الجامعة الإسلامية العالمية ماليزيا
INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA
اونيۏرسيتي اسلام انتارابڠسا مليسيا

TRAFFIC SHAPING AND CONGESTION CONTROL IN A DIFFSERV DOMAIN

BY

MOHAMED AWAD ELSHAIKH MOHAMED

A DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMNET FOR THE DEGREE OF
MASTER OF SCIENCE IN COMPUTER AND INFORMATION ENGINEERING

KULLIYYAH OF ENGINEERING
INTERNATIONAL ISLAMIC UNIVERSITY
MALAYSIA

OCTOBER 2001

ABSTRACT OF THE DISSERTATION

As the Internet grows and becomes universally available, it becomes very important to deal with real time service delivery to applications. Thus an interest has developed in having the Internet to provide some degree of Quality of Service (QoS). The Differentiated Services (DiffServ) architecture has been proposed as scalable solution for providing service differentiation among flows without any per-flow buffer management inside the core of the network. A diffserv domain at its edge may control the amount of traffic that enters or exits the domain at various levels of drop precedence, such traffic conditioning may include traffic shaping, discarding packets and reassigning of packets to another traffic class. This dissertation investigates the effect of using various shapers and markers on different stream of traffic types (VBR, CBR, and VBR with a background traffic) on the overall diffserv domain. Simulations are conducted using NS-2 simulation platform applying the traffic stream shaped by shaper and followed by marker. The shaped traffic then applied across a diffserv domain under normal traffic conditions to determine the effect on the level of the congestion due to a particular shaper and marker used. The most suitable one, for a particular application, generates a maximum number of conformed packets and less loss packets, thus reducing the potential of congestion in the network. Shaping the traffic stream, in general, can reduce the percentage of losses and at the same time increased the transmitted number of conformed packets and thus better performance during congestion time. Also we can see that for the VBR applications that run without any background traffic shaped by tswTCM combine with dual token bucket shaper has a better performance during congestion periods than using srTCM and trTCM, but for applications that are more sensitive to delay jitter trTCM is more suitable. For a VBR application that runs beside background traffic, trTCM scheme has a better performance during congestion periods in terms of delay jitter but tswTCM measures a better performance in terms of the number of conformed packets and the percentage of losses.

ملخص البحث

مع ازدياد حجم استخدام الإنترنت وشموليتها اصبح من الضروري التعامل مع توصيل الخدمة فوريا للتطبيقات. لذلك توجه الاتجاه في الإنترنت لاعطاء درجات لجودة الخدمة(QoS). هيكلية الخدمات المميزة (diffserv) طرح كحل قياسي لاعطاء خدمة مميزة لتدفق البيانات من غير ادارة ذاكرة لكل تدفق بداخل الشبكة الرئيسية. مجال الخدمة المميزة يمكن ان يتحكم في كمية البيانات الداخلة والخارجة من المجال بدرجات مختلفة من اسقاط الرزم (packets). منظمات الحركة هذه يمكن أن تشمل منظمات الحركة، اسقاط الرزم، اعادة تخصيص الرزم لنوعية حركة اخرى. هذا البحث يمحّص تاثير استخدام مجموعة مختلفة من المنظمات (shapers) والملونات (markers) على نوعيات مختلفة من الحركة(VBR, CBR, VBR with background traffic) على مجال الخدمة المميزة(diffserv). استخدم برنامج المحاكاة (NS-2) وطبق النظام باستخدام الحركة منظمة باستخدام المنظم (shaper) ثم اتبع بالملون(marker). الحركة المنظمة بعد ذلك أدخلت على مجال الخدمات المميزة (diffserv) في وضع الخدمة الطبيعي وذلك حتى تتم دراسة تاثير المنظمات (shapers) والملونات (markers) المستخدمة على درجة الاختناق. انسب واحد لتطبيق بعينه هو الذي ينتج أكبر عدد من الرزم المطابقة وأقل رزم مفقودة وبالتالي نقصان درجة الاختناق في الشبكة. تنظيم الحركة بصورة عامة ينقص نسبة فقدان الرزم وفي نفس الوقت ازدياد عدد الرزم الموافقة المرسلة وبالتالي افضل اداء في زمن الاختناق. ايضا يمكن أن نرى في تطبيقات (VBR) والتى تستخدم من غير حركة داخلية ومنظمة باستخدام (tswTCM) ومنظم (dual token bucket) افضل اداء من استخدام (srTCM) و (trTCM) في زمن الاختناق، ولكن للتطبيقات الاكثر حساسية بالنسبة (delay jitter) تكون (trTCM) هي الافضل. بالنسبة لتطبيقات (VBR) مع حركة داخلية لتطبيق اخر (trTCM) يكون لها الأفضلية في زمن الاختناق بالنسبة ل(delay jitter)، ولكن (tswTCM) لها الأفضلية بالنسبة لعدد الرزم الموافقة ونسبة الفقدان.

APPROVAL PAGE

I certify that I have supervised and read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Master of Science in Computer and Information Engineering.

13-10-2001
Farhat Anwar
Supervisor

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Master of Science in Computer and Information Engineering.

Sufyan Al-Irhayim
Internal Examiner

16/10/01
Kasmiran B. Jumari
External Examiner

This dissertation was submitted to the department of Electrical and Computer Engineering and is accepted as partial fulfillment of the requirements for the degree of Master of Science in Computer and Information Engineering.

Khalid Al-Khateeb
Head, Department of Electrical and Computer Engineering.

This dissertation was submitted to the Kulliyyah of Engineering and is accepted as partial fulfillment of the requirements for the degree of Master of Science in Computer and Information Engineering.

Ahmad Faris Ismail
Dean, Kulliyyah of Engineering

DECLARATION

I hereby declare that this dissertation is the result of my own investigations, except where otherwise stated. Other sources are acknowledged by giving explicit references and a bibliography is appended.

Name: Mohamed Awad Elshaikh Mohamed

Signature.............................. Date 12-10-2001........

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

DECLARATION OF COPYRIGHT AND AFFIRMATION OF FAIR USE OF UNPUBLISHED RESEARCH

Traffic Shaping and Congestion Control in A Diffserv Domain.

No part of this unpublished research may be reproduced, stored in a retrieval system, or transmitted, in any means, electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the copyright holder except as provided below.

1. Any material combined in or derived from this unpublished research may only be used by others in their writing with due acknowledgement.

2. IIUM or its library will have right to make and transmit copies (print or electronic) for institutional and academic purposes.

3. The IIUM library will have the right to make, store in a retrieval system and supply copies of this unpublished research if requested by other universities and research libraries.

Affirmed by Mohamed Awad Elshaikh

.... Signature

...16.10.2001........ Date

ACKNOWLEDGEMENTS

Firstly I would like to thanks Malaysia and the IIUM for offering me the chance of gaining knowledge in this country in execution of the exchange program between IIUM and the higher education in Sudan. Then my grateful thanks goes to Sudan University of Science and Technology (SUST).

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

Further I must record my thanks to my co-supervisor Dr. Junaid Zubairi for his valuable help on putting this research on reality.

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

I express appreciation to my external and internal examiners for their effort and time and their comments and suggestions.

I acknowledge the several helpful discussions with my colleagues especially network research lab members.

I express warm appreciation to my teachers through 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 academic life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYMS

ABR	Available Bit Rate
AF	Assured Forwarding
AF PHB	Assured Forwarding Per Hop Behavior
AS	Assured Service
ATM	Asynchronous Transfer Mode
BA	Behaviour Aggregate
BE	Best Effort
BISDN	Broadband Integrated Service Digital Network
CAC	Connection Admission Control
CBR	Constant Bit Rate
CBS	Committed Burst Size
CDV	Cell Delay Variation
CIR	Committed Information Rate
CIR_th	Committed Information Rate Threshold
CLP	Cell Loss Priority
CPE	Customer Premises Equipment
CRCS	Composite Rate Control Scheme
DE PHB	Default Per Hop Behaviour
Diffserv	Differentiated Service
DS	Differentiated Service
DSCP	Differentiated Service Code Point
dSRED	Differentiated Service Random Early Detection
EARnew	Estimated Average Rate (New)
EARold	Estimated Average Rate (Previous)
EBS	Excess Burst Size
EF	Expedited Forwarding
EF PHB	Expedited Forwarding Per Hop Behavior
FEC	Forwarding Equivalence Class
FIFO	First In First Out
GLB	Generalize Leaky Bucket
IETF	Internet Engineering Task Force
Intserv	Integrated Service
IP	Internet Protocol
IPv4	Internet Protocol Version 4
IPv6	Internet Protocol Version 6
IS	Integrated Services
LAN	Local Area Network
LB	Leaky Bucket
LBF	Leaky Bucket With Feed Back
LSP	Label Switch Path
LSR	Label Switching Router
MBR	Mean Bit Rate
MBS	Maximum Burst Size
MCFC	Minimum Cost Flow Control
MF	Multi Field Classifier
MIR	Maximum Information Rate
MIR_th	Maximum Information Rate Threshold

MPLS	MultiProtcocol Label Switching
NNI	Network-Network Interface
NS-2	Network Simulator Version 2
PBR	Peak Bit Rate
PBS	Peak Burst Size
PC	Peak Counter
PHB	Per Hop Behaviour
PIR	Peak Information Rate
PIR_th	Peak Information Rate Threshold
PQ	Priority Queue
PTR	Peak Target Rate
QoS	Quality of Service
RAS	Rate Adaptive Shaper
READ	Rate Regulation Via Early Adaptive Detection
RED	Random Early Detection
RFCs	Request For Comments
RSVP	Resource Reservation Protocol
RTT	Round Trip Time
RT-VBR	Real Time Variable Bit Rate
SLA	Service Level Agreement
srRAS	Single Rate (Rate Adaptive Shaper)
srTCM	Single Rate Three Color Marker
TB	Token Bucket
TBF	Token Bucket Filter
TCA	Traffic Conditioning Agreement
TCP	Transfer Control Protocol
TE	Terminal Equipment
trRAS	Tow Rate (Rate Adaptive Shaper)
trTCM	Two Rate Three Color Marker
tswTCM	Time Slide Window Three Color Marker
UNI	User-Network Interface
UPC/NPC	Usage/Network Parameter Control
VBR	Variable Bit Rate
VCC	Virtual Channel Connection
VCI	Virtual Channel Identifier
VP	Virtual Path
VPC	Virtual Path Connection
VPI	Virtual Path Identifier
WAN	Wide Area Network
WFQ	Weighted Fair Queue

CHAPTER 1

INTRODUCTION

1.1 Introduction

Traditionally, the Internet has provided only best effort service to every user with out any consideration to any requirements such as delay jitter and loss. Because every user receives the same level of service, congestion in the network often results in serious degradation for applications that require some minimum amount of bandwidth to work properly. 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, video on demand and interactive multimedia. Thus an interest has developed in having the Internet to provide some degree of Quality of Service (QoS). To provide different QoS Commitments, the Internet Engineering Task Force (IETF) developed many technologies that require resources such as bandwidth and buffers to be explicitly reserved for a given data flow to ensure that the application receives its requested QoS.

This dissertation is divided into six different chapters. Chapter one gives a general introduction and describes the problem statement or dissertation statement and gives a literature review of the current works on the same area. Chapter two gives a background to various new technologies such as MPLS, ATM, Intserv, Diffserv. Chapter three gives an overview to network traffic, congestion control, traffic shaping, and overview of different shapers, markers, and meters schemes such as leaky bucket, token bucket, srTCM, trTCM, RAS, TSW. Chapter four gives an introduction to simulation software NS-2 that has been used as a simulation tool for the current problem; also it contains the experimental Setup for the problem. The result of the experiments and discussion has been shown in chapter five. Finally, chapter 6 gives the conclusion and recommendations for future research.

1.2 Literature Review

Many work and papers have looked at traffic shaping and congestion control in Asynchronous Transfer Mode (ATM) networks, but a few investigate the same problem with the new diffserv architecture.

In 2000, Sahu, Nain, Towsley and Diot examined whether it is possible to provide service differentiation among a set of TCP flows by choosing appropriate marking profiles for each flow and under which circumstances the marking profiles are able to influence the service that a TCP flow receives and how to choose a correct profile to achieve a given service level. The paper has derived an analytical model for determining the achieved rate of a TCP flow when edge routers use a token bucket packet marking and core routers use active queue management for packet dropping. It is for computing the send rate as a function of token bucket parameters. A token bucket marking profile for 2-color is described by (green (confirm) and red (violate)) they assume:

A: token rate

B: token bucket size

P_1, P_2: loss probability for green and red packet respectively

T: average round trip time for the TCP connection

R: Rate

The paper considers two cases

1. Under Subscription case: $P_1=0$, $0<P_2<1$

$$R = \begin{cases} (A + \text{sqrt}(A2 + 6/P_2T2))/2 & :A <= 1/T\ (\text{sqrt}(2(B+1/P_2)) + 2\ \text{sqrt}(2B)) \\ 3A/4+(3\text{sqrt}(B+1/P_2)/(2\text{sqrt}(2T)) & : A>1/T\ (\text{sqrt}(2(B+1/P_2)) + 2\ \text{sqrt}(2B)) \end{cases}$$

2. Over Subscription case: P1>0, P2=1

$$R = \min \left[A,\ (3(A+\text{sqrt}(2B)/T))/4,\ 1/T(\text{sqrt}(3/2P_1)) \right]$$

The paper validates this model using NS-2 and then used this validated model to examine the effect of token bucket parameters on achieved rate and it concludes that the achieved rate depends on the assured rate and bucket size in a non-linear manner.

In 2000, Li. N. Borrego and Li. S. presented a traffic conditioner based on rate Regulation via Early Adaptive Detection (READ). The paper described the emergence of diffserv and gives an overview for its architecture and about Assured Forwarding (AF) and Expedited Forwarding (EF) Per Hop Behaviour (PHB), and then gives a description of the new scheme READ (Fig 1.1).

There are many benefits of choosing the shaper to regulate TCP throughput:

1. The shaper is the first one to know when the TCP sending rate exceeds the threshold to trigger a packet drop.

2. The FIFO queue let the shaper remembers the accumulated past arrival rate once the arrival rate exceeds the shaping rate, without any additional information the shaper has enough information to decide when a packet should be dropped.

3. The queue makes it easy to use a drop from front strategy to speed up the control.

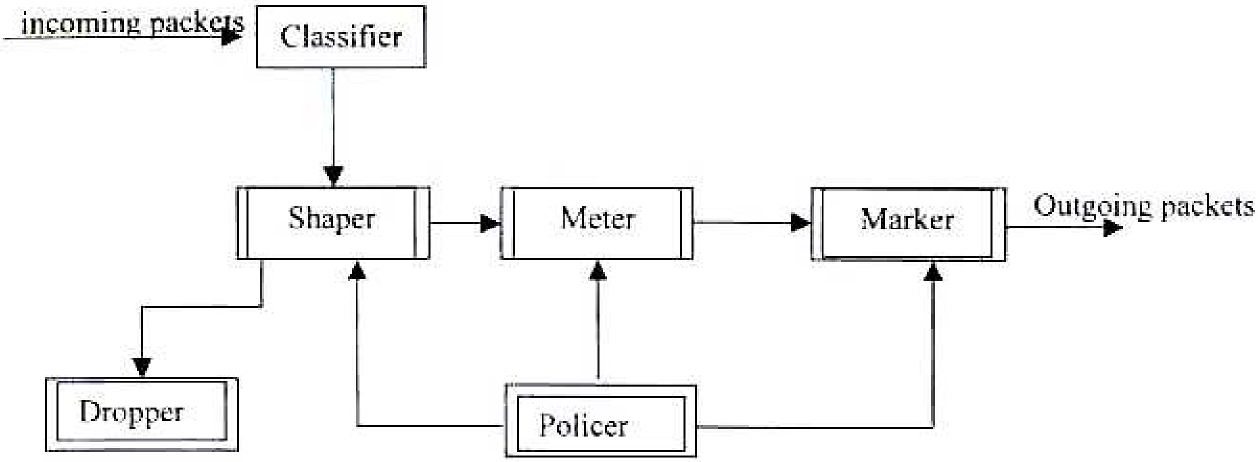

Fig. 1.1 The structure of READ

Authors suggested adding the feedback mechanism as an optional component to the scheme (READ) in order to increase the overall performance. The simulation study shows the effectiveness of READ at boundary nodes within diffserv domain and also shows that READ significantly outperforms Rate Adaptive Shaper (RAS). Also the simulations proved that usage of a feedback mechanism with READ will improve the TCP performance but using as adaptive marker doesn't give a significant improvement.

In 1995, Chan and Liu investigated different types of policing functions including the well-known leaky bucket scheme (LB), Generalized leaky bucket (GLB), peak counter scheme (PC), and their modified version leaky bucket with feedback (LBF). The GLB scheme and PC scheme are both effective in shutting of a single abusive source, in the limiting case PC is very similar to the GLB scheme. However, PC requires more complex implementation that makes GLB the more desirable scheme.

When multiple sources are considered for LB, GLB and LBF, both GLB and LBF show improvement, but GLB perform better in handling packet loss sensitive traffic and LBF perform better for burst loss sensitive traffic.

In 1996, Standfield and Carvey in their simulation showed the effectiveness of the leaky bucket in controlling the cell traffic into the network, and that the two significant design parameters (token refresh rate, and token buffer size) are capable of providing the necessary policing mechanism for the traffic. An input queue can be added in order to provide shaping or smoothing to the cell traffic. The additions of the input queue have the effect of decreasing the cell loss probability but increasing the cell delay.

In 2000, Ferraria focused on the mechanisms used to implement the Expedited Forwarding Per Hop Behavior (EF PHB). The paper analyzed these mechanisms with respect to two main QoS parameters: delay and delay variation. The paper studied the impact of buffering within a diffserv node and compares the effectiveness of several different scheduling algorithms for treatment of EF traffic under different network configurations with a range of different parameters and traffic conditions for the analysis of classification, metering, packet marking, policing and scheduling. It shows that, according to the results, delays are function of many PHB parameters such as EF queue size and the EF queue service rate, as well as a diffserv node parameters such as the transmission queue size, the average background packet length and the EF scheduling algorithm adopted. Also it shows that, priority queuing has a better delay characteristics than Weighted Fair Queue (WFQ) but no significance different in jitter.

In wide area tests, jitter is particularly influenced by the degree of EF flow aggregation. Also it shows that the aggregation of multiple combinations of EF flows can produce burstiness and consequently cause heavy EF packet loss.

In 1998, Schmid, Scott, Hutchison and Froitzheim explored the effectiveness of the next generation Internet protocol, IPv6, in sensitive real time traffic media streaming; especially live audio over the Internet. It determines the efficiency gain for packet classification. It used the audio streaming application called web audio which make use of Resource Reservation Protocol (RSVP) of the IETF'S Integrated Service (Intserv) architecture to achieve QoS for audio stream flow. Label based packet classification within RSVP decrease the processing cost in routers in order of 3-6 times with respect to standard IPv6 and about 2-4 times with regards to IPv4 classification. The results vary depending on the available processing architecture. As a result, utilization of the flow label with resource reservation enables routers to handle significantly more flows before reaching their processing limit and reduces the end to end delay due to simpler and hence, faster packet processing at each intermediate router along the transmission path.

In 1998, Golestani and Bhattacharyya have formulated an end-to-end congestion control as a global optimization problem. Based on this formulation, a class of minimum cost flow control (MCFC) algorithms for adjusting session rates or window sizes are proposed. Two algorithm versions are discussed, the first one is a coarse version (coarse realization) geared towards implementation in the current Internet, relying on the end-to-end packet loss observations as indication of congestion. The second one is a more complete version (exact realization) which anticipates an session can get explicit congestion information through a concise probing mechanism; some of data packets of each session include a short congestion field that is modified by each switch that the packet visit. The paper also has provided a comparison between the MCFC algorithm and some of the congestion control schemes previously proposed for the Internet such as TCP congestion control, binary feedback scheme, random early detection and the dynamic adaptive window.

In 2000, Bhaniramka, Sun and Jain compared the service received by TCP and UDP flows when they share a link or MultiProtocol Label Switching (MPLS) traffic trunks. The simulation results shows that for UDP and TCP mixed in a trunk, if UDP (congestion-unresponsive) flows increase their rates, the TCP (congestion-responsive) reduce service and affect the overall throughput. Therefore, different types of flows should be isolated in different trunks in order to guarantee quality of service. The trunks need to be end-to-end, otherwise the advantages of isolation in other parts of the network will be reduced significantly.

In 2000, Ferrarib evaluated Differentiated Services (DS) QoS architecture. The experiments were setup in presence of EF aggregation (without traffic conditioning), the results shows that the limitation of EF load of the aggregation degree and the presence of the small packet sizes greatly contribute to minimize the occurrence of the occasional bursts, and the increasing in the EF queue size has a minor effect on delay. In the case of limited load the instantaneous packet delay variation is minimized. Also it shows that the weighted fair queue scheme (WFQ) is less burstiness-prone if the EF departure rate approximately equals the arrival rate, while it converges to priority queue (PQ) when the EF queue weight and the corresponding service rate increase, thus with WFQ a tradeoff between one delay minimization and burstiness avoidance has to be identified.

In 1998, Mitchell, Van and Place used linear algebra queuing theory to model the behavior of token leaky bucket as a traffic shaper. It focuses on the behavior of the token leaky bucket mechanism with respect to second order statistics of the departure process under various correlated cell arrival and deterministic token arrival distribution. It derives an expression for the lag-k auto-correlation of the inter departure times for the token leaky bucket cell departure stream and produced numerical examples for the lag-k auto-correlation of the token leaky bucket inter-departure times for several different arrival distributions exhibiting for both positive and negative auto-correlations. This study shows that, slight correlation does exist in the output process of a token leaky bucket even when the input processes are renewal processes, and the analysis of departure process of the token leaky bucket under non renewal cell arrival distributions which are shown to be correlated, shows that the token leaky bucket provides a little traffic shaping with regard to correlations to alleviate detrimental impact of highly correlated arrival traffic on buffer requirements.

In 1998, Goyal, Jain, Fahmy and Narayanaswamy described the OPNET models that has been developed for ATM and Available Bit Rate (ABR) design and analysis in ATM, which it has been designed to fairly distribute all unused capacity to data traffic. It has been specified in the ATM forum's traffic management (TM4.0) standard.