

~~•••~~

※ 支援 QoS 路由器的封包排程與服務分流之設計與實作 ※



1

39

至

孫

付作

- ☐

管 理

年

行政院國家科學委員會專題研究計劃成果報告

支援QoS路由器的封包排程與服務分流之設計與實作

Design and Implementation of Packet Scheduler and Service Classifier for Broadband Internet QoS Router

計畫編號：NSC89-2213-E-002-079

執行期限：88年8月1日至89年7月31日

主持人：孫雅麗 國立台灣大學資訊管理學系副教授

E-mail: sunny@irn.ntu.edu.tw Fax:02-23621327

一、中文摘要：

在本篇研究中，我們定義在差異化服務網路 (Differentiated Services Networks) 的一種新的服務模式叫 “Statistical Expedited Forwarding Service”，該服務可以保障封包延遲和封包遺失率不超過某個上限。我們針對此一新服務模式提出一個新的動態頻寬及緩衝區資源分配的方法與機制。我們的設計理念是希望能在連接提供差異化服務的骨幹網路的邊際路由器 (edge router) 上提供使用者整合性服務般的效能保證，卻不需在網路的核心路由器 (core router) 維持 per-flow 的資訊。目標應用是同時對封包延遲和封包遺失兩項服務品質敏感的應用，如網路電話、即時影像傳輸和線上電子商務等等提供 QoS 的保證。

為了保障封包延遲和封包遺失不超過上限值，我們根據給定的 QoS 要求，來替每個類別保留 “effective resource” (包括頻寬和緩衝區大小)。除此之外，我們更進一步探討在同一個類別中所有資料流的公平性。數個聲音和影像應用的實驗顯現出我們提出機制的適用性。

關鍵詞：差異化服務、傳輸服務品質保證、資源管理

Abstract

In this paper we propose a new scheme Statistical Expedited Forwarding to support a new class of service that guarantees bounded delay and bounded packet loss performance. The goal of the proposed scheme is to provide the Integrated Services type of performance guarantees, but *without* maintaining per-flow state in the network core. The target applications are both delay-sensitive and loss-sensitive like IP telephony, video applications and emerging on-line electronic business-related applications. In order to guarantee delay and loss bounds, the effective resource (consisting of bandwidth and buffer) is assigned to each class according to the given QoS requirement. In addition, we further study the fairness among flows in a class. The experiments on several voice and video applications show the strength of the proposed scheme.

Keyword: Differentiated Services, QoS, resource management

Part I: 動態頻寬及緩衝區資源分配 for “Statistical Expedited Forwarding” 服務模式

1. 研究動機與目的

現今的網際網路（Internet）只提供一種稱作best-effort的服務，在這個架構下每個使用者都得到平等的對待。但是已經有越來越多的應用需要不同的服務品質，例如即時應用和檔案傳輸所需要的網路服務品質就大不相同，前者要求的是低延遲，而後者的最大考量則是正確性，所以我們可以看到best-effort服務架構改進的必要性。

整合性服務（Integrated Services）和資源保留協定（RSVP）於是應運而生，但是整合性服務架構的擴充性（scalability）問題被普遍認為是其無法在短期內被廣泛應用的重要原因，這點在需要快速傳輸封包的骨幹路由器上尤其明顯。

差異化服務（Differentiated Services）則是目前另一個提供服務品質保證（QoS）的架構。在這個架構下，我們將路由器區分為邊緣路由器（edge router）和核心路由器（core router）兩種類型。複雜的工作都只發生在邊緣路由器，讓核心路由器可以快速處理封包，以改善整合性服務擴充性的問題。

目前差異化服務定義了 EF（Expedited Forwarding）、AF（Assured Forwarding）兩種類型的服務，這兩者都是“質化”的服務品質保證，使用者並沒有辦法“明確”知道自己將接收到的服務品質，而“明確”的數值（如 1%、10ms）對於許多網路應用的設計將有相當大的幫助。

不同網路應用要求的服務品質往往是不相同的，在同一個網路架構下整合不同的應用因而可以帶來許多效益，例如具有時效性的多媒體應用可以較高的優先權先行傳送，而時效性要求較低的一般資料就可以利用多媒體應用停止輸出資料的時候再傳送封包。如果我們再從另一個服務品質指標——封包遺失率來

看的話，多媒體應用可以忍受一定程度的封包遺失，而一般資料對正確性的要求較高，也就是不希望有封包遺失的情況發生，所以我們在分配路由器緩衝區大小的時候，多媒體應用與一般資料可以依據他們不同的需求來分配。由這些例子我們可以看出，在同一個網路架構下，整合了不同的應用將可以更有效率的使用網路資源，避免資源的閒置。

針對以上數點，本研究提出了一個差異化網路下的架構，該架構整合不同的應用，並且針對不同的應用分別提供了“量化”的服務品質保證。並且我們選定以封包最大延遲上限和封包遺失率上限來當作我們的服務品質保證指標。

2. 相關研究

在本研究中我們融合了許多觀念，包括整合性服務[RFC 2205] [RFC 2210] [RFC 2211] [RFC 2212]、差異化服務[RFC 2474][RFC 2475] [RFC 2597] [RFC 2598]、“Summed TSpec”[Schmitt 1999]、“Effective Bandwidth”[Roch 1991]等等，在這裡我們介紹後兩者，我們加以變化並應用了後兩者所提出的方法。

(1) “Summed TSpec”[Schmitt 1999]

[Schmitt 1999]的出發點在於體認到整合性服務的最大問題在於擴充性，而一個解決的方法是把整合性服務中的網路資料流集合（aggregate）起來。那麼分配資源時就只需要針對集合後的網路資料流，而不再是為每個網路資料流個別分配資源，所以降低了整個架構的複雜度。

使用者提出封包最大延遲上限的要求和自己本身交通量特性的描述（這裡使用TSpec，包括平均速率 r 、權杖深度（token depth） b 、最大速率 p 、最大封包大小 M 等參數），來導出所需保留的頻寬，而集合的方法使用RFC 2212中定

義的”Summed TSpec”

(2) “Effective Bandwidth” [Roch 1991]

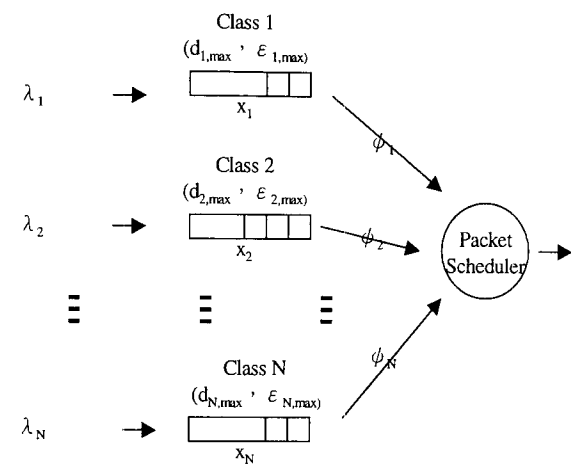
該篇文章中，以數理統計方法來導出網路資料流的”equivalent capacity”，意義是給定網路資料的流量特性和要求的服務品質，那麼網路節點需要保留的頻寬即是”equivalent capacity”。

該篇文章對網路資料流交通流量特性的描述是使用 two-state ON/OFF Markov model服務品質保證的要求則是佇列長度超過給定緩衝區大小機率要小於 ε 。

3. 服務模式之架構設計

本研究中，提出一個以應用為導向的服務模式，整合不同應用（具有不同服務品質保證要求），想法是將差異化網路架構下的路由器排程分成多個佇列，每個佇列代表一種服務品質保證要求的應用，不同佇列之間則以WFQ（weighted fair queueing）來排程以達到分配頻寬和緩衝區大小的目的。（如下表）

Queue	適合的應用	Packet delay (QoS目標)	Packet loss (QoS目標)
Q1	Strict QoS requirement	Max delay bound	No
Q2	real time voice	Max delay bound	Predictable
Q3	real time video	Max delay bound	Predictable
Q4	Data	Reasonable	Very low



整個架構可以用上圖來表示，不同Class表示不同的應用，經過Classifier後，被分配到不同的佇列，然後我們會分配頻寬R和緩衝區B兩項資源給各個佇列。以下我們詳細介紹每個佇列：

- Q1：以RFC2212中的”summed TSpec”來為集合後的資料流一起保留頻寬，每個資料來源需要被shape至符合宣稱的TSpec，然後可以求出所需要的頻寬和緩衝區大小。適用的應用要求不能有封包遺失，和封包要在一定時間內到達，如即時影像醫療。
- Q2：VoIP加入silence suppression後，其行為可以用two-state ON/OFF Markov model來描述，而且Voice封包具有時效性，所以適合Q2該類應用。我們將使用”equivalent capacity”算出需要的頻寬來使得封包遺失機率小於某個值。不同於[Roch 1991]該篇文章的地方是，我們接著再導入最大封包延遲上限這項服務品質指標，來求出所需要的頻寬和緩衝區大小，使得可以同時滿足封最大包延遲上限和封包遺失率上限兩個服務品質保證目標。
- Q3：Video應用經過我們抓取實際應用程式產生的封包資料後，發現也有ON/OFF的特性，所以我們將嘗試

使用和Q2一樣的機制來保留、分配資源，並探討其適用性。

- Q4：Q1、Q2、Q3所使用的緩衝區大小和頻寬都可以明確算出，使用剩下的資源為Q4所有，因為Q4對封包遺失比較敏感而且頻寬是使用Q1、Q2、Q3剩下來的，所以要有較大的緩衝區來暫存封包，但是為了避免封包無止盡的等待或是終於造成封包遺失，整個架構必定要加入允入控制。
- 我們整個架構所希望達到的目標有：增加擴充性、整合不同應用、提供量化服務品質、容易在實際網路上實做。

4. 效能評估

● Voice

“Effective Bandwidth”是以(式1)中各個參數求出所需保留的頻寬c，再配合封包最大延遲Dmax是緩衝區大小x除以頻寬c（式2），兩個等式可以求出要保留的頻寬和緩衝區大小。

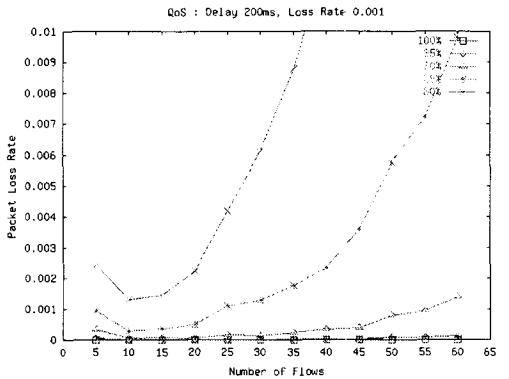
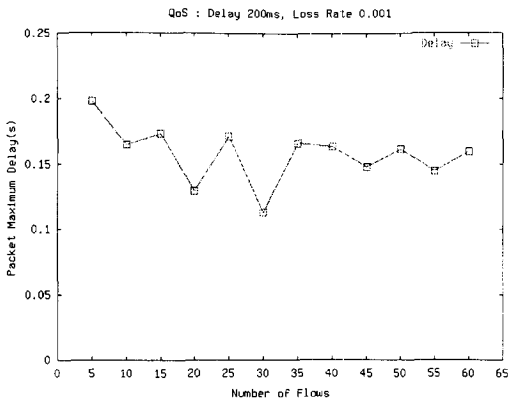
$$c = F(R_{peak}, \rho, b, \epsilon, x).....(1)$$

$$D_{max} = x / c.....(2)$$

如果是多個資料流的情形，我們使用集合後的effective bandwidth是個別的加總，來當作第一個等式(式3)。

$$Cn = N * F(Xn).....(3)$$

依此設定的實驗結果如下兩圖所示，資料流要求的服務品質是封包延遲不超過0.2秒，封包遺失率不超過0.001。我們可以看到不管資料流的數目多寡，都可以達到要求的服務品質。



我們進一步使用RED來增進整個系統的公平性，同時降低連續封包遺失的

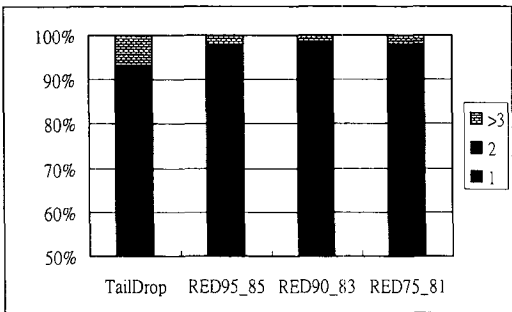
$$D_{max} = Q_{max} / c.....(5)$$

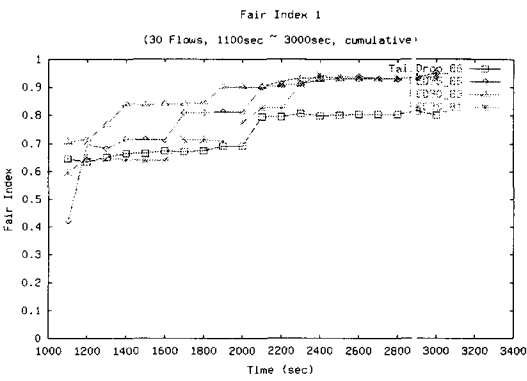
$$Q_{min} = Y * Q_{max}.....(6)$$

機率。

$$c = F(Q_{min}).....(4)$$

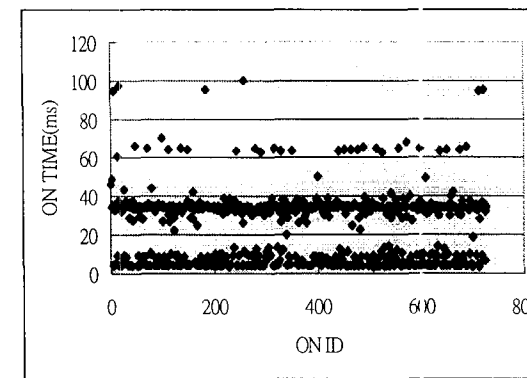
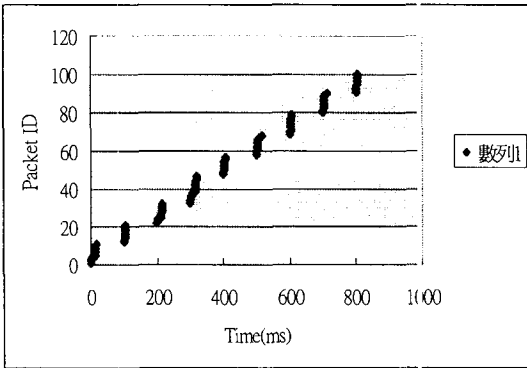
我們使用上述三個式子來求出所需保留的資源，其中Qmax、Qmin是RED中的兩個Threshold。實驗結果如下所示，連續封包遺失的個數大為降低，而公平性指標較快趨近到1(1是最公平的狀況)。





● Video

我們實際量測了使用IP/TV播放儲存的VCD和Live畫面，和使用rmedia player播放VCD，三者都呈現近似下圖的封包到達狀況，所以我們假設皆具有exponential on/off的特性，並使用前述Voice的方法來保留資源，實驗結果顯示我們保留的資源超過實際需要保留的資源，所以使用exponential on/off模型在Video應用上可以當作一個保留資源的上限參考值。我們進一步分析各應用的流量狀況後，顯示和exponential on/off模型有些差距，所以造成我們保留資源過多。



此圖顯示使用IP/TV播放Live的ON持續時間分佈圖，我們可以發現並沒有呈現exponential分配。

5. 結論

在本研究中，我們主要探討了：(一) 使用”Summed TSpec”來為要求封包遺失率為0的應用保留資源。(二) 使用”Equivalent Capacity”來為可以忍受一定封包遺失的網路聲音、影像應用保留資源。並且在保障封包遺失率不超過要求值的同時，封包的最大延遲也不能超過要求值。由以上的實驗結果可以得知，我們提出的以應用為導向的服務模式，的確可以在差異化網路架構下，提供明確的量化服務品質保證。針對同時滿足兩項服務品質保證指標是本研究的特點。

我們的架構適用環境在：(一) 骨幹網路上。(二) 企業網路，管理者可以決定想要保障其服務品質的應用，並為其保留資源，不屬於這些應用的其他網路資料流就歸到Best-Effort佇列。而應用導向的緣故，也適合企業拿來支援某些特定應用。例如在企業內推動視訊會議、分公司間以網路電話聯繫等等。

我們的研究未來的工作有以下數點：

- 在考慮系統的負載能力和應用服務品質保證需求下，制訂適當的類別個數與類別服務品質保證值。
- 經過多路由器後，端點到端點的服務品質保證研究。
- 使用更適當的模型來描述網路影像應用交通流量特性。

參考文獻

● [RFC 2205] R. Braden, Ed., et. al., "Resource Reservation Protocol (RSVP) - Version 1

Functional Specification", RFC 2205, September 1997.

- [RFC 2210] J. Wroclawski, "The Use of RSVP with IETF Integrated Services", RFC 2210, September 1997.
- [RFC 2211] J. Wroclawski, "Specification of the Controlled Load Quality of Service", RFC 2211, September 1997.
- [RFC 2212] S. Shenker, Partridge C., and R. Guerin, "Specification of Guaranteed Quality of Service", RFC 2212, September 1997.
- [RFC 2474] K. Nicols, S. Blake, F. Baker, D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", RFC 2474, December 1998.
- [RFC 2475] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, W. Weiss, "An Architecture for Differentiated Services", RFC 2475, Dec. 1998.
- [RFC 2597] J. Heinanen, F. Baker, W. Weiss, J. Wroclawski, "Assured Forwarding PHB Group", RFC 2597, Jun. 1999.
- [RFC 2598] V. Jacobson, Kathleen Nichols, Kedernath Poduri, "An Expedited Forwarding PHB", RFC 2598, Jun. 1999.
- [Schmitt 1999] J. Schmitt, Karsten, M., Wolf, L., Steinmetz, R., "Aggregation of guaranteed service flows", Quality of Service, 1999. IWQoS '99. 1999 Seventh International Workshop, 1999, pp. 147-155
- [Guerin 1991] R. Guerin, H. Ahnadi, and M. Naghshineh, "Equivalent capacity and its application to bandwidth allocation in high-speed networks," IEEE J. Select. Areas Commun., vol. 9, pp. 968-981, 1991.
- [Sriram 1990] K. Sriram, "Dynamic bandwidth allocation and congestion control schemes for voice and data multiplexing in wideband packet technology", IEEE

International Conference on Communication, 1990, Page(s): 1003-1009 vol.3

- [May 1981] C. E. May and T. J. Zebo, unpublished paper, AT&T Bell Laboratories, December 1981.
- [Bowker 1987] D. O. Bowker, unpublished paper, AT&T Bell Laboratories, Holmdel, New Jersey, 1987.
- [Doulamis 2000] N. D. Doulamis, Doulamis, A. D., Konstantoulakis, G. E., Stassinopoulos, G. I., "Efficient Modeling of VBR MPEG-1 Coded Video Sources", IEEE Transactions on Circuits and Systems for Video Technology, Volume: 10, 1, Feb. 2000, Page(s): 93-112
- [Lazar 1993] A. A. Lazar, Pacifici, G., Pendarakis, D. E., "Modeling Video Sources for Real-Time Scheduling", GLOBECOM '93., 1993, Page(s): 835-839 vol.2
- [Lee 1996] C. B. Lee, Ha, K.B., and Park, R.-H. "Computation of effective bandwidth of aggregated VBR MPEG video traffic in ATM networks using the modified equivalent capacity", ICC '96, Volume: 2, 1996, Page(s): 627-631 vol.2

Part II: Policy-Based QoS 路由 器實作

1. 簡介

為了提供傳輸的服務品質(quality of service, QoS)，擔任封包遞送的路由器(router)必須提供封包傳輸控制的機制，為資料流(flow)保留足夠的資源，並記錄資料流使用資源的狀況。而要達到上述目標，必須在路由器加入三個重要元件：一是傳輸允諾控制(admission control)，二是封包分類器(packet classifier)，三是封包排程器(packet

scheduler)。

我們的研究目的在於建立一個提供傳輸服務控制的路由器的雛型(prototype)，除此之外，我們並導入”Policy”和”Hierarchical link sharing”的觀念，設計了一個稱做”Policy Manager”的系統，讓使用者以圖形化、直覺的操作方式來建立整個網路管理的政策，同時得以掌控該路由器雛形的運作。

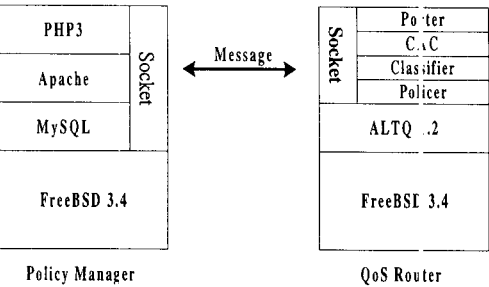
我們並會在此路由器雛型和Policy Manager上進行一系列實驗以驗證整個系統的性能。

2. 系統使用情境

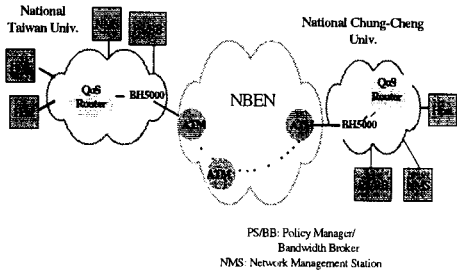
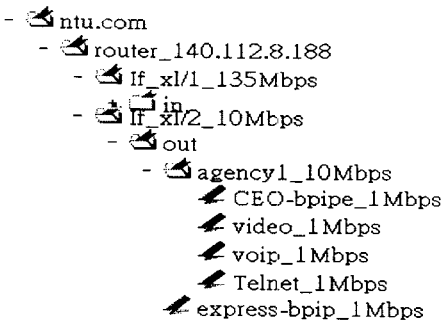
3. 系統架構

在QoS Router系統架構方面，我們選擇以FreeBSD 3.4 Release此作業系統為系統發展的平台，並將ALTQ 1.2安裝進作業系統的核心 (kernel) 中，以作為實作封包分類器及封包排程器的基礎架構。

在”Policy Manager”系統架構方面，我們一樣使用FreeBSD 3.4 Release此作業系統，並安裝MySQL資料庫軟體來儲存政策，整個系統的架構可以用下圖來表示：

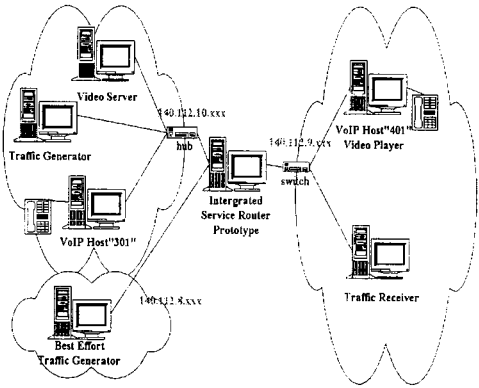


4. Policy Manager系統介面



5. 實驗概述

為驗證整個系統的性能，我們設計了一系列實驗，實驗環境如下圖所示



在我們實驗中，我們由”Policy Manager”建立了階層式的政策，觀察QoS Router 是否能夠根據這些政策正確地被設定和啟動。同時我們使用 Video 和 VoIP 等應用來測試 QoS Router 對服務品質保證的性能。

從實驗結果可以驗證幾個結論：(a) ”Policy Manager”可以讓使用者直覺地、快速得設定階層式網路管理政策。(b) ”Policy Manager”和QoS Router之間可以正確地溝通。(c) QoS Router可以

根據”Policy Manager”設定的各類別服務品質水準給予保障。

6. 結論

本研究以ALTQ為系統架構，實作了QoS Router雛形。並設計了”Policy Manager”來管理QoS Router，同時我們也設計數個實驗驗證這整個系統的正確性與效能，由實驗結果發現整個系統可以保障各類別要求的服務品質。

7. 參考文獻

- [1] Kenjiro Cho, "A Framework for Alternate Queueing: Towards Traffic Management by PC-UNIX-Based Routers", Proceedings of USENIX 1998 Annual Technical Conference, New Orleans LA, June 1998
- [2] Wroclawski, J., "The Use of RSVP with IETF Integrated Services", RFC 2210, Sep. 1997
- [3] Parekh, A., Gallager, R., "A generalized processor sharing approach to flow control – the single node case", IEEE/ACM Transactions on Networking, Vol. 1, No.3, pp.344-357, June 1993
- [4] Demers, A., Keshav, S., and Shenker, S., "Analysis and simulation of a fair queueing algorithm", Journal of Internetworking Research and Experience, pp.3-26, Oct. 1990
- [5] Blake, S., Black, D., Carlson, M. Davies, E., Wang, Z., Weiss, W., "An Architecture for Differentiated Services", RFC 2475, Dec. 1998