2.1. One Giga Ethernet Fat Tree with 0.3 Mbps Wireless Interpod Links 46
5.2.2.2. Ten Giga Ethernet Fat Tree with 5.28 Gbps Wirless Interpod Links 46
5.2.3. Delay based performance evaluation 47
5.2.3.1. Capacity Assignment problem 47
5.2.3.2. CVX solution 48

CHAPTER 6 : SIMULATION SETUP AND PERFORMANCE EVALUATION 50

6.1. SIMULATION SETUP 50
6.2. COMMUNICATION PATTERNS 52
6.2.1. Random 52
6.2.2. Worst case 1 52
6.2.3. Worst case 2 52
6.3. NUMERICAL RESULTS 54
6.3.1. Added Links at One Level 54
6.3.2. Added Links at Both Levels 56
6.3.3. FitFat Tree with Removed Core Switches 59
6.3.4. Wireless Extension Evaluation and results 60
6.3.4.1. Discussion 62
6.3.4.2. Recommendation 63

CHAPTER 7 : CONCLUSION AND FUTURE WORK 64

7.1. CONCLUSION 64
7.2. FUTURE WORK 65

REFERENCES 66

APPENDIX A: MATLAB CODE FOR CONVEX PROBLEM SOLVING 69