FAIR TRAFFIC SCHEDULING AND SHAPING ALGORITHMS FOR DIFFSERV NETWORKS

MOHAMED AWAD ELSHAIKH MOHAMED

DOCTOR OF PHILOSOPHY
UNIVERSITY PUTRA MALAYSIA

2006

FAIR TRAFFIC SCHEDULING AND SHAPING ALGORITHMS FOR DIFFSERV NETWORKS

By

MOHAMED AWAD ELSHAIKH MOHAMED

Thesis Submitted to the School of Graduate Studies, University Putra Malaysia, in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

July 2006

This work is dedicated to my beloved Father, Mother, Brothers, Sisters and my family Raja, Waad and Sajid

Abstract of thesis presented to the Senate of Univesiti Putra Malaysia

in fulfillment of the requierments for the degree of Doctor of Philosophy

FAIR TRAFFIC SCHEDULING AND SHAPING ALGORITHMS FOR DIFFSERV NETWORKS

By

MOHAMED AWAD ELSHAIKH

December 2006

Chairman : Aossociate Professor Mohamed Othman, PhD

Faculty : Computer Science and Information Technology

Traditionally, the Internet has provided only best effort service to every user without any consideration to any requirements. As the Internet grows and becomes universally available, it becomes very important to deal with real time service delivery to application such as IP telephony. Thus, an interest has developed in having the Internet to provide some degree of Quality of Service (QoS). To provide different QoS commitments, the IETF developed many technologies that requires resources such as bandwidth and buffers to be explicitly reserved for a given data flow to ensure that the application receives its requested QoS.

DiffServ is proposed by the thought that per-flow moved to the edge of the network with very simple functionalities left at the core network. However, services provided

by DiffServ networks have lower flexibility and fairness among the aggregates sharing the network and among the flows of those aggregates. Per Hob Behavoiurs (PHBs) are implemented at DiffServ networks nodes using some scheduling and queuing mechanisms that are predecessor by markers and shaper. The current proposed DiffServ have unfairness problem that is caused by these elements. This thesis investigates the effect of different scheduling mechanisms and usage of a different marker algorithms then develop and implement a fair scheduling mechanism and integrate it with a shaper with the usage of a very fair efficient marker at the edge of the DiffServ networks with the usage of simulation and analytical techniques.

In this thesis, I^2tswTCM is proposed as an enhanced version of the famous tswTCM marker scheme used in the current DiffServ architecture, this new algorithm improves fairness in the excess bandwidth among different aggregates in a DiffServ networks and it has a better fairness than all other algorithms for different network provision levels. In addition, the algorithm marks the right amount of yellow traffic into the network proportional to the CIR. Nevertheless, it is not sensitive to the number of flows enters the networks. The thesis also proposed a new scheduling algorithm (FWFQ) which suite the DiffServ architecture and it gives a better performance in terms of delay and jitter. The research shows the importance of considering the scheduler type when dealing with different types of traffic sources in DiffServ networks. The interaction of different traffic type flows in a DiffServ networks using the proposed integration model between FWFQ and the dual token

bucket shaper algorithm is analyzed. The result shows that the proposed model can significantly improve the performance.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagi memenuhi keperluan untuk ijazah Doktor Falsafah

ALGORITMA PENSKEDULAN DAN PEMBENTUKAN TRAFIK ADIL UNTUK RANGKAIAN-RANGKAIAN DIFFSERV

Oleh

MOHAMED AWAD ELSHAIKH

Disember 2006

Pengerusi : Professor Madya Mohamed Othman, PhD

Fakult : Komputer Sains dan Teknologi Maklumat

Secara tradisi, Internet menawarkan perkhidmatan terbaik kepada setiap pengguna tanpa mengambilkira sebarang pertimbangan terhadap ke perluan. Apabila Internet berkembang dan digunakan secara meluas, ianya menjadi penting untuk menangani penghantaran perkhidmatan secara nyata ke aplikasi seperti telefoni *IP*. Lanjutan itu, tercetusnya minat untuk *Internet* menawarkan sesuatu peringkat perkhidmatan berkualiti (QoS). Untuk menawarkan kesungguhan *QoS* yang berbeza, IETF telah membangunkan banyak teknologi yang berasaskan sumber seperti jalur lebar dan storan sementara dikhaskan untuk aliran data yang diberi bagi memastikan aplikasi tersebut menerima *QoS* yang diperlukan.

***DiffServ* dicadangkan kerana idea operasi pra aliran harus digerakkan ke hujung rangkaian-rangkaian dan fungsi yang mudah dibiarkan dalam rangkaian utama.**

Walaubagaimana pun, perkhidmatan yang ditawarkan oleh rangkaian- rangkaian *DiffServ* mempunyai keanjalan yang rendah dan adil di antara perkongsian agregat dalam rangkaian dan di antara aliran agregat tersebut. *Per Hob Behavoiurs* (PHBs) dilaksanakan pada nod rangkaian-rangkaian *DiffServ* menggunakan penjadualan dan mekanisma barisan yang menjadi *predecessor* oleh penanda dan pembentuk. Cadangan *DiffServ* yang terkini mempunyai masalah ketidakadilan yang disebabkan oleh unsur-unsur tersebut. Tesis ini mengkaji kesan mekanisma penjadualan yang berbeza dan penggunaan algoritma penanda yang berbeza untuk membangunkan dan melaksanakan mekasnisma yang berkesan dan mudah dan diintegrasikan dengan penajam yang menggunakan penanda yang adil dan saksama di penghujung rangkaian-rangkaian *DiffServ* dengan bantuan simulasi dan teknik penganalisaan.

Dalam kajian ini, I^2tswTCM dicadangkan sebagai versi tambahan untuk skema penanda tswTCM yang terkenal dengan menggunakan rekabentuk semasa. Algorithma baru ini memperbaiki keadilan lebar jalur yang berlebihan di antara agregat yang berlainan dalam rangkaian-rangkaian *DiffServ.* Ia mempunyai keadilan yang lebih baik daripada algorithma yang lain untuk lapisan rangkaian yang berbeza. Tambahan pula, algoritma ini memasukkan jumlah trafik kuning yang betul ke dalam rangakain memadani *CIR*. Namun begitu, ianya tidak sensitif ke atas jumlah aliran yang memasuki rangkaian. Tesis ini juga mencadangkan satu algorithma penjadualan baru (*FWFQ*) yang sesuai dengan rekabentuk *DiffServ* dan ia memberikan kesan yang lebih baik dalam penangguhan dan penangguhan tak

sekata (*jitter*). Kajian ini juga menunjukkan kepentingan untuk mempertimbangkan jenis penskedulan apabila menangani jenis sumber trafik yang berbeza dalam rangkaian-rangkaian *DiffServ*. Analisa dilakukan ke atas interaksi jenis aliran trafik yang berbeza dalam rangkaian-rangkaian *DiffServ* yang menggunakan model integrasi yang dicadangkan di antara *FWFQ* dan kaedah pembentuk *dual token bucket.* Keputusan menunjukkan bahawa model yang dicadangkan boleh meningkatkan prestasi secara ketara.

ACKNOWLEDGEMENTS

All Praise is due to Almightly Allah as He is all Mercifull, Most Gracious and Most Compassionate who gathered all knowledge in His Essence and who is the Creator of all knowledge for eternity. May God praise our Prophet Muhammad, and render him and his household safe and secure from all evil. All Thanks belong to Almightly Allah.

I would like to thank the government of Malaysia and the University Putra Malaysia for offering me the chance of gaining knowledge in this country. Then my gratitude goes to Sudan University of Science and Technology (SUST).

I wish to express a deep appreciation and gratitude to my advisor and supervisor Associate Proffessor Dr. Mohamed Othman for his extensive help and valuable comments and advice at every stage of my research.

Further I must record my thanks to my co-supervisors Dr. Shamala Subramaniam and Dr. Jalil Md. Desa for their valuable help in making this research a reality.

I would also like to thank my examination committee members for their time and suggestions.

I should not forget the helpful discussions with my colleagues especially network research labrotary members.

I express warm appreciation to my teachers throughout my whole academic life.

I reserve my deepest gratitude to my parents, brothers, sisters and other family members for their preserving efforts towards my entire life.

Last but not least I reserve my deepest gratitude to my wife Raja, and my children Waad and Sajid for their preserving efforts towards my social and academic life.

I certify that an Examination Committee has met on 19 December 2006 to conduct the final examination of Mohamed Awad Elshaikh on his Doctor of Philosophy thesis entitled "Fair Taffic Scheduling and Shaping Algorithms for DiffServ Networks" in accordance with Universiti Pertanian Malaysia (Higher Degree) Act 1980 and Universiti Pertanian Malaysia (Higher Degree) Regulations 1981. The Committee recommends that the candidate be awarded the relevant degree. Members of the Examination Committee are as follows:

Ali Mamat, PhD
Associate Professor
Faculty of Graduate Studies
Universiti Putra Malaysia
(Chairman)

Hamidah Ibrahim, PhD
Associate Professor
Faculty of Graduate Studies
Universiti Putra Malaysia
(Internal Examiner)

Ramlan Mohamad, PhD
Associate Professor
Faculty of Graduate Studies
Universiti Putra Malaysia
(Internal Examiner)

Kasmiran Jumari, PhD
Professor
Faculty of Graduate Studies
Universiti Putra Malaysia
(External Examiner)

_____
HASANAH MOHD GHAZALI, PhD
Professor/ Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

This thesis submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfillment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee are as follows:

Mohamed Othman, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Shamala Subramaniam, PhD
Lecturer
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

Jalil Md. Desa, PhD
Lecturer
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

_____
AINI IDERIS, PhD
Professor/Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

DECLARATION

I hereby declare that the thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at UPM or other institutions.

_____

MOHAMED AWAD ELSHAIKH

Date:

TABLE OF CONTENTS

		Page
DEDICATION		ii
ABSTRACT		iii
ABSTRAKT		v
ACKNOWLEGMENTS		viii
APPROVAL		x
DECLARATION		xii
LIST OF TABLES		xvi
LIST OF FIGURES		xvii
LIST OF ABBREVIATION		xx
CHAPTER		
1	INTRODUCTION	1
	1.1 Environment of Discourse	1
	1.2 Differentiated Service Architecture	
	1.2.1 Differentiated Services	5
	1.2.2 Traffic Classifier, Profile and Conditioner	8
	1.2.3 Differentiated Services Field	
	1.3 Differentiated Services and Integrated Services	14
	1.4 Problem Statement	15
	1.5 Objectives	16
	1.6 Scope	17
	1.7 Methodology	17
	1.8 Outline of the Thesis	19
2	LITERATURE REVIEW	21
	2.1 Network Traffic Engineering	
	2.1.1 Congestion Control	22
	2.1.2 Traffic Shaping	23
	2.2 Differentiated Services Marker algorithms	
	2.2.1 Leaky Bucket	24
	2.2.2 Token Bucket	25
	2.2.3 Single Rate Three Color Marker	25
	2.2.4 Two rate Three Color Marker	
	2.2.5 Rate Adaptive Shaper	
	2.2.6 Time Sliding Window Three Color Marker	
	2.2.7 Improved Time Sliding Window Three Color Marker	
	2.3 Per Hop Behavior in DiffServ	
	2.3.1 Principle of PHB	41
	2.3.2 PHB and Scheduling Algorithms	41
	2.4 Related Schedulers	42
	2.4.1 First Generation	43

2.4.2 Second Generation 45
2.4.3 Third Generation 49
2.5 Summary 53

3 FRAMEWORK AND PERFORMANCE MODEL 54
3.1 Performance Analysis Approaches 54
3.2 Performance Patterns
3.2.1 Confidence Interval 56 57
3.2.2 Summarization 59
3.3 Simulation 60
3.3.1 Simulation Types 60
3.3.2 Simulation Techniques 62
3.3.3 Network Simulator 64
3.4 DiffServ Architecture Modules 67
3.5 Performance Metrics and Discussions 68

4 FAIR TSWTCM ALGORITHM 71
4.1 Introduction 71
4.2 Type of Markers 72
4.2.1 Token Bucket Based 73
4.2.2 Average Rate Estimator Based 74
4.3 Unfairness of Existing Markers 75
4.4 The Proposed I^{2}tswTCM Algorithm
4.4.1 Algorithm Description 77
4.4.2 Assumptions and Analysis 78
4.5 Marking Probability 81
4.5.1 Importance of Yellow Packets 81
4.5.2 Marking Probability for Different Algorithms 82
4.6 Model Development and Performance Criteria 84
4.6.1 Simulation Model 84
4.6.2 Model Scenarios 86
4.6.3 Performance Metrics
4.7 Results and Discussions 87 88
4.7.1 Case 1 88
4.7.2 Case 2 90
4.7.3 Case 3 91
4.8 Summary 95

5 FAIR DIFFSERV SCHEDULER ALGORITHM 96
5.1 Introduction 96
5.2 Weighted Fair Queuing 97
5.2.1 WFQ Analysis
5.2.2 Unfairness of Weighted Fair Queuing 98 100

5.3 The Proposed Fair Weighted Fair Queuing
5.3.1 Flow Charts 102
5.3.2 Analytical Description and Analysis
5.4 Model Development and Performance Criteria 108
5.4.1 Simulation Model 108
5.4.2 Model Scenarios 109
5.4.3 Performance Metrics
5.5 Results and Discussions 111
5.6 Summary 115

6 INTEGRATED APPROACH FOR DUAL SHAPER SCHEDULER 117
6.1 Introduction 117
6.2 Integration Approach
6.3 Proposed Integrated Model 118
6.3.1 FWFQ 120
6.3.2 Dual Token Bucket Module
6.4 Model Development and Performance Criteria 122
6.4.1 Simulation Model 122
6.4.2 Model Scenarios 124
6.4.3 Performance Metrics
6.5 Results and Discussions 125
6.6 Summary 132

7 CONCLUSIONS AND FUTURE WORKS 133
7.1 Conclusions 133
7.2 Future Works
135

REFERENCES 136
APPENDICES 142
A Nortel DiffServ Architecture in NS-2 142
B Functions 148
BIODATA OF THE AUTHOR 152

LIST OF TABLES

Table		Page
1.1	Recommended AF Codepoints Values	7
4.1	Case 1 Fairness of the Different Algorithms	88
4.2	Case 2 Fairness of the Different Algorithms	90
4.3	Case 3 Fairness of the Different Algorithms	92
5.1	Average Delay, Loss and Jitter	114

LIST OF FIGURES

Figure		Page
1.1	DiffServ Classifier and Traffic Conditioner	8
1.2	Classifier	9
1.3	Profile Example	10
1.4	A Generic Meter	11
1.5	The Structure of DiffServ Field	14
2.1	Token Bucket Behavior	27
2.2	srTCM Color-Blind Mode	27
2.3	srTCM Color-Aware Mode	27
2.4	trTCM Color-Blind Mode	29
2.5	trTCM Color-Aware Mode	29
2.6	RAS with srTCM	32
2.7	RAS with trTCM	34
2.8	Block Diagram for the tswTCM	35
2.9	Rate Estimator Algorithm	36
2.10	tswTCM Algorithm	38
2.11	ItswTCM Algorithm	40
3.1	Scientific Method for Knowing a System	57
3.2	DiffServ Modules Architecture	67
4.1	I^2tswTCM Algorithm	78
4.2	tswTCM Marking Probability	83
4.3	ItswTCM Marking Probability	83

4.4	**I^2tswTCM Marking Probability**	**84**
4.5	**Network topology**	**85**
4.6	**Equal Number of Flows for all Aggregate**	**89**
4.7	**Fixed Rate Aggregate With a Lesser Number of Flows**	**90**
4.8	**Fixed Rate Aggregate With Greater Number of Flows**	**92**
4.9	**Effects of Gamma Constant Value**	**93**
4.10	**Equal Number of Flows in Aggregates**	**94**
4.11	**Different Number of Flows in Aggregates**	**94**
5.1	**FWFQ Algorithm Description**	**101**
5.2	**Packet Arrival Flow Chart**	**102**
5.3	**Packet Departure Flow Chart**	**103**
5.4	**DiffServ Network Model**	**109**
5.5	**Idrop**	**112**
5.6	**Edrop**	**113**
5.7	**Total Drop**	**113**
5.8	**WFQ Fairness Index for Different Classes**	**114**
5.9	**FWFQ Fairness Index for Different Classes**	**115**
6.1	**Integration Approach**	**119**
6.2	**The New Integration Architecture**	**121**
6.3	**The New Integration Algorithm**	**121**
6.4	**DiffServ Model used in the Simulation**	**123**
6.5	**Different Shaper Levels with Different Schedulers**	**126**
6.6	**UDP Average Delay and Jitter**	**127**

6.7 UDP Losses 128

6.8 TCP Average Delay and Jitter 129

6.9 TCP Losses 130

6.10 Fairness of Different Aggregate with Different Schedulers 131

LIST OF ABBRIVIATION

ACE	**Accelerated**
AF	**Assured Forwarding**
AUX	**Auxiliary**
BA Classifier	**Behavior Aggregate Classifier**
BE	**Best Effort**
CBS	**Committed Bursty Size**
CIR	**Committed Information Rate**
CIR_th	**Committed Information Rate_threshold**
DiffServ	**Differentiated Services**
DSCP	**Differentiated Services Code Point**
EBS	**Excess Bursty Size**
EF	**Expedited Forwarding**
FI	**Fairness Index**
FWFQ	**Fair Weighted Fair Queue**
GPS	**Generalize Processor Sharing**
HoL	**Head of Line**
HoL_PJ	**Head of Line with Priority Jumping**
I^2tswTCM	**Double Improve Time Sliding Window Three Color**
IETF	**Internet Engineering Task Force**
IID	**Independent and Identically Distribution**
ItswTCM	**Improve Time Sliding Window Three Color Marker**
MF Classifier	**Multi-Field Classifier**

MIR	**Maximum Information Rate**
MIR_th	**Maximum Information Rate_threshold**
MLT	**Maximum Laxity Threshold**
PBS	**Peak Bursty Size**
PEC	**Performance Evaluation Checklist**
PHB	**Per Hob Behaviour**
PIR	**Peak Information Rate**
PQ	**Priority Queuing**
PQWRR	**Priority Queue Weighted Round Robin**
QoS	**Quality of Services**
RAS	**Rate Adaptive Shaper**
RED	**Random Early Detection**
RSVP	**Resource Reservation Protocol**
RTT	**Round Trip Time**
SLA	**Service Level Agreement**
SLS	**Service Level Specification**
srRAS	**Single Rate - Rate Adaptive Shaper**
srTCM	**Single Rate Three Color Marker**
TD^{2}FQ	**Token Driven Delay-Sensitive Dynamic Fair Queuing**
ToS	**Type of Service**
trTCM	**Two Rate Three Color Marker**
trRAS	**Two Rate – Rate Adaptive Shaper**
tswTCM	**Time Sliding Window Three Color Marker**

VC	**Virtual Clock**
WFQ	**Weighted Fair Queue**
WIRR	**Weighted Interleave Round Robin**
WRR	**Weighted Round Robin**