

TAMPERE UNIVERSITY OF TECHNOLOGY  
Department of Information Technology

MERJA MÄKI-SÄNTTI  
INTRA- AND INTERDOMAIN QOS IN THE DIFFERENTIATED SERVICES  
ARCHITECTURE  
Master of Science Thesis

Subject approved by the department council  
on 7.5.2003

Supervisors: Prof. Pekka Loula

Prof. Jarmo Harju

## **PREFACE**

This Master of Science thesis was done as a part of the ICEFIN Research Laboratory at Tampere University of Technology, Institute of Communications Engineering. ICEFIN research laboratory concentrates on network level issues that have turned out to be problematic in current and future Internet. The main areas of research are Quality of Service, WLAN, multicast and IPv6. This work was done during the years 2002-2003.

I pay my respect and thanks especially to my advisors Pekka Loula and Jarmo Harju for their guidance and support during this work. I would also like to thank Virpi Laatu for giving suggestions, insight and support during this project.

This work would not have been possible without continuous support and understanding of me parents Anja and Heikki. Also the love and friendship of my sister Sirpa and brother-in-law Ilkka gave me the trust to walk through the hard times. I want also to thank my friends for their empathy and love.

Finally I want to thank my beloved men Jarmo, Iikka and Sameli. You gave me the real world full of moments. You were the source of love and joy that helped me to go through this work.

Pori, June 24, 2003

Merja Mäki-Säntti  
Vierutie 9-11 as 7  
28600 PORI  
Finland

## TABLE OF CONTENTS

<b>PREFACE .....</b>	<b>I</b>
<b>TABLE OF CONTENTS.....</b>	<b>II</b>
<b>ABSTRACT .....</b>	<b>IV</b>
<b>TIIVISTELMÄ.....</b>	<b>V</b>
<b>LIST OF ACRONYMS.....</b>	<b>VI</b>
<b>1. INTRODUCTION .....</b>	<b>1</b>
1.1 Background .....	1
1.2 Objectives and the outline of thesis.....	3
<b>2. DIFFERENTIATED SERVICES .....</b>	<b>5</b>
2.1 An overview of Differentiated Services model.....	5
2.2 Terminology related to a Differentiated Services domain .....	6
2.3 Per-Hop Behaviors (PHBs) .....	7
2.4 Traffic classification and conditioning.....	8
2.5 Standardized Per-Hop Behaviors .....	10
2.5.1 Default PHB .....	10
2.5.2 Class Selector PHB .....	10
2.5.3 Expedited Forwarding PHB .....	11
2.5.4 Assured Forwarding PHB .....	12
<b>3. SERVICE LEVEL AGREEMENT .....</b>	<b>15</b>
3.1 Prevalent types of SLA.....	15
3.2 Network level SLA.....	16
3.3 The content of network level SLA .....	18
3.4 SLA and Differentiated Services.....	20
3.5 Static SLA .....	22
3.6 Dynamic SLA.....	22
<b>4. DIFFERENTIATED SERVICES AS A TOOL FOR OPERATOR.....</b>	<b>24</b>
4.1 Intradomain QoS .....	25
4.1.1 IntServ over DiffServ .....	25
4.1.2 DiffServ over MPLS .....	26
4.1.3 Less than best effort class of service .....	28
4.1.4 Intradomain metrics and monitoring .....	29
4.1.5 Vision of Differentiated Services Network.....	30
4.2 Interdomain QoS .....	35
4.2.1 SLAs between network service provides .....	35
4.2.2 Per-Domain Behavior.....	36
4.2.3 Exchange points .....	39
4.2.4 Policy-based networking in a DiffServ network .....	40
4.2.5 Interdomain metrics and monitoring .....	42
4.2.6 Means to build the QoS .....	43

<b>5. DIFFERENTIATED SERVICES FROM THE CUSTOMER POINT OF VIEW</b> .....	<b>44</b>
5.1 User-metrics and monitoring.....	44
5.2 QoS in customer network.....	45
5.3 Fairness and the end-to-end QoS with TCP and UDP.....	47
<b>6. CONCLUSIONS</b> .....	<b>49</b>
<b>REFERENCES</b> .....	<b>51</b>

## **ABSTRACT**

### **TAMPERE UNIVERSITY OF TECHNOLOGY**

Department of Information Technology

Information Technology (Pori)

**MERJA MÄKI-SÄNTTI:** Intra- and interdomain QoS in the Differentiated Services architecture

Master of Science thesis: 55 pages

Supervisors: Prof. Pekka Loula and Prof. Jarmo Harju

August 2003

Keywords: Differentiated Services, DiffServ, Service Level Agreement, SLA, intradomain QoS, interdomain QoS, QoS

UDC: 004.738.5, 366.544, 658.56

The traditional Internet was designed to support the Best Effort (BE) service that does not distinguish between applications or users. Due to the business use and real-time applications the Internet BE service does not satisfy all the users in today's Internet. The focus of the Internet is on a converged network, where more advanced and time-critical applications (such as videoconferencing, Internet telephony, video on demand) can be served in terms of Quality of Service (QoS).

The Differentiated Services (DiffServ) concept has caught a lot of attention as a means to bring the QoS into the Internet. DiffServ offers different levels of QoS to customers according the Service Level Agreement (SLA). The ultimate goal of QoS is to provide end-to-end QoS. Since there are already a number of proposals for intradomain end-to-end QoS mechanisms, the question of providing interdomain end-to-end QoS is an interesting research direction. The deployment of end-to-end service requires more than technical decisions. It is necessary also to deal with issues such as standardized SLA definitions and forms, QoS monitoring, peering SLAs and service negotiation.

This thesis in hand introduces the DiffServ concept. Today's SLAs are discussed and future SLAs are highlighted. DiffServ is dealt with both from point of view of the operator and the customer. The thesis addresses the shortcomings of existing QoS and suggests some practical guidelines for bringing QoS by means of DiffServ and SLAs into intradomain and interdomain networks. The work concentrates on techniques that offer QoS in intra- and interdomain networks. Such techniques are for example different service classes, SLA negotiation, Per-Domain Behavior and policy-based networking.

## TIIVISTELMÄ

### TAMPEREEN TEKNILLINEN YLIOPISTO

Tietotekniikan osasto

Tietotekniikka (Pori)

MERJA MÄKI-SÄNTTI: Intra- and interdomain QoS in the Differentiated Services architecture

Diplomityö: 55 s

Tarkastajat: prof. Pekka Loula and prof. Jarmo Harju

Elokuu 2003

Avainsanat: Differentiated Services, DiffServ, Service Level Agreement, SLA, intradomain QoS, interdomain QoS, QoS

UDK: 004.738.5, 366.544, 658.56

Perinteinen Internet suunniteltiin tukemaan Best Effort (BE) - palvelua, joka ei toiminnassaan erottele sovelluksia eikä käyttäjiä. Internetin lisääntynyt käyttö liiketoiminnan tarpeisiin sekä reaaliaikaiset sovellukset edellyttävät parempaa palvelunlaatua kuin BE palvelu tarjoaa. Internetin painopiste on verkossa, jossa yhä kehittyneemmät ja aikakriittiset sovellukset (kuten videokonferenssi, Internet puhelut) voidaan toteuttaa käyttäen erilaisia laatukriteereitä.

Differentiated Services (DiffServ) konsepti on saanut osakseen paljon kiinnostusta, koska se on lupaava tapa toteuttaa palvelunlaatua Internetissä. DiffServ tekee mahdolliseksi käyttää erilaisia palvelunlaatusojuja, jotka perustuvat palvelunlaatusopimukseen. Useita ehdotuksia on tehty verkon sisäisen palvelunlaadun toteuttamiseksi päätepisteiden välille, joka onkin mielenkiintoinen tutkimussuunta. Päätepisteiden välinen palvelunlaatu vaatii kuitenkin muutakin kuin vain teknisiä ratkaisuja. On tarpeellista käsitellä myös sellaisia asioita kuten standardisoidujen palvelusopimusten määrittely ja muoto, palvelunlaadun valvonta, verkkojen väliset palvelusopimukset ja palvelunlaatusopimuksen neuvottelu.

Tämä diplomityö esittelee DiffServ konseptin. Konseptia käsitellään operaattorin ja asiakkaan näkökulmasta. Työssä pohditaan nykyisiä palvelunlaatusopimuksia, sekä kuvataan tulevaisuuden palvelunlaatusopimuksia. Käsillä oleva diplomityö esittelee nykyisen palvelunlaadun puutteita ja ehdottaa joitakin käytännön suosituksia, joiden avulla verkkojen sisällä ja välillä voidaan toteuttaa palvelunlaatu. Sellaisia ovat esimerkiksi erilaiset palveluluokat ja palvelusopimusten neuvottelu.

**LIST OF ACRONYMS**

AF PHB	Assured Forwarding Per-Hop Behavior
AQM	Active Queue Management
AR PDB	Assured Rate Per Domain Behavior
ATM	Asynchronous Transfer Mode
BA	Behavior Aggregate
BB	Bandwidth Broker
BE	Best Effort
BE PDB	Best Effort Per-Domain Behavior
BE PHB	Best Effort Per-Hop Behavior
CBQ	Class Based Queuing
CBR	Constant Bit Rate
CBS	Committed Burst Size
CIR	Committed Information Rate
CLP	Cell Loss Priority
COPS	Common Open Policy Service
CPE	Customer equipment
CS PHB	Class Selector Per-Hop Behavior
CTR	Committed Target Rate
CU	Currently Unused
DE	Discard Eligibility
DLCI	Data Link Connection Identifier
DS	Differentiated Services
DSCP	Differentiated Services CodePoint
EBS	Excess Burst Size
ECN	Explicit Congestion Notification

EF PHB	Expedited Forwarding Per-Hop Behavior
E-LSP	Exp-Inferred-PSC Label Switching Path
EWMA	Exponential Weighted Moving Average
EXP/LU	Experimental and Local Use
FEC	Forwarding Equivalence Class
FTP	File Transfer Protocol
GPS	Global Positioning System
HTTP	HyperText Transfer Protocol
IETF	Internet Engineering Task Force
IP	Internet Protocol
IntServ	Integrated Services
ISP	Internet Service Provider
IXP	Internet Exchange Point
LAN	Local Area Network
LBE PHB	Less than Best Effort Per-Hop Behavior
LDAP	Lightweight Directory Access Protocol
LE PDP	A Lower Effort Per-Domain Behavior
LE PHB	A Lower than Best-Effort Per-Hop Behavior
L-LSP	Label-Only-Inferred-PSC Label Switching Path
LPDP	Local Policy Decision Point
LSP	Label Switching Path
LSR	Label Switching Router
MAI	Metropolitan Area Interconnectivity
MB	Megabyte
Mbps	Megabits per second
MDT	Mean Down Time
MF	Multi-Field classifier

MOS	Mean Opinion Score
MPLS	Multiprotocol Label Switching
MTBF	Mean Time Between Failures
MTTR	Mean Time to Repair
MTU	Maximum Transfer Unit
NSP	Network Service Provider
OA	Ordered Aggregate
PBN	Policy-based networking
PBS	Peak Burst Size
PDB	Per-Domain Behavior
PDP	Policy Decision Point
PEP	Policy Enforcement Point
PHB	Per-Hop Behavior
PIR	Peak Information Rate
PMT	Policy Management Tool
PQ	Priority Queuing
PSC	PHB Scheduling Class
PTR	Peak Target Rate
QoS	Quality of Service
RAP	Resource Allocation Protocol
RFC	Request For Comments
RED	Random Early Detection
RIO	Random Early Detection with In-and-Out
RSVP	Resource ReServation Protocol
RTT	Round Trip Time
SLA	Service Level Agreement
SLS	Service Level Specification

SMTP	Simple Message Transfer Protocol
SNMP	Simple Network Management Protocol
srTCM	Single Rate Three Color Marker
TA	Traffic Aggregate
TC	Traffic Class
TCA	Traffic Conditioning Agreement
TCP	Transmission Control Protocol
TCS	Traffic Conditioning Specification
TOS	Type of Service
trTCM	Two Rate Three Color Marker
TSWTCM	Time Sliding Window Three Color Marker
TTL	Time-To-Live
UDP	User Datagram Protocol
VBR	Variable Bit Rate
VCI	Virtual Channel Identifier
VoD	Video on Demand
VoIP	Voice over Internet Protocol
VPI	Virtual Path Identifier
VPN	Virtual Private Network
VW PHB	Virtual Wire Per-Hop Behavior
WAN	Wide Area Network
WFQ	Weighed Fair Queuing
WG	Work Group
WRED	Weighted Random Early Detection
WRR	Weighted Round Robin
WWW	World Wide Web

## 1. Introduction

This chapter gives an introduction to the work including the basic issues related to the services that Internet provide. This chapter also reviews the background in the Internet QoS (Quality of Service) area. This is followed by the objectives and novelty of the thesis. Finally the outline of the thesis is presented.

### 1.1 Background

The Internet was originally designed to support a best effort (BE) service. The BE service does not make any difference among users and applications. BE service is sufficient for elastic traffic, such as email, file transfer and Web traffic, because applications have somewhat loose bounds on QoS target (e.g. time requirements). However, with time the services that the Internet provides have been changing. Beside traditional data transfer services, more advanced and performance-critical applications, such as videoconferencing, video on demand (VoD) and Internet telephony (VoIP, Voice over IP) have emerged. These new applications often have strict quality requirements for bandwidth, packet loss rate, delay and delay variation called jitter.

It is widely accepted that traditional BE service does not meet the delivery requirements for these new real time and performance-critical services. There is an increasing demand for providing QoS support in the Internet. Multimedia communication has gained increasing attention, both from the user side and the service provider side. Internet service providers (ISPs) are increasingly pressured by their customers to provide a range of QoS levels in the Internet. QoS measures are not just for real time communications but also for the transfer of documents of various data types that may have different bandwidth demands.

The common way to meet the bandwidth and QoS requirements is to use the overprovisioned network. However, the just added excess bandwidth can quickly be consumed by an increasing amount of traffic. Despite the use of overprovisioning, some advanced applications need QoS guarantees in order to be effective for their users, especially in the case of instantaneous traffic peak loads. Network connections will be bursty and the total network load varies much across the day, which means that even if there is enough capacity on average, it will be necessary to either maintain substantial excess capacity almost all of the time or to introduce mechanisms to manage situations when the peak traffic from all users occurs at the same time.

Within the last ten years, the Internet Engineering Task Force (IETF) has proposed a number of architectures to provide QoS in IP (Internet Protocol) networks, such as Integrated Services (IntServ) [1], Differentiated Services (DiffServ, DS) [2] and Multiprotocol Label Switching (MPLS) [3]. This thesis mainly focuses on the DiffServ based solutions. The DiffServ standard is a quite open and flexible framework, only the basic mechanisms and forwarding treatments are standardized. This gives a great flexibility to operators to choose suitable signaling, provisioning, policy functions and charging schemes for their own purposes. DiffServ framework assumes that there is a bilateral service level agreement (SLA) between customers and their respective service providers. A customer can be an end-user, an organization or another service provider. Typically a small or medium sized organization is the customer of an ISP and the ISP is the customer of a network

service provider (NSP). The contents of SLAs are different between a customer and an ISP than between service providers. An SLA between a customer and its ISP for example defines the traffic profiles and describes the services provided by the service provider to the customers. An SLA between service providers includes above-mentioned issues, but concentrates on throughput of traffic. In a broader sense, SLAs clarify responsibilities and alleviate communication between contracting parties.

So far, research and development efforts in the area of technologies for QoS support are mainly concentrated on the definition of QoS architectures and protocols. Most customers would accept to be charged on a service basis, if performance guarantees are sufficiently offered. An end-user sees QoS through applications. However, there is not currently a standardized form for specification and mapping of application and user QoS preferences into evolving service profiles. There is a clear lack of standardization to give guidelines that put DiffServ into practice. In order to an NSP to configure and provision its network, it needs general, standardized and simple rules for their configuration.

Another challenge in the Internet is to provide end-to-end QoS guarantees. If the DiffServ architecture and mechanisms are fully deployed in the Internet, packets in the most case traverse multiple DiffServ domains to reach their ultimate destination. Since there are already a number of proposals for intradomain end-to-end QoS mechanisms, the question of providing interdomain end-to-end QoS is an interesting future research direction.

Typically QoS assurances are provided over a single DS domain, but end-to-end QoS guarantees are ignored. QoS guarantees need to be constant along the path between source and ultimate destination to offer the end-to-end QoS. This can only be ensured using a chain of the SLAs between all adjacent DS domains on the entire path from source to destination. In a wellprovisioned network, an NSP could contract a profile from its downstream NSP equal to or greater than the sum of all the upstream profiles it has contracted out.

However, the deployment of end-to-end service guarantee requires more than technical decisions. It is necessary to also deal with issues such as standardized SLA definitions and forms, QoS monitoring, peering SLAs and service negotiation. The user obviously wants to make comparison between the service offerings available in different ISPs. The comparison will be easier if there are standardized SLAs. Another important aspect with respect to SLAs is the customers' expectations and perceptions of QoS. SLAs must be also verifiable, otherwise it is just marketing hype. Performance monitoring will be easier if there are standard SLAs that provide a basis for collecting statistics and conducting benchmarking analysis. The QoS monitoring is a large subject and has many perspectives. An end-user is typically interested in the performance of specific applications, whereas an ISP needs to have a view of the state of the network. In particular, an ISP needs to monitor the usage of their network in order to ensure optimal and predictable performance of the network services. Interconnection of domains is also open issue that needs to be resolved before a reasonable level of service guarantees can be offered to the customers. When providing QoS guarantees between ISPs, the hardest problem is policy, not technology. This is reflected in the importance of peering relationships (defined in peering SLAs) between ISPs. To date SLAs are established mainly statically. However, there is an increasing demand to negotiate and establish the SLAs dynamically.

Service providers have responded to QoS challenges in a number of ways. Service providers have established bilateral peering SLAs that can support QoS

commitments. While most ISPs offer some kind of services that can be extended from their domain to another DS domain, an end-to-end QoS is not realised. Vendors have developed test equipments and third parties have appeared in marketplace to make it easier for customers to diagnose IP performance problems.

While most of the customers are satisfied that their QoS requirements are met, some customers demand also that their data is encrypted while traversing the ISP's network. Security can be seen as a part of guaranteeing network services. These security aspects are described in a network security SLA, which determines the point in the network, where the data or applications of the customer are encrypted, what encryption keys are used, and which applications or data require protection. Security aspects, even if extremely important for the deployment of QoS are not considered in this thesis due to its high complexity and extent.

## **1.2 Objectives and the outline of thesis**

In order to be effective, QoS guarantees need to be constant along the entire path between source and destination, i.e. end-to-end, across several domains. There is a clear need to understand the end-to-end performance in a DiffServ network. This thesis addresses the issue of providing QoS in the Internet. It is important that every domain understand service requirements in the same way, so that a unique end-to-end behavior may be achieved, no matter the sort of different QoS technologies domains are using to implement it. In this thesis it is put in evidence that QoS has several aspects, not all related to the IP level packet treatment. Specific areas are outlined, where there is a need to concentrate research and engineering efforts in order to improve today's user perception of QoS.

This thesis addresses the shortcomings of existing QoS and suggests some practical guidelines for bringing QoS by means of DiffServ and SLAs into intradomain and interdomain networks. This thesis introduces techniques and efforts that exist to reach the QoS.

The rest of this thesis is organized as follows. In Chapter 2 the DiffServ architecture and the basic components needed in the architecture are introduced. The focus of the chapter 2 is to present briefly the Differentiated Services architecture and building blocks behind it. First, the general DiffServ architecture model and the terminology related to a DS domain are explained. This is followed by the introduction of the most important traffic classification and conditioning functions, which mainly are performed at the boundary routers of an ISPs' network. In addition, some of the available DiffServ schemes are introduced in Chapter 2.

In Chapter 3 an overview of the SLA concept is given. The prevalent types of SLAs are introduced. However, the main focus in this thesis is the network level SLA. The content of SLA is very important to ensure the operation of the SLA in point of view of customers as well as in point of view of service providers. The issues related to the content of the SLA are highlighted. Technical side of the SLA, such as Service Level Specifications (SLS) is discussed. Dynamic SLAs that are made by software agents without the presence of human have been under a lot of research. Static SLAs and dynamic SLAs are compared and the benefits of dynamic SLAs are highlighted.

Chapter 4 describes the main ideas related to intra- and inter-domain QoS. The technical approaches are also introduced. DiffServ is a mechanism that can be connected with other techniques like MPLS and IntServ, whose advantages and disadvantages are discussed. A vision is given for intradomain QoS to be implemented using the standardized tools. Applications are mapped to existing

DiffServ traffic classes and guidelines to bring QoS to intradomain network are given. Also the means to build the interdomain QoS are highlighted and problems and advantages are introduced.

Chapter 5 discusses the DiffServ from the customer point of view. User-metrics are introduced. Also the possibilities of customer to provide QoS in local area network (LAN) are discussed. The main issues of fairness in DiffServ network are discussed.

The work closes with conclusions in Chapter 6.

## 2. Differentiated Services

The focus of this chapter is to introduce briefly the Differentiated Services (DiffServ, DS) architecture and building blocks behind it. First, the general DiffServ architecture model and the terminology related to a DS domain are explained. This is followed by the presentation of the most important traffic classification and conditioning functions, which are performed mainly at the boundary routers of ISP networks. The IETF has standardized some Per-Hop Behavior (PHB) groups<sup>1</sup>, such as Expedited Forwarding (EF) PHB group and Assured Forwarding (AF) PHB group. The former is designed for real time traffic, which requires a hard guarantee on the delay and jitter while the latter is for real- and non-real time traffic that has no such hard quantitative guarantees. Standardized PHB groups are discussed in Section 2.5.

### 2.1 An overview of Differentiated Services model

The main aim of developing the DiffServ model was to provide QoS to the network users. The DiffServ model is based on the assumption that Internet is a set of independent networks that are administered by ISPs. A single network can be considered as a homogenous region with the administrative control, technology and bandwidth [4]. The DiffServ architecture consists of many elements, such as edge elements and network core elements. Because these elements are logically specified in the DiffServ standard [2], it gives ISPs the freedom to create a wide set of services.

The essential components for the DiffServ architecture are traffic classification and conditioning, queue management and scheduling techniques in core routers. The DiffServ architecture is based on a simple model where traffic passing through the DiffServ network is metered, classified in different service classes and marked at the boundaries of the network according to an SLA.

After the packets have been classified at the boundary of the network, they are forwarded through the core network nodes according to the PHB associated with the Differentiated Services CodePoint (DSCP) field. PHB is the forwarding treatment that the packets experience in the network node when traversing the core DS network, i.e. PHBs describe the treatment the packets experience in a single network node. The core network is usually a homogenous area of the domain and its main duty is to forward the packets according the rules given. The packets are expected to experience the similar treatment as they cross the domain through the core network nodes. The operation of a core network node is mainly based on the information in the DSCP field. Core network nodes implement scheduling, buffering and forwarding functions.

Aggregation and packet marking makes the core network simple. Packets are treated based on the DSCP value. Packets with the same treatment (DSCP value) and the same direction<sup>2</sup> form a behavior aggregate (BA). Using BAs in the core of the DS domain does not need so much resources and intelligence in the core network than

---

<sup>1</sup> A PHB group usually consists of a set of two or more PHBs.

<sup>2</sup> The same direction means in this context from source to destination, exactly from a node to another node.

flow-based mechanisms. Moreover, it is more scalable and effective than for example IntServ architecture and RSVP- (Resource ReSerVation Protocol) based solutions. No signaling is needed and no reservations are made for these BAs. The essential difference in the DiffServ model compared to other QoS models is that aggregates are forwarded rather than flows.

## **2.2 Terminology related to a Differentiated Services domain**

According to the definition given in [2], a DS domain consists of one or more networks administrated by the same administration, for example organization's intranet or ISP. Usually only one operator administers the DS domain. In addition, a DS domain has common PHB definitions and service policy. Service policy means the set of rules that administer, manage and control access to network resources. DS domains together compose the DS region. A DS region is a continuous chain of DS domains, which can offer DiffServ along the path from source to destination.

Basic elements of a DiffServ network are DS nodes. The most complex tasks in a DS network, such as traffic classification and conditioning are pushed to the boundary nodes. In DiffServ terminology, a boundary node is determined as the node that is located in the edge of the DS domain. A boundary node connects the DS domain to other DS domains or a non-DS-capable domain [2]. A boundary node can be either a DS ingress node or a DS egress node depending on the direction of a traffic stream. The incoming traffic enters the DS domain through the DS ingress node and leaves the DS domain through the DS egress node. In the core of the DS domain are the interior nodes. These nodes are kept to be as simple as possibly and there is usually no traffic handling. Packets are classified to aggregates using the DCSP value. The main purpose of an interior node is to forward the PHB aggregates according the rules given with every PHB. The DiffServ framework is shown in Figure 2.1. The figure presents two DS domains composing a DS region.

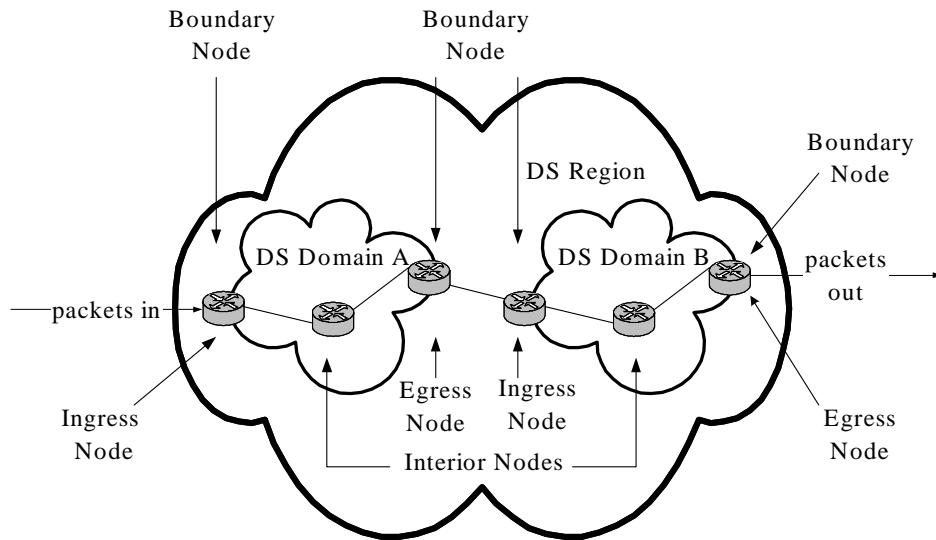


Figure 2.1. The DiffServ framework

### 2.3 Per-Hop Behaviors (PHBs)

The notion of Per-Hop Behaviors (PHBs) has been defined in the DiffServ standard [2]. PHBs define how traffic belonging to a particular DS behavior aggregate is treated at a DS network node.

After the packets have been classified and possibly conditioned according to an SLA made between the customer and its ISP, they are assigned to a particular DS behavior aggregate. This is done by marking the DS field of the headers of IP packets with the corresponding DSCP values. For example packets carrying a VoIP application can be marked with different value than packets carrying an email application.

PHBs are implemented in network nodes by means of buffer management and packet scheduling mechanisms. The core network nodes use active queue management (AQM) techniques e.g. to provide different drop priorities among the service class. There are many AQM candidates which can be used to this purpose, such as a simple RED (Random Early Detection) mechanism [5], RED with In-and-Out called RIO mechanism [6] and WRED<sup>3</sup> (Weighted Random Early Detection) mechanism.

DiffServ has adopted the 6-bit field of IP header to convey the treatment the packets should experience in a DS domain. Every IPv4 packet carries a Type of Service (TOS) octet. TOS octet is not widely used and therefore it has been redefined to support DiffServ. IPv6 packets carry Traffic Class (TC) octet, which has similar meaning than IPv4 TOS octet. Both the TOS and TC octets are renamed as a DS field. The DS field consists of six DSCP and two Currently Unused (CU) bits. CU bits are not used in DiffServ packet marking. CU-bits have been allocated to the use of explicit congestion notification (ECN<sup>4</sup>) [8]. The location of the DSCP field bits, which is equal for both IPv4 and IPv6 protocols, is shown in Figure 2.2. [7]

<sup>3</sup> WRED is Cisco's implementation of RED. It combines the RED algorithm with IP Precedence to provide preferential traffic handling for higher priority packets.

<sup>4</sup> ECN is used as the mechanism for signalling congestion.

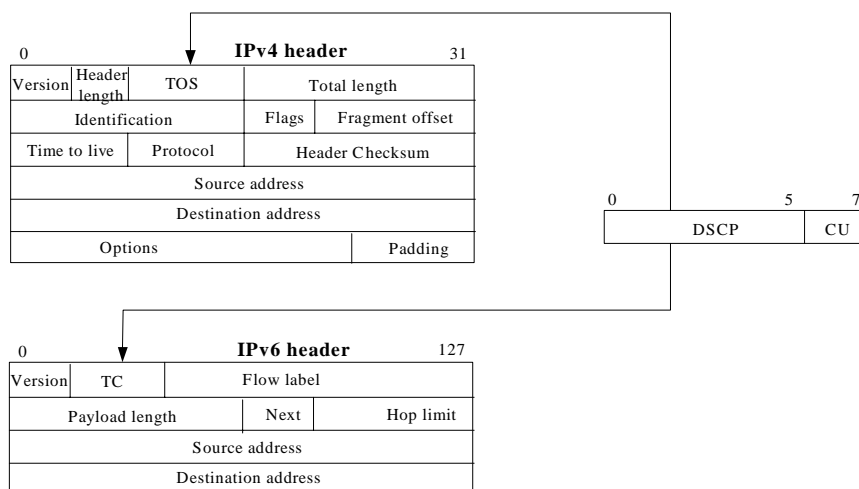


Figure 2.2. Location and bit allocation of DS field in IPv4 and IPv6 headers

The DiffServ Work Group (WG) has standardized a small number of PHBs. Because, the DSCP is six bits wide, it is possible to mark 64 different DSCP values to select the appropriate forwarding behavior. Network experts argue that all the values will hardly ever be needed [4]. The common consensus is to use fewer classes in order to get more granularity within each class. These 64 different DSCP values are divided into three different pools, based on their assignment policy, to help the management of codepoints shown in Table 2.1. Half of these 64 values have been reserved for standard action. Two pools of 16 values have been reserved for experimental and local use (EXP/LU). The difference with these pools is that pool three is also reserved for the use of standard actions. The values of pool three are used as standard actions in the case the values in pool one are already all in use.

Table 2.1. The use of the DSCP values [2]

Pool	DSCP	Assignment Policy
1	xxxxx0	Standard Actions
2	xxxx11	EXP/LU
3	xxxx01	EXP/LU

## 2.4 Traffic classification and conditioning

To support service differentiation for individual or aggregated flows, the DiffServ architecture consists of traffic classification and conditioning functions. Traffic conditioners may contain meters, markers, droppers and shapers. The elements of classification and traffic conditioning are depicted in Figure 2.3.

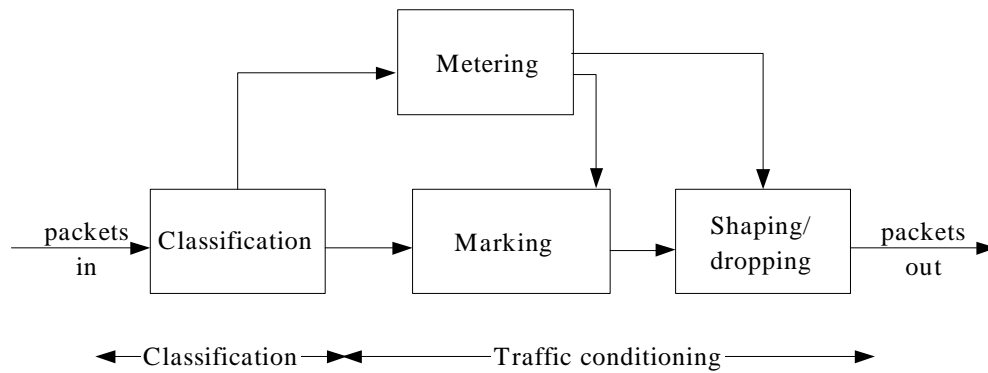


Figure 2.3. Logical schematic of a traffic conditioner [2]

Traffic classification and conditioning actions are mainly done at the boundary nodes, but these functions may also be implemented at nodes within the interior of a DS domain.

Packets are classified as they enter to the boundary node. Classification is a process to select packets according to the specified rules. There are two types of classifiers: Behavior Aggregate (BA) and Multi-Field (MF) classifiers. BA classifiers select the packets based only on the DSCP value, whereas MF classifiers select packets based on one or more fields of IP, User Datagram Protocol (UDP) and Transmission Control Protocol (TCP) packet headers. It is possible to use so-called 5-tuples (source address, destination address, source port, destination port and protocol type) as the means to identify an individual traffic flow. MF classifiers can be used to differentiate different application flows from each other.

Boundary nodes, i.e. ingress nodes perform traffic conditioning using metering, shaping and policing functions in order to ensure that the traffic entering the DS domain conforms to rules specified in the Traffic Conditioning Agreement (TCA). TCA determines the actions to happen in classifiers and traffic conditioners. The TCA is usually derived from the SLA.

The main function in metering is to determine whether packets are in-profile<sup>5</sup> or out-of-profile. It sorts the classified traffic to different importance levels. Traffic meter measures the classified traffic and compares it to the traffic profile. There are many ways to measure the incoming traffic flows. Token bucket<sup>6</sup> and exponential weighted moving average (EWMA<sup>7</sup>) based meters are the most commonly used meters.

Marker sets the appropriate DSCP value to a packet. It takes into account the state of the meter. If the packet is marked as out-of-profile, it will receive a higher drop precedence value within the same service class. In the case that packet is marked as in-profile, it will receive a low drop precedence and will be the last one to be dropped within the service class in the case of congestion. Marker adds the DSCP-marked packet to the right aggregate. Packets might arrive to the marker as pre-marked, if host or other boundary node has marked the packets. The marker may change the DSCP value, if needed, according the state of the meter. This operation is called remarking. Several markers are proposed in literature, such as a srTCM (Single Rate Three Color Marker) [9], a trTCM (Two Rate Three Color Marker) [10]

<sup>5</sup> Profile means in this context metrics that have been stated in SLA.

<sup>6</sup> Token bucket is described e.g. [www.ietf.org/rfc/rfc3290](http://www.ietf.org/rfc/rfc3290).

<sup>7</sup> EWMA is described e.g. [www.ietf.org/rfc/rfc3290](http://www.ietf.org/rfc/rfc3290).

and a Fair Marker [11]. These meters/markers are discussed in more detail in Section 2.5.4.

The function of shaper is to delay some or all of the packets to bring the stream into compliance with traffic profile. If there is no room to delay packets in buffer, they may be discarded. Droppers police the stream by dropping some or all of the packets. The purpose of policing is to ensure that traffic does not exceed certain bounds.

## 2.5 Standardized Per-Hop Behaviors

As already mentioned in section 1.1, current DiffServ architecture allows the definition of many different PHBs that can provide different QoS levels. Currently the main standardized PHB groups are Expedited Forwarding (EF) service and Assured Forwarding (AF) service. In addition, it is assumed that all DS-compatible nodes support Default PHB.

### 2.5.1 Default PHB

Default PHB, i.e. best effort (BE) PHB has been introduced in [7]. The Default PHB is a basic PHB, whose forwarding characteristics meet the requirements of today's best effort traffic. This traffic is forwarded as soon as possible, if there is nothing else more important to send. These are the minimum requirements, which are given to the Default PHB. For the ISP, it is often practical to give some minimum bandwidth and buffer space for the Default PHB in order to prevent the starving of the best effort traffic.

The recommended codepoint value for the Default PHB is '000000'. Every DS-compatible node should have this PHB available. Usually this PHB represents the lowest priority that the traffic may receive in QoS network.

### 2.5.2 Class Selector PHB

Class Selector PHB (CS PHB) has been introduced in [7]. The main purpose of the CS PHB is to offer backward compatibility with today's IP Precedence field of IPv4 TOS octet. When a DSCP node communicates with an IP Precedence node, only the first three bits in DSCP field are used to prioritize the packets. TOS octet and precedence bits are shown in Figure 2.4. As mentioned earlier, to date the TOS octet is not widely used. Thus the backward compatibility has not much significance.

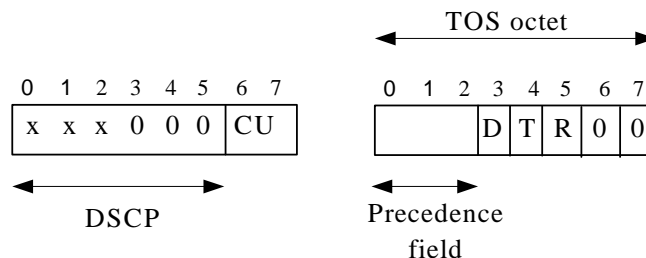


Figure 2.4. CS PHB backward compatibility

The IP Precedence field uses the first three bits of the TOS octet (bits 0-2). The CS PHB uses the first three bits of DS field values ('xxx000') having numerical values from 0 to 7. The CS PHB with the higher numerical DSCP value gets higher priority

than one with the lower numerical value. For example the packets with the DSCP value of '111000' will get priority over the packets with the DSCP value of '101000'. The CS PHB can be implemented through many different queuing and scheduling algorithms such as Priority Queuing (PQ), Weighted Fair Queuing (WFQ), Class Based Queuing (CBQ) and Weighted Round Robin (WRR).

The CS PHB codepoint values are taken from DSCP pool one (Table 2.1). It should be noted that Default PHB ('000000') is actually a CS codepoint. The CS PHB is a considerable class to be used in communications between the NSPs or in the DiffServ network of NSPs to classify the control traffic of the NSP. The intradomain QoS is built using the appropriate combination of the EF PHB, the AF PHB Group and the default PHB. The above-mentioned PHB classes have the specific characteristics for a DS domain and are natural to be used inside the DS domain. The interconnection of ISPs should be done using a standardized way and the CS PHB is a good alternative to implement connections like peering. The CS PHB is a valuable PHB to classify the control traffic of the ISP to the class of own to get higher priority over the other traffic classes in the case of congestion. The use of CS PHB would be essential to be standardized to help the way to communicate between the ISP through DiffServ means.

### **2.5.3 Expedited Forwarding PHB**

Expedited Forwarding PHB (EF PHB) has been introduced in [12] and specified in more details in [13] and [14]. The idea behind the EF service is to provide low loss, delay and jitter guarantee in order to meet the strict requirements of real time services, such as VoIP and videoconferencing.

The EF PHB is a rate controlled PHB. Packets marked to the EF PHB are guaranteed to have service at least at configured rate. The rate of packets is strictly controlled both in the boundary nodes and in the interior nodes. Boundary nodes control arriving packets and shape the rate to be the configured rate and prevent the arriving packets to enter faster. Interior nodes forward the packets as fast as possible. The means to reach the goal with low loss, jitter and delay is to keep queues short or, if possible, almost empty. When the queues are almost empty there will also be a low packet loss. The EF PHB has only one importance level and therefore no marking is needed during traffic conditioning. Recommended DSCP value for the EF PHB is '101110'. The EF service can be considered some kind virtual leased line service, because EF provides the dedicated link of fixed bandwidth between the two edge nodes. Figure 2.5 shows a single EF service queue and its position among other queues.

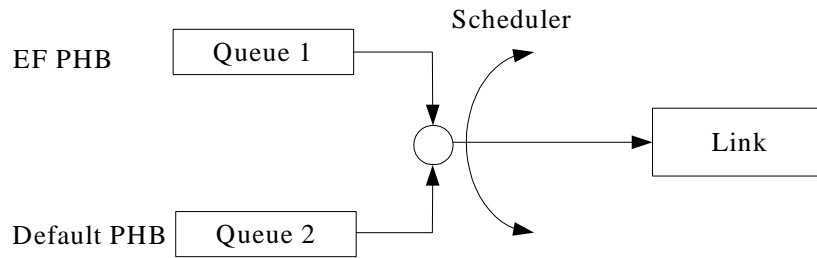


Figure 2.5. EF PHB with default PHB

The EF specification provides only the formal definitions for an EF service, it does not tell the details of how to build such a service. An EF service could be implemented in the network node using different mechanisms. One way to implement the EF service is to use PQ mechanism. Another example of an implementation of EF is a WRR scheduler. In addition, WFQ-like schedulers can be considered as implementing an EF-like service.

The EF PHB needs very careful configuration of traffic conditioning, queuing and scheduling mechanisms. The ingress node has to use very strict traffic conditioning to ensure the suitable traffic handling in the core network node. In the core network node there should be no need to drop packets, if the traffic handling has been successfully implemented in the ingress node.

#### 2.5.4 Assured Forwarding PHB

Assured Forwarding PHB (AF PHB) group is introduced in [15]. Many variations of AF PHB have published. However, the main interest in this section is in the AF service proposed in [15]. The AF PHB has attained a lot of interest among researchers. The influence of increasing or decreasing the number of classes or drop precedence levels has been an issue of research as well as co-operation with other classes. The implementation of AF class has many possible alternatives and there is lot of research of this issue.

The AF PHB offers an assured forwarding of packets. The AF service [15] defined by IETF provides the delivery of IP packets in four independent traffic classes (AF classes), each with three levels of drop precedence (low, medium, high) as shown in Table 2.2. In the AF service, each packet has a codepoint encoded in the DS field, which identifies the AF PHB. In all, there are twelve DSCPs reserved for AF PHB group. Figure 2.6 shows AF service queues and their position among other queues.

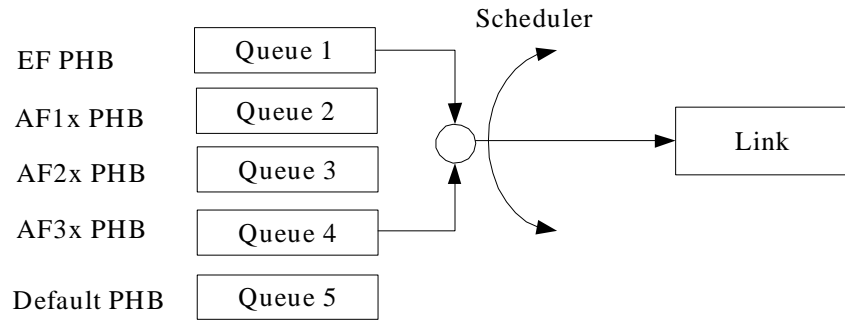


Figure 2.6. AF PHBs with other PHBs

Each AF class has been allocated a certain buffer space and bandwidth in a DS node. Arriving packets can be pre-marked by a customer or by the ISP to have a certain drop precedence value. In the case of congestion within the AF class, the drop precedence of each packet determines the relative importance of the packets. Packets with high drop precedence are dropped first during the congestion and the last ones to be dropped are packets as marked low drop precedence. If there is free capacity in an ingress node, the capacity can be shared among AF classes. Packets with low drop precedence level are served first. AF class may also get extra capacity from other PHBs.

DS nodes may not aggregate different AF classes together. Every AF class should be served and forwarded independently. A DS node must avoid reordering of the microflows that belong to the same AF service class. One way to avoid reordering is to use only one queue for each AF service class. If the congestion is rare in the network, it is possible to use only two different drop precedence levels. When using two different drop precedence levels, the two lowest levels are combined.

Table 2.2 presents DSCP values recommended for the AF PHB group. First three bits of DSCP field describe the class and next two bits indicate the drop precedence. The last bit (zero) describes that this DSCP belongs to pool one and is for standard actions. The drop precedence levels of different classes are different from each other while these classes are independent.

Table 2.2. Recommended DSCP values for the AF PHBs

	Class 1	Class 2	Class 3	Class 4
Low	001010	010010	011010	100010
Medium	001100	010100	011100	100100
High	001110	010110	011110	100110

A crucial part of the AF PHB is marking. Many different marking mechanisms have been introduced till today and some of them have also been standardized. Three commonly used standardized marking mechanisms are a Time Sliding Window Three Colour Marker (TSWTCM), a Single Rate Three Color Marker (srTCM) [9] and a Two Rate Three Color Marker (trTCM) [10]. TCM assumes the packets arriving to the ingress node to be either color-blind or color-aware mode. Color-aware mode means that packets have pre-marked green, red or yellow before they reach the ingress node. Color-blind mode supposes that no marking has been done beforehand.

The TSWTCM meters a traffic stream and marks packets with different colors (green, yellow or red). This action is based on the measured throughput relative to two specified rates: Committed Target Rate (CTR) and Peak Target Rate (PTR) [16].

The srTCM is configured by means of Committed Information Rate (CIR), Committed Bucket Size (CBS) and Excess Burst Size (EBS). The srTCM is useful, if only burst size matters.

The trTCM mechanism meters the incoming traffic of the user and marks its packets based on two rates, (CIR) and Peak Information Rate (PIR), and their associated burst sizes, CBS and PBS (Peak Bucket Size) respectively. CBS is used as the green token bucket size, whereas PBS is used the yellow token bucket size. The trTCM is useful, if peak rate needs to be enforced. The trTCM with color-blind mode is shown in Figure 2.6.

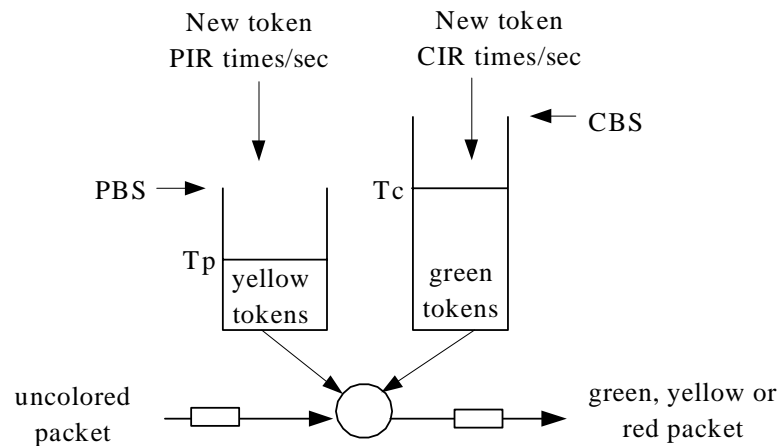


Figure 2.6. Color-blind mode of trTCM [10]

Two token buckets are used to meter the incoming packets. When an aggregate's measured traffic is within its contracted average sending rate (CIR), its packets are marked as green. When traffic exceeds its CIR, but falls below its maximum contracted sending rate (PIR), packets are marked as yellow, i.e., there are tokens available ( $T_p$ ) in token bucket for PIR, but not enough tokens for CIR. A packet is marked as red if it exceeds the PIR. PIR rate should always be equal or greater than CIR [10].

There has been much discussion (such as in [17]) about the usefulness of three different drop precedence levels within a single service class. It is observed that a three color marker does not offer significantly better performance than marker using only two drop precedence levels when dealing with TCP traffic. The reason for this is that TCP does not distinguish between packet losses of different drop precedence.

### 3. Service Level Agreement

A Service Level Agreement (SLA) is determined as a bilateral contract between an service provider and a customer or between two service providers. An SLA specifies the expected service that a customer will receive in the network. On the other hand, an SLA is a basis for a service provider to account for the usage of network resource.

The Internet consists of customers, ISPs and NSPs. Customers may be either residential or business customers. NSPs own and maintain high speed and high capacity networks and provide global access and interconnectivity. NSPs sell wholesale Internet connectivity to ISPs. ISPs offer network services to their customers. However, NSPs may also sell retail Internet connectivity to the customers [19]. Service provider is used as a general term for ISP and NSP.

In this thesis an ISP means a service provider that sells the network connectivity services to customers. An NSP means the global service provider that sells capacity to ISPs or other NSPs. The core of the Internet consists of NSPs and the ISPs are in the edge of the Internet offering the network services to customers. Figure 3.1 clarifies the terminology.

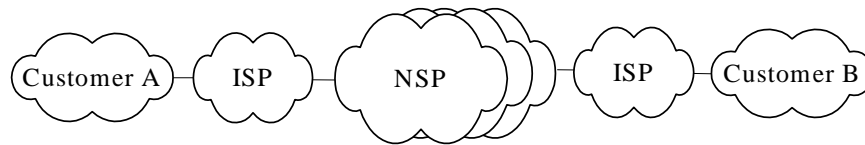


Figure 3.1. Terminology related to service providers

NSPs interconnect with each other through Internet Exchange Points (IXPs) or using direct connections. Most often ISPs, as they are considered in this thesis, are too small to fulfill the requirements of IXPs and ISPs find the interconnection through NSPs effective.

This chapter explains some general issues related to the SLA concept. The SLA is mostly discussed from the network level point-of view. The content of SLA is considered and the role of SLA and its technical part Service Level Specification (SLS) is introduced in DiffServ context.

An SLA may be either static or dynamic. A static SLA is negotiated between human representatives, whereas, a dynamic SLA is negotiated between automated agents without the presence of human. SLAs are today mostly static. Both static and dynamic SLAs are introduced.

#### 3.1 Prevalent types of SLA

SLAs can be generally categorized into three prevalent types according the service the provider will offer: an application SLA, a network level SLA and a service provider SLA. Figure 3.2 shows the SLAs that service providers offer.

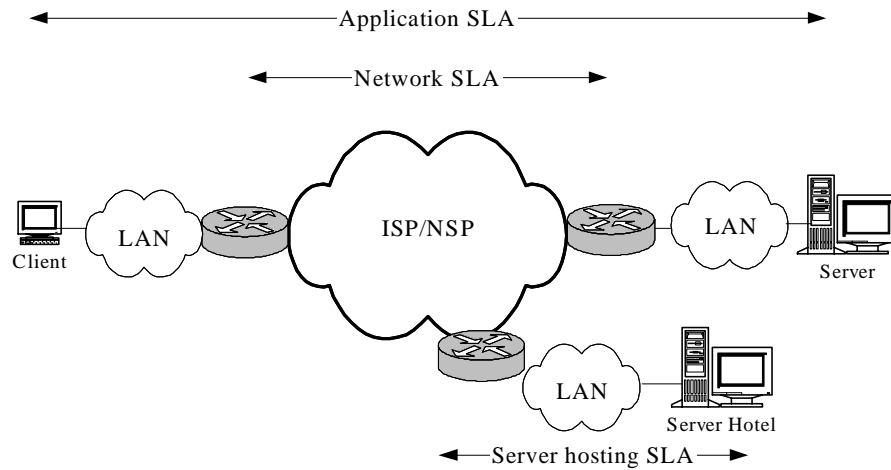


Figure 3.2. Prevalent types of SLA

Application SLAs describes the performance of specific applications in terms of response time, time the server is expected to be available or time of transactions. As seen in Figure 3.2 an application SLA is under the influence of a network SLA. The client communicates with the server through the network of service providers. The performance of service provider network effects on the quality of application the end-user experiences.

The server hosting SLA means in this context a service where the provider hosts the server (e.g. Web server) of a customer on behalf of a customer. Provider usually offers a certain access to the network of ISP. The SLA includes the bandwidth reserved for the server of the customer. In normal case a provider usually controls only over the server side access to the network not the client side access. Performance is determined in terms of access capacity and server capacity. The server hosting SLA is also under influence of the network SLA.

A network level SLA deals with the part between the access points of network. Network SLA describes the performance of network between the defined points of the network. This thesis focuses on the network level SLA.

### 3.2 Network level SLA

A network SLA is the fundamental SLA and needs to be determined carefully to satisfy the expectations of the customers. Network performance can be depicted using approaches like a funnel SLA, a tunnel SLA and a cloud SLA shown in Figure 3.3.

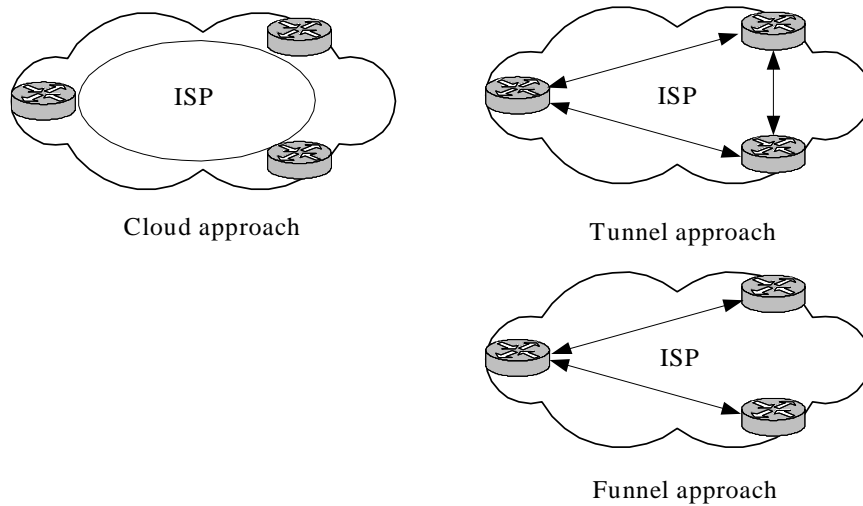


Figure 3.3. The areas of network SLA [20]

SLAs are determined in many networks of ISPs in terms of a cloud approach. The cloud approach determines the SLA parameters (e.g. average delay in the network) from any point to any point in the network of the ISP. It might be guaranteed e.g. the delay inside the network of the ISP to be below 50 msec. The value is usually monthly average. The funnel approach specifies the SLA from one point to any point of the network. It might state that the delay from a certain point to any other point in the network of ISP is below 50 msec. The tunnel approach could be used between the sites of the enterprise, while it specifies the SLA between two specified points.

Interconnection among customers, ISPs and NSPs is formed using appropriate network level SLA. Customers conclude an access SLA with ISP. NSPs and ISPs conclude a peering or a transit SLA with ISPs or with other NSPs. Peering SLA is done through Internet Exchange Points (IXP). Figure 3.4 highlights the position of terms mentioned above. The service provider in the Figure 3.4 could be either ISP or NSP.

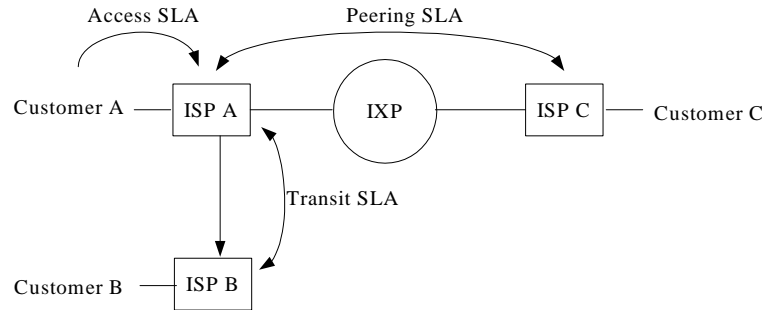


Figure 3.4. Network SLAs between service providers

An access SLA is done between the customer and the ISP. It describes the service in terms to satisfy the business and technical expectations of the customer and the ISP. The end-to-end QoS is under the influence of access SLAs, transit and/or peering SLAs. For instance, if the customer A in Figure 3.4 wants to have an end-to-end connection with the customer B, it will be under influence of access SLAs in both sites. Peering SLA between the ISP A and the ISP C will have a significant meaning to the QoS the customers will experience.

Peering is the most powerful way for NSPs to interconnect with each other. There are both public and private peering points available. The main idea in peering is that two NSPs of equal size exchange traffic destined to their customers. In general, there is no SLA present for interconnection today and no money will be charged except annual fees. However, some private IXPs are offering SLAs for traffic that is exchanged. Issues of peering and transit are discussed in more detail in Section 4.2.3.

For smaller ISPs transit SLA is a suitable way to enter the Internet. In the case of transit the downstream provider (like small ISP) pays to the upstream provider (like NSP) to reach the peers and customers of the upstream providers. The transit SLA may also be a mean to enlarge the interconnection of bigger ISPs that already have some peering SLAs.

### 3.3 The content of network level SLA

An SLA is a commercial formal agreement including both business and technical information. Today's SLAs are mainly established using a long-term period (a year to three years) and therefore are an important base on relationship between the customer and the ISP. It is the guarantee to a customer to get the service that has been agreed. Alike it is also a guarantee to a ISP to charge the customer according to the SLA. The ISP gives the resources to build the services for the customer, to plan and establish the service for a certain period of time according the SLA. The SLA is a common deal that satisfies the needs and expectations of the customer and the ISP.

The SLA includes at least the following information: service description, service pricing, responsibilities, reporting, contact lists, actions related to failures, security and accounting procedures [21] as shown in Table 3.1. The foregoing parts of the SLA are described in more details in this section.

Table 3.1. The content of SLA

<b>Content of SLA</b>
service description
pricing
responsibilities
reporting
failure procedure
contact information
security
accounting

The offered service needs to be carefully described. This includes all the QoS commitments including e.g. service reliability. It could contain an information such as sites to be connected to the service, total capacity of the access link, the QoS the applications are expected to experience and time the SLA to be valid.

The offered service including service pricing are the key components in bilateral negotiations. Service pricing is described and listed in the SLA and represents an important part of the commercial side of the SLA. SLAs may include also prices for future enlargement. SLAs may also have the list for prices to be used in the case the capacity or the QoS parameters will be changed.

The SLA describes the responsibilities for the ISP and for the customer. The customer responsibilities could be e.g. to give information to the ISP for certain type of changes in the network such as a new IP address space to be implemented, new applications to be installed or new sites to be connected. The responsibilities of the ISP could be to maintain the hardware and software of the customer equipment (CPE) such as an access router and to inform the customer of significant changes to be done to the CPE. One of the responsibilities of the ISP is also to inform the customer, if there will be a failure in the network of the ISP that affects to the traffic of the customer. Reporting is also a very important part of the responsibilities of the ISP.

The SLA determines how often reports are sent to the customer and what events reports should contain. Reports are a good tool for both the customer and the service provider. The traffic profile of customers such as traffic sent during a day and traffic peaks are easy to examine due to the reports. Reports are a good tool to help the customer to estimate future capacity needs and also to plan service upgrades. Reports give the view of traffic growth or maybe traffic fall. Reports are usually the main tool to verify whether the goal of the service is reached. Reports are usually delivered once a month and give the backward information of the measured traffic. Often reports give the average values and will not be accurate enough to visualize the situation during a particular moment between the different applications. Today there are many ISPs, who offer timely reporting. There might be only a minor delay of the information given. Customers may log in the report center and list the reports of their sites in real-time [22]. Reports help the experts of customers to clarify the problems in the network by offering a valuable tool to define the area of problems in the case of failure to deliver the traffic.

In the case of failure, there is the way to act for both the customer and the service provider, e.g. where the customer shall send his or her email or where to call. The SLA contains the time limits for the service provider to start the investigations and the time limit the service should be fixed. There might also be the time limit for the service provider to inform the customer about problems in network of ISP. The penalties for a service provider preventing to reach the service are added in the SLA.

A notable part of the SLA is the contact list for both the customer and the ISP. Contact list contains both the commercial and the technical persons. The persons and their contact information are defined. The ISP has a help desk for customers. Usually the telephone number, fax and email of the help desk of the ISP are placed to the SLA.

Security part identifies issues such as the level of encryption, the point in the network where data is encrypted and use of encryption keys. Also the applications are services that are protected as they traverse through the network of ISP are defined.

Accounting procedures depend on the capabilities of the service providers as well as the needs of the customers. Customers may have a need to subdivide the account among the departments.

The customers usually set requirements to the ISPs to measure QoS reliably, to provision the expected QoS and to optimize the resource usage. The efficient management gives the value to SLAs. SLAs need to be understandable for the customer and the service should be described using terminology that the customer is familiar with.

### 3.4 SLA and Differentiated Services

There are many elements that are associated with SLA like Traffic Conditioning Agreement (TCA), Service Level Specification (SLS) and Traffic Conditioning Specification (TCS). These elements and their relationship are clarified in [23] from the DiffServ point of view. The relationship between the elements is described in Figure 3.5.

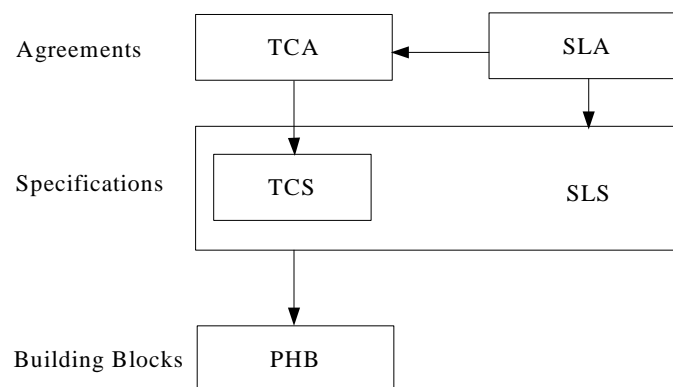


Figure 3.5. Terminology related to the SLA concept

The SLS is the technical part of the SLA. SLSs can be an accurate technical copy of issues stated in the SLA or it can be an interpretation of the SLA. The SLS defines the service the DS domain offers to a traffic stream. It includes parameters and their values.

The TCA describes generally classifier rules, traffic profiles and traffic conditioning rules, i.e. it describes DiffServ characteristics. These rules might have been specified in the SLA or they are relevant to service described or they can describe the service provisioning policy of the DS domain. The TCS is a part of the SLS and it is derived from the TCA. Together above described elements define PHBs, DSCPs and the treatments of the traffic for DSCPs.

These elements can be divided into three categories: agreements, specifications and building blocks. Agreements are the first to be done and they consist of the TCA and the SLA. The TCA is often derived from the SLA. Specifications are derived from agreements and they consist of the SLS and the TCS. The TCS is a part of the SLS. PHBs are the building blocks for the service and they are based on the SLS.

Some proposals have been made from SLS parameters and their semantics in the Tequila<sup>8</sup> [24] and the Aquila<sup>9</sup> [25] projects. These parameters offer a way to describe the areas of classification, traffic conditioning and the service guarantees. Parameters are service-related or network-related.

Service-related parameters are parameters that describe performance guarantee such as delay, loss, jitter and throughput. Reliability is described using maximum allowed time to repair or mean down time. An important parameter is also the one that describes the way to handle excess traffic. Parameter points out whether excess traffic should be shaped, dropped or remarked. The time the service will be available is determined through these parameters.

DiffServ-related or network-related parameters describe the way the network should operate. It is described whether BA or MF classifier is used and whether the connection is formed from an ingress node to other egress node or from one ingress node to any egress node. Traffic conditioning operation is described using parameters. Parameters describe the characteristics the packets should have to get the service guarantees that have been offered. In-profile and out-of-profile packets are identified. Parameters may describe e.g. token bucket rate, maximum transport unit (MTU) and the maximum allowed peak rate.

Four predefined types of SLS have been determined [25]. The focus for predefined SLS types is to offer a certain type of SLS to a certain type of traffic need. Predefined SLSs types are described for Constant Bit Rate (CBR)-type traffic, Variable Bit Rate (VBR)-type traffic, multimedia traffic and mission critical traffic. CBR-type SLS may be used in applications like VoIP and VLL-service. VBR-type SLS may be used in video- and teleconferencing whereas multimedia SLS may be used with streaming multimedia.

Standardized SLS parameters and semantics are essential in building the end-to-end QoS. The standard should include a large number of possible parameters and it should be general to be useful. The standardized SLS parameters are the base for negotiation between the ISP and the customer.

---

<sup>8</sup> Tequila: Traffic Engineering for Quality of Service in the Internet, at Large Scale. Its main objective is to study, specify, implement and validate service definition and traffic engineering tools for the Internet.

<sup>9</sup> Aquila: QoS Architecture for Adaptive Resource Control. It is aimed to enable dynamic end-to-end QoS provisioning in IP networks for QoS sensitive applications.

### 3.5 Static SLA

Static SLAs are the ones in use today. Static SLAs are done between organizations by people who represent the organization. Static SLAs are updated seldom, maybe once a year. There is always at first a negotiation between the ISP and the customer. Usually some technical operations are needed before new negotiated characteristics are in use. In worst case it might take from few days to weeks to get e.g. more bandwidth. Reports are a good tool to foresee these needs.

Static SLAs meet the case if customer needs no time-varying requirements and conditions. The network is presumed to act the way stated in the SLA. Applications are satisfied by certain amount of bandwidth, delay and loss characteristics. There is no need to negotiate parameters for a particular moment. The network of customer is very stable and traffic profile can be predicted. In the case of insufficient resources, new SLA has to be negotiated between the representatives of the ISP and the customer.

The SLA stands for the expectations the customer will have on the performance of the network of ISPs. Congestion situation in the network will effect on the opinion the customer will have from the performance of the network. For instance in the case of VoIP traffic the traffic handling is essential. What will happen, if there is already enough VoIP sessions and five new ones are trying to get capacity? Is the capacity shared among all the VoIP sessions and the performance lost for all of them, or are the new five sessions discarded? In the case of static SLA, there will always be end-users that are not satisfied during the time the amount of bandwidth will be underprovisioned because of congestion. To fix the situation, the customer and the ISP need to negotiate and some operations need to be done in the network e.g. more capacity may be determined to VoIP applications. However, this negotiation is time consuming and the end-user has to wait the improvement.

Static SLAs are today the way to build the contract between the customer and the ISP. However applications requiring all the more interactive actions will be satisfied using static SLAs no longer. Interaction will increase among applications and more demand will be on direct communication between the customer and the ISP without human interaction.

### 3.6 Dynamic SLA

Dynamic SLAs are established using automated machines or software agents without a human to take part to the negotiation. Dynamic SLAs offer a better utilization of network resources and faster time scale to the negotiating. When the SLA is static, it changes maybe once a month or a year. Dynamic SLAs make it possible to negotiate SLA e.g. per every session.

There is a need to have a dynamic SLA both in intra-domain and inter-domain networks. It is essential to have dynamic SLAs to build a functional end-to-end QoS. The tools that the static SLA offers for inter-domain QoS are incomplete.

Dynamic SLAs can increase network utilization through more efficient resource use. ISPs are able to minimize costs of network resources and maximize profits through efficient use of all resources. ISPs need not to prepare for the worst case e.g. over-provisioning the network. The need for accurate network resources is known at every moment in the network of ISPs through dynamic SLAs.

Customers can get benefits of using dynamic SLAs because they will always have the resources needed at that particular moment. Customers may use network resources based on momentary needs. Billing might be based on transmitted data (Megabyte) or used bandwidth (Mbps). The SLA might have thresholds for dynamic negotiating like determine the upper limit to be used for bandwidth. It will give the best benefit to an enterprise to use the required amount of network resources to make his/her business. Thresholds stated in the SLA could protect the network of the customer from overusing available network resources

Dynamic SLAs need some kind of a Policy Server to control network resources that are available. Interaction with the customer and the Policy Server needs a way to signal each other. Bandwidth Brokers(BBs) are proposed a way to handle network resources. BBs are discussed in more details in Section 4.2.4.

Dynamic SLAs will be a part of the future interoperation between customers and ISPs. A lot of research is done concerning dynamic negotiation of network resources. More research will be needed to bring effective and standardized model for use. Policy-based Networking (PBN) is the main research area and it is discussed more in Section 4.2.4.

## 4. Differentiated Services as a tool for operator

This chapter focuses on the issues related to techniques that use DiffServ to implement the QoS in the network. The chapter is subdivided into two sections that discuss the intradomain and interdomain QoS.

Intradomain QoS is discussed first, because the techniques introduced are mostly usable also in the interdomain area of the QoS. Different mechanisms can be connected with DiffServ like IntServ and MPLS. IntServ has many advantages, but it has lost a lot of interest, since it does not scale so well because of the flow-based mechanism. IntServ and DiffServ have been united in [26]. The advantages and disadvantages of combination of IntServ and DiffServ are presented in this chapter.

MPLS has got a lot of interest among ISPs and NSPs, because of the versatile traffic engineering characteristics. MPLS is also largely used inside the networks. Benefits of using MPLS with DiffServ are discussed later more.

It has come up that there would be also a need to use a traffic class that would serve the traffic worse than a BE traffic or a default class. The idea is to offer the unutilized bandwidth for use, after the other traffic classes are served. The applications to use this class of service would have to be tolerant of delay and packet loss. This Less than Best Effort (LBE) class of service is described and its possible use is highlighted.

In the end of intradomain QoS section the vision of DiffServ network is given. The focus is on the traffic classes that provide the QoS in the intradomain network and serve the need to classify the traffic by means of DiffServ. Applications are mapped to the service classes. Service classes are implemented in DiffServ using certain PHBs. Main issues are discussed and some examples are given to utilize the service classes and to map them to DiffServ PHBs.

Interdomain section introduces the problems as well as possibilities to offer the end-to-end QoS. The proposals of Per-Domain Behavior (PDB) are reviewed. PDBs describe the end-to-end QoS inside a single DS domain. In other words, PDB describes the edge-to-edge QoS. The concept of PDB is discussed as well as problems and needs to implement PDB into the DiffServ mechanism.

The main interconnection points in the Internet for NSPs are the IXPs. Traffic between the NSPs traverse most often through the IXP. The service the IXP will offer to the NSP will have a significant effect on the end-to-end QoS that the customer will experience. The need to support the QoS in the IXPs is considered more in subsections.

The resources have always the final thresholds in networks, though it might in most cases never be met. PBN is a concept that gives the tools for policing the network e.g. classifying the traffic, setting the policy rules, maintaining the available resources and policing the arriving traffic. PBN elements are presented and the use of BBs is discussed.

Finally the means to build the end-to-end QoS in interdomain networks is considered. Both problems and challenges are brought out.

In both, intradomain and interdomain cases, useful metrics are shown. Intradomain and interdomain metrics are enlarged separately, because the metrics and the purpose of metrics differ from each other in these cases.

## 4.1 Intradomain QoS

### 4.1.1 IntServ over DiffServ

IntServ over DiffServ [26] offers a model to build the end-to-end QoS using IntServ [27], DiffServ and RSVP [28, 29]. It is an approach, whose goal is to combine the benefits of DiffServ and IntServ models to support the delivery of the end-to-end QoS. IntServ offers a per-flow mechanisms, but faces significant problems to scale in large networks. DiffServ offers the scalability properties in large networks, but does not have reliable mechanism to manage the network resources. The advantages of IntServ and DiffServ are combined to meet the QoS using IntServ in small access networks and DiffServ in larger core networks. The framework is presented in Figure 4.1. DiffServ is used in the middle of the network and it may or may not to support IntServ. DiffServ region can be either statically or dynamically provisioned.

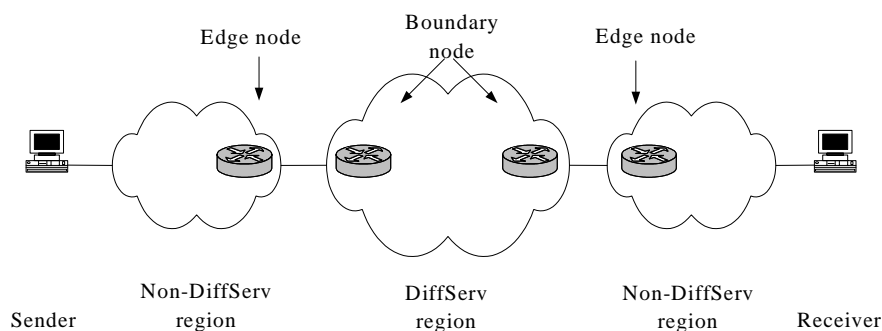


Figure 4.1. IntServ over DiffServ

In the case the DiffServ region is statically provisioned, it does not have to be aware of RSVP. It is assumed that RSVP messages are delivered transparently from the sending host to the receiving host. The boundary node is a pure DiffServ node and the admission control is done in the edge node based on the SLS negotiated between the DiffServ region and the non-DiffServ region. There is no signaling between the boundary nodes and the SLSs are static.

In the case the DiffServ region is provisioned dynamically, boundary nodes have to be aware of RSVP. Boundary routers should participate in RSVP signaling and do admission control in order to make reservations of resources for network traffic. RSVP might also be supported in core routers inside the DiffServ region. Yet, these core routers classify and schedule traffic based on aggregates not per-flows.

IntServ over DiffServ framework is a significant approach to fix the end-to-end QoS using existing models. However, both ways to implement DiffServ region in the middle of network, with or without RSVP signaling, have many disadvantages. The main drawback in statically provisioned DiffServ region is that edge nodes do not know the traffic load inside the DiffServ region. As the SLS is static, it can not effect on resource availability in the DiffServ region. If the DiffServ region is dynamically provisioned, signaling and provisioning might have overhead, which in turn decreases scalability of DiffServ [30]. RSVP aware boundary nodes will increase the complexity of the DiffServ region and this might adversely affect the performance of the network in terms of delay.

### 4.1.2 DiffServ over MPLS

Multiprotocol Label Switching (MPLS) [3] is a label switching technique, which has largely increased its popularity. MPLS offers benefits to ISPs by providing efficient means for traffic engineering (TE). TE requires specific routes to enable versatile routing, traffic load balancing and other means of optimizing network resources [31]. MPLS enforces traffic to specific routes thus minimizing to reduce the situations of network congestion and getting the network resources in efficient use. MPLS combined with DiffServ offers a fast forwarding technique as well as a good service differentiation.

The main elements of the MPLS network are Label Switching Routers (LSRs) and Label Switching Paths (LSPs) as seen in Figure 4.2. Traffic enters the MPLS network at Ingress Edge LSR. Ingress Edge LSR adds an MPLS header to the IP packet and assigns a label. Ingress Edge LSR assigns an incoming packet to a particular Forwarding Equivalence Class (FEC) according to the label.

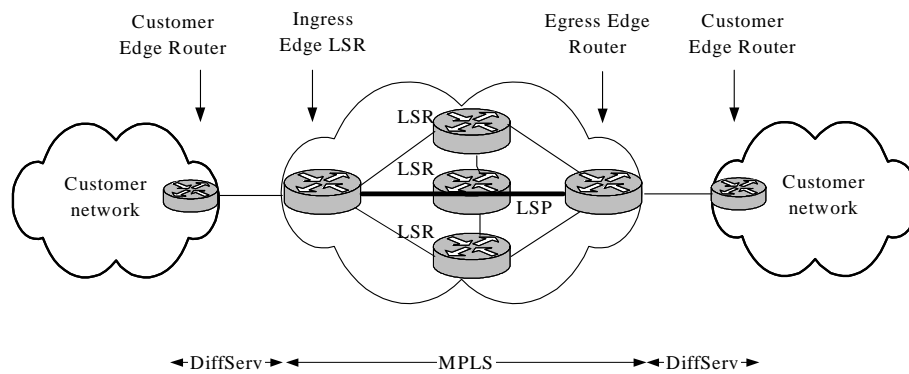


Figure 4.2. DiffServ over MPLS

MPLS is based on the labels. The ingress LSR marks all the incoming packets with the label. Packets are forwarded in the MPLS network based only on that label. The forwarding is fast, because only one value needs to be checked. This improves the delay characteristics of the MPLS network. The label is a 32-bit wide and it is called shim header. MPLS locates the shim header between the layer 2 and layer 3 headers. The shim header is shown in Figure 4.3. There are four different fields in the shim header: the label (20 bits), the EXP field (3 bits), the S field (stack field 1 bit) and the TTL (Time-To-Live -field (8 bits). The label field carries the actual label value. The EXP field contains an information of packet differentiation and can affect on the queuing and discard algorithms. The EXP field can be used to map the DiffServ DSCP value to the MPLS. The S field supports label stacking. The TTL field offers time-to-live operation like in the IP header.



Figure 4.3. MPLS shim header

FEC is encoded as a value of a label. FEC is a set of packets that usually traverses the same path in an MPLS domain. Packets are then forwarded along the LSP. The LSP is a specific path from the ingress LSR to the egress LSR. The packet

forwarding is based on a label, without analysis of the IP header. Every LSR in the MPLS core network removes the old label and replaces it with a new one. The egress Edge LSR removes the label and forwards the packet to its destination. LSPs build explicit circuits inside the MPLS network and offers an effective TE for ISP to operate the network.

DiffServ over MPLS approach allows the ISP to select the way to map BAs to LSPs to match the DiffServ, traffic engineering and protection objectives within their network [32]. Two different ways to implement MPLS with DiffServ is to use Exp-Inferred-PSC<sup>10</sup> LSPs (E-LSP) or Label-Only-Inferred-PSC LSPs (L-LSP). Because MPLS is based on label switching, it does not interpret the DS field in the IP header. E-LSP and L-LSP fields are used to carry the PHB information. [33]

The E-LSP uses the EXP field of the MPLS shim header to convey the PHB information. The LSR uses the EXP field to determine which PHB to apply to the packet. The EXP field contains the information of both the PHB scheduling Class (PSC) and the drop precedence. Mapping from the EXP field to the PSC can be pre-configured or signaled at a label set-up. Figure 4.4 presents mapping between the DSCP and the EXP field. The E-LSP field can support up to eight PHBs of a given FEC per LSP.

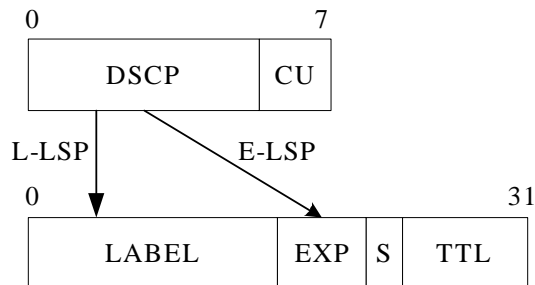


Figure 4.4. DSCP mapping to the MPLS shim header EXP-field

The L-LSP field is used when more than eight PHBs are needed to be supported or different LSPs are used with a FEC. In this case the label conveys the PHB information instead of EXP field and the PSC is signaled at the time of the label establishment. Only one PHB can be mapped to the LSP. In case of Asynchronous Transfer Mode (ATM) and Frame Relay, that do not use shim header, the label is carried in Virtual Path Identifier (VPI) or Virtual Channel Identifier (VCI) field of the ATM header and in Data Link Connection Identifier (DLCI) field of the Frame Relay header. An AF Class makes an exception here. A drop precedence has to be mapped somewhere. If an EXP-field is available, it can be mapped there. In the case of Frame Relay and ATM, the drop precedence is conveyed inside the link layer. ATM conveys the drop precedence in the Cell Loss Priority (CLP) bit and Frame Relay in Discard Eligibility (DE) bit.

Bandwidth reservations can be established using both the E-LSP and the L-LSP fields. Bandwidth reservation with the E-LSP field is associated with the whole LSP and with all the possible PSCs in the LSP but it can not be associated with the specific queue in the case of E-LSP. In the case of L-LSP, bandwidth reservations

<sup>10</sup> PSC is a set of one or more PHBs that are applied to the BAs belonging to a given ordered aggregate. Ordered aggregate contains the set of BAs that have same ordering constraint. EF is a PSC containing only one PHB, instead AF1x is a PSC containing AF11, AF12 and AF13 PHBs.

can be made to a specific PSC. This means that e.g. an EF Class carrying VoIP traffic can be determined to have the certain bandwidth guarantees. [34]

DiffServ over MPLS framework offers a significant approach to combine the traffic engineered paths of MPLS with the capability of differentiate traffic using DiffServ. It alleviates the steps to move from MPLS to DiffServ. The DiffServ operations are needed only at the edge nodes and there is no need to update the core network. MPLS offers the ISP a solution to administer and maintain the core of DiffServ in the efficient way.

#### 4.1.3 Less than best effort class of service

There is a need to have a service class that offers service less than the default class meant for BE traffic. This service class will offer the unutilized resources for use through a Less than Best Effort (LBE) class. Less than best effort (LBE) class of service is a category for traffic, which is placed below the default class and given a higher drop precedence than the BE class. The LBE class may use all the available resources left from higher classes. The LBE will not bring more quality to higher classes. If there is congestion in the network and all the bandwidth is used, the LBE will bring no more help for that traffic. The LBE class is meant for applications that are not sensitive to delay and loss. The end host of the sender or a boundary router usually premarks the LBE DSCP value to the packets. It is also recommended by appropriate RFC documents to determine a small amount of network resources to the LBE class like one percent that prevents the traffic to starve totally.

Many applications may take advantage of the LBE class. Network backups, file mirroring, bulk data transfers, or experimental data transfers are the area to get benefit of the LBE class. Applications using the LBE class must be tolerant of loss, jitter and delay. There are applications that are used only during the night time to ensure that the traffic will not disturb the already existing traffic. The LBE class makes it possible for these applications to be run during active time without fear to steal resources from the critical data traffic. The benefits of the LBE class have been evaluated in [35]. The authors in [35] concluded that LBE does not have effect on performance of the BE class and higher classes. The test results show that the LBE traffic does not have effect on performance of the BE class and higher classes.

The LBE class has been proposed in RFC (Request For Comments) documents like LE PHB [36] and LE PDB [37]. The LBE class has been implemented and tested by GEANT TF-NGN working group [38] and Internet2 working group [39].

It has been supposed to use standard DSCP value of '001000' for LBE traffic [35]. The value is one of the values used in the Class Selector PHB Group. Some sites in Europe (e.g. Funet<sup>11</sup> network and the GEANT<sup>12</sup> backbone network) and US (Abilene<sup>13</sup> network) are already using the LBE service. The GEANT and Abilene in USA enabled also the LBE service across the transatlantic links.

---

<sup>11</sup> Funet (Finnish University and research NETwork) [40] is the Finnish national research network, and forms part of the NORDUnet network.

<sup>12</sup> GÉANT is a project co-funded by the European Union set up by a Consortium of 27 European national research and education networks (NRENs).

<sup>13</sup> Abilene [41] network is a backbone network that supports the work of Internet2 universities to develop advanced Internet applications. It is a partnership of Internet2, Qwest Communications, Cisco Systems, Nortel Networks, Juniper Networks, and Indiana University.

#### 4.1.4 Intradomain metrics and monitoring

Metrics are a remarkable area related to SLAs between the ISP and the customer. Different metrics are needed to validate the SLA from the customer and the ISP point of view. This section concentrates on the metrics that are important from the intradomain perspective. The interdomain metrics deal with the SLAs made between the IPS and the NSP or between the NSPs and serves the purposes of the NSPs. The interdomain metrics are discussed in Section 4.2.5.

Generally metrics can be divided in operational and performance metrics as shown in Figure 4.5. Performance metrics can be further divided in availability, traffic profile metrics and responsive metrics. The ISP needs to be aware of all the areas of performance metrics to serve the customer according to the SLA.

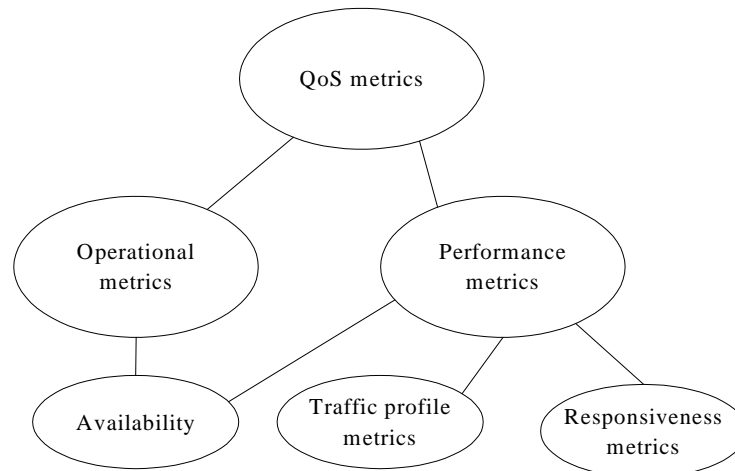


Figure 4.5. QoS metrics related to an Intradomain network

Operational metrics are the traditional metrics that the enterprise is interested in like Mean Time To Repair (MTTR), Mean provision time, Mean Time Between the Failures (MTBF), service delivery time and Mean Down Time (MDT). These metrics have been used to describe the service ISP is delivering. Certain thresholds are usually set for metrics in the SLA.

Performance metrics contain availability, traffic profile metrics and responsive metrics. Availability metrics usually contain network connectivity, outage count, outage resolution time, error rates and packet loss ratio [42]. Error rates and packet loss ratio are performance metrics. Network connectivity means the percentage of time the network is reachable. Outage count describes the number of times network resources could not be reached. Also the time to repair the failure is usually included to the metrics. Error rate is measured through the packets dropped or the packets corrupted. Packet loss rate is measured as the fraction of packets to fail to reach the destination.

Traffic profile metrics describe the performance of network from the DiffServ point of view. Traffic profile metrics contains e.g. CIR, PIR and their associated token bucket parameters. These metrics may be derived from the TCS and are used to describe the traffic classification and conditioning in the domain of the ISP. The incoming traffic might also adhere to a certain traffic profile.

Responsiveness metrics contain parameters such as one-way delay, round trip time (RTT), jitter, throughput and allowable bandwidth. The one-way delay can be measured using a tunnel, a funnel or a cloud approach.

The SLA determines the service that the ISP has committed to deliver to the customer. The SLA must support monitoring and verification to be useful. The promises made in the SLA must be enforceable and there have to be metrics to be used to validate the SLA. In many cases these metrics are specified in the SLA and they describe the quality of service the ISP will guarantee. Metrics are the indicator of the service provided and is the basis for discussion of penalty in the case the metrics are beyond the committed thresholds. Metrics can be considered also as the indicator for the traffic of the customer.

The measurements done in the network of the ISP can be divided in two parts, those that serve the validation of the SLA and the others that serve the purposes of the ISP, such as planning, dimensioning and maintaining the network. The ISP needs to ensure that the end-to-end QoS between the sites of the customer is reached to validate the SLA. In the case of intradomain service this means that the ISP needs to measure the network of its own. The traffic that enters the network of the ISP has to fulfill the traffic profile the customer has committed in the SLA. The incoming traffic has to be checked using the measurements. The measurements may be needed also for accounting, billing and security purposes. The ISP needs the measurements in order to plan the future services, to provision the network capacities and to optimize the usage of network resource to have as much profit of the network as possible.

The network has usually many measurement points depending on the metrics that are needed. Measurement points can be situated in different layers, such as node, link and network levels [43]. Node level measurements are related to the single node and the performance of the node. Metrics may contain router-specific statistics or DiffServ-related classification and traffic conditioning metrics. In the ingress node the incoming traffic of the customer is measured and compared against the traffic profile. The leaving traffic could be measured in the egress node and it can be shaped according the SLA. Core nodes can be measured to gather the information that is important for ISPs to plan the future services and capacity allocation for services. Link level measurements consider the link between the nodes. Link level related issues are e.g. link and aggregate utilization. Aggregate utilization can be measured using the DSCP value. Network level measurements occur between the ingress and egress nodes. Network level metrics describe the edge-to-edge QoS that the intradomain network can offer to the customer. Network level metrics contain one-way-delay, RTT and loss concerning a whole network or a specific traffic aggregate. The customer is mostly interested in network level metrics, whereas node and link level metrics help the ISP to design and maintain the network to serve the customer and to make the revenue. [43]

#### **4.1.5 Vision of Differentiated Services Network**

The networks of ISPs and the corporate LANs are usually lightly loaded networks. The bottleneck is more often situated in the access networks' side. The customer usually buys the amount of bandwidth to meet the needs of the traffic of that moment. There is normally no excess capacity to spend and without a proper classification there might occur some disappointment among the users of certain applications like mission-critical TCP application or interactive voice applications.

The means to respond to the increasing demand of capacity is to increase the capacity of the access network. This might satisfy customers and users of applications for a while, but it will not solve the actual problem. The applications have become more and more greedy and all the available bandwidth is used. There are always applications that will be the first to suffer in the case of congestion.

DiffServ offers an approach to organize the traffic using classification. The classification of the traffic usually corresponds with the business goals of the enterprise. The business critical applications are the first to be served. DiffServ gives the benefits for both the customer and the ISP. The customer can get benefits by better availability of services in the case of congestion. The ISP may serve the customer better, use the bandwidth more effectively and avoid over-provisioning.

The traffic is classified to different classes in accordance with the characteristics of the traffic. The applications are sensitive to different metrics. It is important to clarify the metrics of applications in use. In Table 4.1, it is highlighted some QoS requirements of the most common applications used in the Internet.

Table 4.1. QoS requirements for different applications

Application	QoS requirements		
	delay	jitter	packet loss
VoIP	low	low	medium
video conferencing	low	low	medium
streaming media	medium	medium	medium
Bussiness data	medium	variable	low
e-mail	high	high	low
file transfer	high	high	low

The QoS requirements presented in Table 4.1 can be used as a guideline for application-specific classification. VoIP, videoconferencing and streaming video are all sensitive to delay and jitter. However, these applications have different thresholds e.g. for delay parameters and needs to be treated different way.

A delay is an important parameter for interactive applications. A one-way delay below 150 ms is acceptable for most applications [44]. VoIP and video conferencing applications are very sensitive to one-way delay and it should be below 150-200 ms. VoIP applications are satisfied with a minor amount of bandwidth per session, while video conferencing requires more bandwidth. The difference in requirements for bandwidth advocates the isolation of VoIP and video conferencing to the different traffic classes.

Streaming video applications are more tolerant for delay and packet loss and should not be set to the same traffic class with VoIP and video conferencing. VoIP applications and video conferencing can operate effectively with packet loss rate below one percent and can be acceptable even if packet loss ratio is five percent. However, packet loss over the five percent range is acceptable for typical Web browsing.

The classification is in most cases based on the standardized PHBs. Each standardized PHBs have their own characteristics. The EF PHB for instance is determined to serve the applications that require low delay, jitter and packet loss. It is natural to use EF PHB with VoIP applications, because EF PHB offers the metrics

that are especially important to the VoIP applications. The AF PHB, the CS PHB, the default PHB and the LBE PHB are not so clearly allocated to the certain applications. There are many different alternatives to implement these PHBs into the network. The ISP may use all of the PHBs or a subset of PHBs to offer the QoS to the customers.

The VoIP applications are most often the main reason to have the QoS. VoIP applications suffer from the congestion and users experience the service quality unacceptable. There is a long history in telephone service and the user expects not less than toll quality when using VoIP call. The ear is also very sensitive to the delay in speech applications. VoIP applications can be isolated to their own PHB class to protect the traffic from the disturbance of other applications. The common way to use the EF PHB is voice applications. The rest of the traffic is assigned to the default PHB class. In this case there are no applications that suffer from the congestion others than VoIP applications. The traffic of the network is served by two queues, one for the EF PHB and the other for example, BE PHB. The EF PHB can be implemented using PQ (Priority Queuing), WFQ, CBQ or WRR. PQ ensures the nominal delay and jitter, but there is a risk to starve the queues of lower priority.

When the need for QoS increases, more classes are needed to serve the requirements of the QoS of the customer. The AF PHB group is often the most efficient way to offer QoS to the customers. The AF PHB group has four independent classes with drop precedence levels of their own. The AF PHB group offers a valuable tool to build the QoS for different applications. Applications can be classified to mission critical applications, network multimedia and general data services. Applications should be managed according to the strategies and needs of the enterprise.

The MF classifier offers many ways to select the traffic to the appropriate PHB. Division into the different classes could be based on the protocol to be used. TCP based traffic should be divided to different classes than the UDP based traffic to handle the distribution of bandwidth more fairly in the situation of network congestion. There may also arise a need to classify the TCP traffic to different traffic classes according to the nature of the traffic. Some studies such as in [45],[46] have shown that isolation of the short-lived (e.g. HyperText Transfer Protocol (HTTP ) type) and the long-lived (e.g. File Transfer Protocol (FTP) type) TCP connections improves the predictability of service and fairness.

There are many applications that use TCP, such as telnet, FTP and Simple Mail Transfer Protocol (SMTP). Applications that hold a user waiting until the application answers are served with nominal delay and loss. When telnet and FTP are compared with each other, there is a difference in characteristics. Telnet is sensitive to the delay and packet loss, while FTP is tolerant of delay. There might be need to differentiate these applications into different service classes or at least to different drop precedences.

Network multimedia uses mainly UDP as transport protocol. Simple Network Management Protocol (SNMP) is a common protocol that uses UDP. Network multimedia is sensitive to different metrics than SNMP and it might not be appropriate to determine these to use the same service class.

There is always traffic that is satisfied by the default class (BE PHB). The default class is served as soon as possible and applications have to be tolerant of moderate delay, jitter and packet loss. The class for best effort traffic is sure needed in networks. However, it will not be effective to let the default class to starve totally, some amount of buffers and bandwidth have to be configured for the default class.

Some customers are likely to be interested in the class to be served worse than default class. This LBE service will get network resources if available while it will not be let to starve totally. No guarantees of QoS are given to applications to use this service class. The advantage of this class comes from the possibility to use the network for applications like site mirroring that might disturb the present traffic. Through the LBE service, applications can be isolated into their own service class and the traffic will not disturb the more important traffic. The LBE service might be directed to certain applications by giving it a price below the default service class or it might be used according the strategies of the enterprise.

There will most obviously arise a need to use specified PHBs in communications between the NSPs. Customers will demand the QoS to be enlarged to the peering connections. None PHBs has been yet standardized for peering connections. It is essential to have a common standard to be used between NSPs for peering purposes. The EF PHB, the AF PHB Group and the default PHB are used to build the QoS service inside the network of the NSP. One possibility to fulfill peering connection needs is to use the CS PHB Group.

Table 4.2 shows one possible scenario to configure the network from business point of view, if all the standardized classes are used to offer the QoS for the customer. The LBE class and peering are implemented using the CS PHB class. The AF PHB is used to offer Olympic service like gold, silver and bronze classes of QoS. In this scenario traffic classes are divided to premium, platinum, gold, silver, bronze and best effort. It can be seen from Table 4.2 that there are the classes of their own for less than best effort traffic and peering.

Table 4.2. Mapping between traffic class and application

<b>Traffic class</b>	<b>PHB</b>	<b>Application</b>
Premium	EF	VoIP
Platinum	AF4	video conferencing, multimedia
Gold	AF3	mission-critical applications
Silver	AF2	Web browsing
Bronze	AF1	timely critical
Best effort	BE	best effort data
Less than best effort	CS1	non-timely critical
Peering	CSx	Peering

The EF PHB is used for VoIP applications and required percent of the bandwidth is assigned to this purpose. The amount of traffic of VoIP class depends on the traffic needs of the customer as well as the approach to build the EF PHB service. If PQ is used, the other traffic classes may be starved. It is practical to configure a maximum limit for the bandwidth to be used for VoIP applications. In the case of no VoIP connections, there will be a waste of capacity, if it is reserved for the VoIP connections only. If a scheduler like WFQ is used, it needs a careful installation to guarantee the VoIP packets to be served well enough and avoid queues of the EF PHB to grow.

Platinum, gold, silver, bronze and best effort classes are used to offer the QoS to the traffic that is today treated as best effort. Video conferencing and multimedia applications are sensitive to delay and packet loss. Most often these applications use UDP protocol. The Platinum class is placed after the Premium class and is served by

the AF4 class. UDP traffic will not disturb other classes while it is isolated into the class of its own. The platinum class is served with controlled rate and delay characteristics.

Mission-critical applications are usually sensitive to delay and loss. These applications usually wait a certain response. These applications are served by gold class, which is implemented using the AF3 class. Web browsing is based on TCP and HTTP. It is often an important application for users and might need a class of its own. In Table 4.2 Web browsing has been configured to use the silver class. Timely critical applications like FTP and SMTP may be situated in the same class, like the bronze class to ensure that they get better service than best effort traffic. The bronze class is implemented using the AF1 PHB class. The Default class can be seen as a class for the traffic that has not been determined to any classes mentioned above. The Default class is implemented using the BE PHB and it will be served like best effort traffic. LBE traffic should be marked by the customer. LBE traffic is treated with no service guarantees.

Traffic may be classified further inside the AF class, if needed. The AF PHB class has been standardized to have three levels of drop precedence within an AF class. For instance the AF1 class can be divided into three levels of drop precedence like AF11, AF12 and AF13. The AF11 PHB class has the lowest drop precedence over the others. Classes AF11 and AF12 could be used to classify the traffic further as shown in Table 4.3.

Table 4.3 Classification inside the AF class

Traffic class	PHB	Application
Bronze	AF11	ftp
Bronze	AF12	smtp

In the case of congestion SMTP traffic will be dropped before FTP traffic. If there is no classification between these traffic types inside the same PHB, the both applications will suffer.

The DiffServ offers many different ways to classify the traffic and serve the needs of the customer. Traffic may be classified into different classes with different characteristics and forwarding behaviors. Every ISP implements the PHBs to be used in its own way using its own thresholds and parameters. PHBs of different ISPs are not comparable with each other although they use the same PHB e.g. EF PHB or AF11 PHB.

The QoS service should be build so that it can offer QoS guarantees to a diverse set of applications. However, this granularity may lead to a complex network, thus reducing the benefits of QoS. Large amount of different classes might increase the complexity to maintain, measure and monitor the service. The BE traffic serves the customer with a single traffic class that gives no guarantees to the customer. It is quite an improvement to the existing best effort traffic to be classified to e.g. five traffic classes instead of the one traffic class. There could be traffic classes meant for VoIP applications, mission-critical applications, UDP-based multimedia and Web browsing. The rest of the traffic could be configured to belong to the best effort class. Table 4.4 presents the classes that might be the most important to be established to serve the customer. In many cases, it might be enough to isolate the most important

traffic to the class of its own and guarantee the forwarding of those packets using the DiffServ model.

Table 4.4 The proposed way to classify traffic

Traffic class	PHB	Application
Premium	EF	VoIP
Gold	AF4	Multimedia
Silver	AF3	Mission-critical
Bronze	AF2	Web browsing
Best Effort	AF1	others

The ISP has to plan carefully the use of the elements in the core of the DiffServ network. The idea of DiffServ is to use aggregates, which are to be forwarded using PHBs. Nodes serve the traffic of different customers aggregating the traffic inside the core. For example, VoIP packets of a customer are aggregated with the VoIP packets of the other customer as served by the EF PHB. It is important that the core is designed to work the way it is agreed in the SLA. In the case of the EF PHB, it is expected that the queues are almost empty in all circumstances. The DiffServ boundary node will not be aware of the load in the core, because no feedback is given from the core. This increases the importance of the network dimensioning.

Due to the things mentioned above, the approach to use the MPLS in the core network has become more general instead of the pure DiffServ model. The MPLS helps the NSP to dimension the resources of the network to correspond with the SLAs that the NSP offers. The traffic engineering skills of MPLS will help the ISP to handle the core of the DiffServ network. The MPLS is a very respectable way to build the core of DiffServ until the mechanisms to handle the core in DiffServ nodes will evolve.

## 4.2 Interdomain QoS

### 4.2.1 SLAs between network service providers

The sites of the customer might be situated in different countries, even in different continents. Traffic usually needs to traverse many domains of NSPs. NSPs are connected with each other either using IXPs or NSPs may have bilateral agreements. The customer transfers business-critical traffic between sites and needs to have an SLA to guarantee the transfer.

The chain of SLAs between the customers is shown in Figure 4.6. There have to be SLAs between NSPs and between the customer and an ISP to build the end-to-end SLA. Today the SLA normally occurs in relationships between the customer and the ISP. Network availability, throughput, delay and packet loss are the parameters that are guaranteed in the SLAs today. The ISP is responsible to the customer of reaching the parameters of the SLA. In some cases, ISPs have SLAs to guarantee the transit parameters of the traffic they will deliver. However, in most cases there are no SLAs for that purpose. SLAs are a fundamental part of the end-to-end QoS to be developed. To get real advance of these SLAs, there should be appropriate tools to deliver the QoS between NSPs as well as between ISPs and NSPs. It is very time-consuming to negotiate the SLAs between NSPs manually.

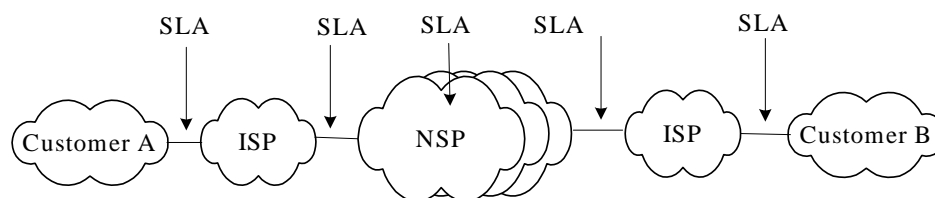


Figure 4.6. SLA through many ISP domains

Some ISPs [47],[48] offer SLAs both on-net and off-net cases. The on-net SLA covers only the network of the ISP. The off-net SLA is offered when the traffic goes from the ISPs network to the network of the specified NSPs. In Figure 4.6 it would mean that off-net SLA is offered to customer A, when its traffic goes from the ISP to the adjacent NSP. No more guarantees are given to traffic traveling further. Performance guarantees are offered using monthly average values. This is a big disadvantage, if we think real time applications, such as VoIP. Off-net SLAs are valuable from point of QoS only if the QoS information can be carried between the NSPs. Business-critical applications must be able to experience the QoS performance all the time. Today's off-net SLAs are valuable for the traffic, which is not business-critical and is tolerant of delay and loss while experiencing most of the time good performance.

#### 4.2.2 Per-Domain Behavior

The ultimate goal for the DiffServ framework is to offer an end-to-end solution for the QoS. The DiffServ WG has designed a Per-Domain Behavior (PDB) to help the evolution towards the end-to-end model. The PDB describes the edge-to-edge solution in a DiffServ network. Thus, the PDB is the next step proposed to develop the Differentiated Services towards an end-to-end solution. The focus of the PDB defined in [49] is to describe the QoS for edge-to-edge in a single DS domain. Technical properties are the same for the end-to-end solution, but there are also commercial facts and agreements that have to be solved out first. It is believed that the first end-to-end solutions will arise in a single DS domain that is administered by a single owner. The solutions will be expanded on domains with different administrators to offer the real end-to-end QoS.

The packets that belong to the same PDB are marked with the same DSCP value and receive the same treatment of PHB. The PDB is closely related to the PHB. As mentioned earlier, the PHB is described to be the forwarding treatment that the behavior aggregate receives in a DiffServ node. The PDB is described to be a definition of forwarding attributes and treatment a traffic aggregate (TA<sup>14</sup>) receives in a network. The PDB is meant as a network definition for a TA, while the PHB focuses on the characteristic of a node or a single link. The Figure 4.7 presents the relation between a PHB and a PDB.

---

<sup>14</sup> TA is a PDB terminology and means a set of packets with a DSCP that maps to the same PHB usually in a DS Domain [49]

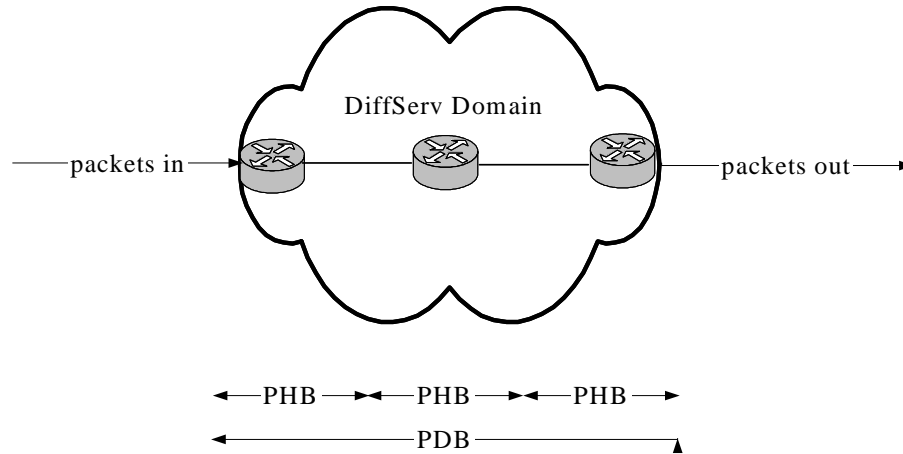


Figure 4.7. Relation between a PHB and a PDB

A PDB contains attributes, which are used to describe what happens to the packets of a particular PDB as they cross the DS domain. The attributes of PDB have the main role in specifying the treatment packets receive in a DS domain. These attributes are very complicated, because they are the sum of classification, traffic conditioning, entering traffic loads and topology of the DS domain. The attributes may be static, absolute or percentile. It is possible that multiple PDBs may use the same PHB and that way the transit characteristics of this PDB will be the same. The differences come from the attributes, which are different for these PDBs

A PDB is composed at the edge of a network by means of classification and traffic conditioning. The target group of packets is distinguished through classification and the packets are marked with DSCP meant to this PDB. The TA is created using traffic conditioning. The TA is then forwarded across the DiffServ network. TAs cross the DS domain splitting and merging on the interior nodes. The attributes that are determined to belong to that PDB have to hold during this operation.

Specific PDBs have been described. These PDBs can be constructed using the existing standardized PHBs, like the EF PHB, the AF PHB group and the default PHB. So far five PDBs have been defined. These PDBs are listed in Table 4.5.

Table 4.5 Different PDB groups

name of the RFC document	Status
Best Effort Per-Domain Behavior	RFC3086 (April 2001)
The 'Virtual Wire' Per-Domain Behavior	Internet draft, expired January, 2001
An Assured Rate Per-Domain Behavior for Differentiated Services	Internet draft, expired August, 2001
An one-to-any Assured Rate Per-Domain Behavior for Differentiated Services	Internet draft, expired August, 2001
A Lower Effort Per-Domain Behavior for Differentiated Services	Internet draft, expired December, 2002

Nichols et al. describe the Best Effort PDB (BE PDB) in [49]. The BE PDB is for today's Internet traffic crossing a DiffServ network. As this BE PDB offers a service

such as today's Internet, no service guarantees are offered, such as certain availability, latency or packet loss. The packets of the BE PDB are marked for the default PHB. Packets of the BE PDB will not be completely starved. Resource that are left from other TA are to be used by the BE PDB. Naturally the NSP can also use bounds to give the BE PDB the quality they want to sell to their customers.

Jacobson et al. have proposed the Virtual Wire PDB (VW PDB) in [50]. The VW PDB describes a service between edge-to-edge in the DS domain. The VW PDB is meant to offer a service similar to dedicated circuits that are used for IP traffic. The VW can replace the dedicated circuit in the DiffServ-aware domain that supports the EF PHB. It is important for the VW PDB that there will be almost empty queues according the typical EF mechanism. The VW PDB is suitable for any packets that are based on traffic that uses fixed circuits (e.g. telephone and leased data lines) and packets that has similar delivery requirements (e.g. IP telephone or video conferencing). The VW PDB can replace some part or all of the physical wire between two points. The DiffServ domain that uses VW PDB is invisible to the sender and the receiver. Picture 4.8 depicts the use of the VW PDB.

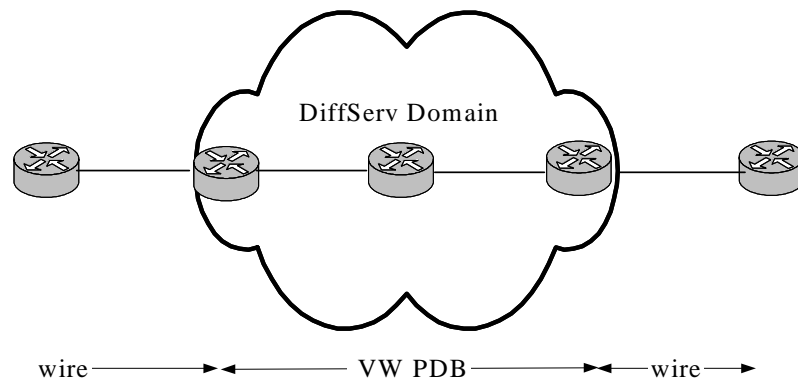


Figure 4.8. The VW PDB as a part of the dedicated circuit

Seddigh et al. have proposed in [51] an Assured Rate PDB (AR PDB) for traffic that require rate assurances, but no delay and jitter bounds. AR PDB traffic might get excess bandwidth beyond the committed information rate, but there is no guarantee of extra bandwidth. The AR PDB is created using the AF PHB. Different types of service topology can be constructed using the AR PDB, such as one-to-one, one-to-few or one-to-any. Attributes used for this PDB are a rate that is assured and low drop probability. Also some parameters must be included to the AR PDB like CIR that is assured with high probability, traffic parameters that are needed to measure CIR and maximum packet size for the aggregate. A policer marks the arriving packets with colours to describe the drop precedence (green, yellow and red). The packets with assurance are marked as green. Green packets are dropped last in the case of congestion. The assumption has been made that the possibility to drop packets marked with green colour is very low. The example use for the AR PDB is e.g. one-to-one or one-to-few VPN (Virtual Private Network)-like services and one-to-any funnel services [51].

Brunner et al. have proposed in [52] a one-to-any AR PDB. This is a special case of the AR PDB. It is very expensive to assure CIR with almost no drops in this one-to-

any case. One-to-any AR PDB considers the possibility that assured rate will not be met with a certain probability. This PDB is used in situation, where one ingress point is sending data to any egress point of the DS domain. The example user might be e.g. a Web site that is wanted to provide its users with high-speed access to its Web pages.

Bless et al have proposed in [37] a Lower Effort PDB (LE PDB) for traffic that is served for a lower priority than best-effort traffic. The LE PDB is meant for sending non-critical traffic across a DS domain. The idea is to delay or drop LE PDB packets, if there are other traffic present. The LE PDB is proposed to be used e.g. as a tool for operators to protect their networks for selected types of traffic. This LE PDB is meant for multimedia applications using UDP, Netnews, content distribution like Napster traffic and for traffic generated by world wide web (WWW) search engines while they gather information from web servers.

There is a need to have a standardized way to traverse the domains, which are administered by different operators. The document [49] describing the PDB is informational, which tries to give answers in the area of intradomain. This document provides a start to introduce end-to-end QoS in future networks. The issue of the PDB is very complicated and more research efforts are likely to be needed before the end-to-end QoS will be reality. It appears that PDB has not got significant interest among researchers, because RFC documents have not been renewed and no experimental documents have been done.

### 4.2.3 Exchange points

Internet exchange points (IXPs) and network service providers (NSPs) compose the heart of the current Internet backbone. An IXP is a place where many domains exchange traffic with each other.

Peering is an agreement made by two NSPs to exchange the traffic of their customers. Normally this agreement is done without payments and it is usually done among NSPs of equal size. Without peering the packets of customers have to travel maybe through many NSPs and the path between the source and destination might have many hops. Delay, loss and jitter increase, when the packets need to traverse many different equipment and links. Peering is the way to offer customers shorter paths and better quality in the network. However, it is impractical to have a peering agreement with every NSP in the Internet. Alternative to bilateral agreements of peering is to join to an Internet exchange point.

The idea behind the exchange points is to offer a shorter route for the Internet traffic between customers of different ISPs. There are non-profit IXP organizations as well as commercial ones. Most of IXPs are non-profit organizations owned by the ISPs and are maintained by annual fees of members. IXPs are operated by a single organization. There are IXPs in every continent and many countries have at least one IXP.

The NSPs are demanded to have a connection of their own to the Internet and an autonomous system number before they can participate to the IXP. The IXP is used only for peering purpose, not the main route to the Internet. Figure 4.9 presents the situation when bilateral SLAs (Path A) are used or peering through IXP (Path B). The advantage of the IXP is the chance to reduce costs of upstream Internet connections and benefit of better quality to offer to the customers. An ISP can join to the IXP directly, if it fulfills the demands of the IXP. Usually ISPs are not large enough to join the IXP and the natural way is then to join the IXP through an NSP.

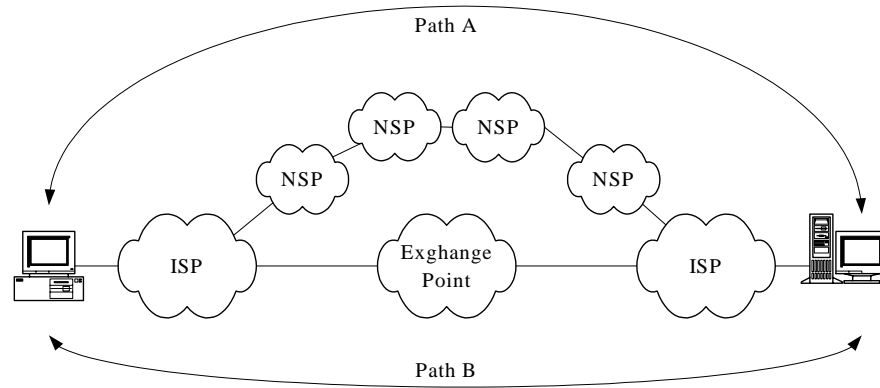


Figure 4.9. Path of packets traversing through networks

An IXP is usually a local area network (LAN) using a fast switch technology. These switches operate at the layer 2. NSPs are free to make connections with each other. Usually no SLAs are made between customers for end-to-end traffic. This restriction comes from the complexity to measure the quality of the end-to-end path, because there is no access to the routers in the network of the NSP. The strategy has been to offer enough bandwidth to meet the requirements. Some IXPs have offered SLAs for end-to-end traffic (such as XchangePoint<sup>15</sup>), but usually no SLAs are made.

IXPs improve the performance of ordinary Internet traffic by reducing the number of hops between customers. However, they will not yet offer remarkable improvement to delay, jitter and loss sensitive applications, such as VoIP, videoconference and other performance-critical applications. Increasing bandwidth is not enough to satisfy the end-users QoS demands that use multimedia applications. There should be a possibility to use DiffServ-enabled functions. ISPs may have an agreement to provide QoS through certain DSCP values. New switches have been implemented to have support to the layer 3 functionality in order to convey DSCP-marked packets. There should also be support for these DSCP-marked packets in the network of the IXP, such as dedicated queues for certain DSCP values.

#### 4.2.4 Policy-based networking in a DiffServ network

Today most of the routers of medium- and large-scale DiffServ networks are configured manually to correspond with high-level abstract notations in SLAs. It is quite difficult, costly, inefficient and in future maybe impossible to configure manually the notations in heterogeneous complex networks and devices. The appropriate parameters for queue mechanisms or drop precedence strategies, admission control and bilateral SLAs are areas of policy that should be implemented through automated management instead of manually configuration to reach the goal of consistent QoS. It is necessary to make the high-level human readable rules and to translate them into the device-oriented commands.

The IETF RAP (Resource Allocation Protocol) working group has developed a framework for Policy Based Networking (PBN). The goal is to offer framework that

<sup>15</sup> XchangePoint is a company dedicated to offering its customers Metro Area Interconnectivity (MAI). XchangePoint is live in London, with MAI planned for Paris and Frankfurt.

is scalable, interoperable and ease to use. The PBN supports the distribution and maintenance of policy rules among a set of network elements.

Policy is determined as a set of rules to administer, manage and control access to network resources [53]. It can also be determined as a method of action to guide and determine present and future decisions [54]. In generally policy is an instruction of network behavior.

The main elements of the PBN are a Policy Decision Point (PDP) and a Policy Enforcement Point (PEP). The PEP is a policy aware node, which will enforce the policy by implementing policy-based admission control for IP traffic. The PDP decides whether or not to admit a specific data flow to enter to the certain DS domain. There exists at least one PDP in each domain. The PEP could be a boundary or core network node depending on assignment it has been implemented. The PDP is usually a separate device, which uses policy repository to store and retrieve the information of policies. Policy repository may be a relational database or LDAP-based (Lightweight Directory Access Protocol) directory. Network administrator sets the policy rules using a Policy Management Tool (PMT). The PMT translates the high-level policies to the format, which network elements can interpret. The PMT is a software component that can be located at independent device or it could be as a part of PDP. Basic elements of PBN are shown in Figure 4.10.

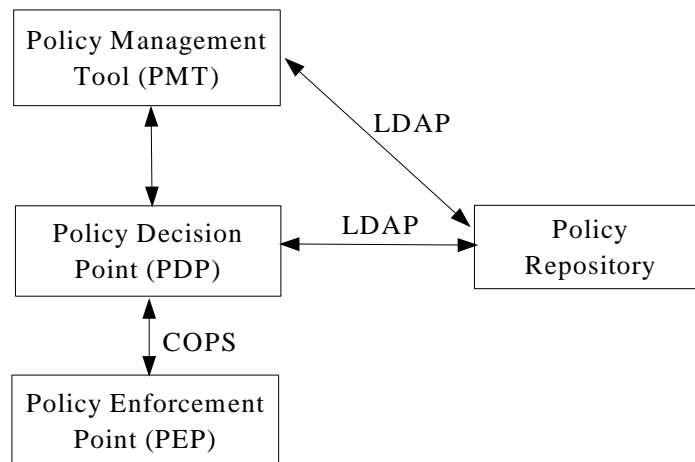


Figure 4.10. Policy framework

The operation between policy elements is mostly the communication between the PEP and the PDP. The PEP receives the request for network resource. It sends the request to the PDP or it might have a local PDP (LPDP). In the case that the PEP has the LPDP, it is first queried. The result of query and the resource request are then send to the PDP. The PDP uses the policy repository in process to make a decision.

The communication between the PDP and policy repository is usually done using e.g. LDAP or SNMP mechanisms. The PDP makes the final decision maybe over the LPDP decision. The PDP returns the decisions to the PEP, which then enforces the policies by accepting or denying the request.

Communication between the PDP and the PEP can be done using Common Open Policy Service (COPS) protocol. The COPS is a TCP-based query and response protocol. The COPS uses TCP as a transport protocol to establish a reliable exchange of messages between policy clients and servers [55].

The goal of the DiffServ framework is the controlled sharing of the bandwidth of the customer. This control can be done using agents, which are aware of policies in the

network and allocate resources according to these policies. Bandwidth Brokers (BBs) are proposed to be these agents. The BB is seen as an entity to be configured with policies, to keep track of current allocation of marked traffic and to interpret new requests to mark traffic using policies and current allocation [56]. BBs are assigned two main functions. They should allocate and maintain the resources in a single DS domain as well as handle messages sent to the adjacent DS regions. Figure 4.11 shows the DiffServ network that uses BBs to allocate network resources.

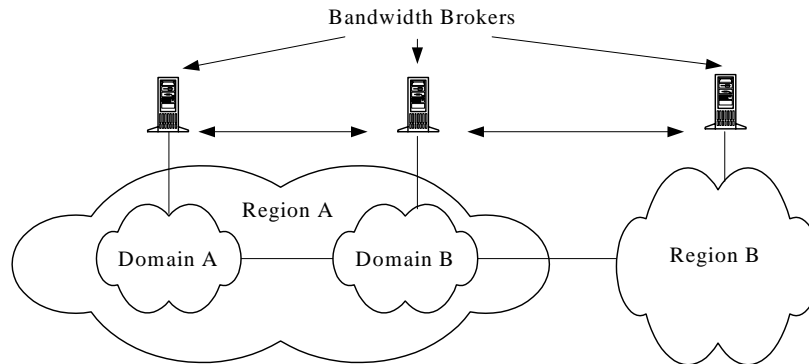


Figure 4.11. Bandwidth Broker architecture

In Figure 4.11 every DS domain has its own BB. BBs communicate with the BB in the region and with the BB of the adjacent region. The BB needs to keep a state of every resource. The requests for resources are sent to the BB, which makes the decision.

The BBs have faced many drawbacks. One main problem is the scalability problem due to the traffic handled by a BB. In some cases the traffic requests directed to one BB are so huge that a single BB can not handle all these requests, thus becoming a bottleneck of the DS domain. This can be improved by employing appropriate policy caching mechanisms [57].

#### 4.2.5 Interdomain metrics and monitoring

The customer has a contract with the ISP, which in turn has a contract with the respective NSP. The NSP might have many different contracts with other NSPs like bilateral transit or peering agreements. This section concentrates on the relation between the ISP and the NSP and the relation between the NSPs. In this section the ISP is always a buyer and the NSP is either a buyer or a seller.

NSPs sell the capacity of network to the ISP, which in turn sells the service to the customer. ISPs are interested in the end-to-end QoS to be provided to the customer. The main interest the NSP has is to design, operate and maintain the network the way that maximizes the revenue and the profits. The above mentioned issues are of course also in focus of the ISP, if the ISP has the domain of its own.

NSPs operate with each other using transit or peering connections. As mentioned earlier in Section 4.2.3 the peering connections usually do not have the SLA between the NSPs. Thus, it is important to carefully meter the incoming as well as outgoing traffic to ensure the maximal benefits to interconnect with each other. Bilateral SLAs would clarify the concept of peering and help in building the end-to-end QoS.

Transit connections are based on bilateral agreements and contain an SLA to be fulfilled. The SLAs made between the NSP needs to be validated. The seller NSP is

expected to verify that the parameters, capacity and characteristics of the traffic transported through the network domain of seller corresponds to the SLA. The seller NSP will also ensure that the traffic entering from the network of buyer NSP will confirm the traffic profile agreed in the SLA. Metric should also give information to the NSP for self-testing, design, develop and resource utilization. Metrics that are useful for this purpose can be e.g. throughput, goodput, loss and delay. Metrics that were mentioned in Section 4.1.4 in case of intradomain metrics are relevant addition also in interdomain case.

The ISP is mostly interested in the end-to-end QoS service that will be delivered between the customer sites. The end-to-end QoS could be verified through measurements in the ingress side of the ISP and the egress side of the other ISP. It is not an easy task for the ISP, because the NSPs and other ISPs are not willing to give the information on their network. One way to resolve this problem is the use of trusted third parties, that would meter the networks of NSPs to provide the independent end-to-end QoS metrics and statistics. The metric needed to verify the end-to-end service performance can be for example end-to-end one-way-delay, RTT, jitter, loss and throughput [43].

#### **4.2.6 Means to build the QoS**

There is a need to carry out the end-to-end QoS in the area of interdomain networks. The basic elements of the end-to-end QoS inside the intradomain network are PHBs, PDBs and DiffServ combined with different techniques like IntServ and MPLS. Most of them are already in use in intradomain area and offer the QoS to meet the needs of the customer described in the SLA. It can be seen that QoS should be brought in a single domain and thereafter tackled an interdomain QoS [58].

Static or dynamic SLAs are the basis to build the interdomain QoS. There is a strong tendency that SLAs will evolve from static into the dynamic SLAs. NSPs need to have bilateral SLAs in the case to exchange the traffic to satisfy the end-to-end QoS. It is essential that IXPs offer a way to deliver QoS through their network. It might be certain DSCP values that are served with separate queue mechanisms to implement the QoS. The NSPs that need the QoS in the IXP, should be served according to the QoS needs and separated from the NSPs that are satisfied by the default class of service.

Standard negotiation processes and standardization of SLS parameters are an important area that is in focus of many research projects. Standardization would help the negotiation between the NSPs and would offer an explicit meaning to different terms to be negotiated.

Technical tools are mainly ready to be used. However, it might take a while to build the concept to use these tool in area of interdomain networks. More efforts are also needed for the business models to develop to correspond to the needs of interdomain end-to-end QoS.

## 5. Differentiated Services from the customer point of view

### 5.1 User-metrics and monitoring

End-users usually experience the performance within the network through the applications and this experience is comprised of end-to-end QoS. The end-to-end QoS contains also the performance of servers, workstations, LANs and WANs (Wide Area Network). A network-level SLA concentrates only on the QoS of the network. The situation within applications and the way applications use the network affect an end-to-end QoS the end-user will experience. There are many QoS measurements that can be recorded, such as throughput, goodput, packet loss, delay and jitter. Throughput and goodput, measured in bytes (or bits) per second often differ since the latter is determined as the amount of user data received in a given interval excluding protocol overhead and bytes retransmitted. Throughput is the metrics commonly monitored by ISPs, whereas more meaningful predictor of application performance is goodput. The metrics such as loss, delay and jitter are the first to have effect on the QoS. Goodput and delay are two of the main relevant features to the user. End-user finds that applications work in the expected way, if the metrics are within bounds. Otherwise end-user needs to wait excess time for the application to answer to the query or there will be time-out in the worst case. If problems occur often in the network the hazard is that end-users get used to that specific problem and will not report to the help desk.

An enterprise is more often concerned on parameters like bandwidth and reliability. Reliability may be presented in terms like Mean Time to Repair (MTTR), Mean Time Between Failure (MTBF) and Mean Down Time (MDT). MTTR, MTBF and MDT are usually stated in an SLA and effect in case of a failure to the billing through penalties of ISP. These metrics are usually monthly metrics and describe the delivered service as a whole.

In a DiffServ environment, it also has to be verified that an SLA is met in traffic differentiation. Different traffic classes have to get the share determined in the SLA and that has to be verified. Also, the extra capacity should be delivered between the classes according to the SLA. ISPs have to verify DiffServ specific characteristics such as maximum delay, minimum packet loss and maximum jitter that ISP offers for the specific traffic class. The characteristics of classes are determined in an SLA and ISPs use suitable metrics to verify agreements.

Traditional subjective metrics, such as Mean Opinion Score (MOS) can be also used to determine e.g. the quality of VoIP sessions. MOS is based on subjective test of the end-user. MOS is performance metrics that complete the Figure 4.5 shown in Section 4.1.4. End-user determines the quality of speech using the scale from 1 to 5, were 1 denotes for bad performance and 5 an excellent performance The weakness of MOS is that it needs human resources to be used in tests.

ISPs use different measurements and tools to verify the quality of their service. Customers might as well use verifiers of their own to see, whether the service the ISP has promised will be delivered to them. The verifiers may be situated so that the measured traffic will go through the verifier. This will give the most reliable results, but it might adversely affect the performance of traffic traversing the verifier. In case of failure in a verifier, there might be a risk to have effect on the performance of traffic. Measurements can be done either using active measurement with probe traffic

or passive measurement of real traffic. Today, many vendors (such as Acterna, Agilent, Brix, Finisar, NetIQ and Sniffer Technologies) offer verifiers that can be used to measure the quality of the speech of VoIP session [59].

Active measurement is based on added traffic in the network and measuring the response. The final result is based on the location of the probes. Probes may be located near the router of an ISP or near the end-system. Active measurement generates the traffic and does not measure the real traffic. An excess load of the traffic is also generated to the network and it might influence to the performance of the real traffic. It is essential to select carefully the place to make measurements, the traffic to be measured and the time interval to be used in measurements to reach the result that corresponds to the real traffic without violating the existing traffic.

Passive measurements do not generate additional traffic load to the network but they analyze the existing traffic. This can be done through probes or using existing information in network nodes like routing tables. Probes can be placed e.g. in a CPE.

A wide selection of measurement tools is available. There are both commercial and non-commercial tools. Many Web sites (such as CAIDA<sup>16</sup> [60] and SLAC<sup>17</sup> [61]) maintain the lists of the available management tools.

## 5.2 QoS in customer network

Ethernet is the widely used LAN technology. Ethernet equipment that are used in enterprises are mostly switches. Traditional Ethernet provided 10 Mbps or 100 Mbps transmission rates in ideal circumstances. Today, Fast Ethernet is offered with 100 Mbps transmission rate as well as 1 and 10 Gigabit Ethernet. Even if high-capacity network is used, there is a possibility that greedy applications will grab the available resources. Extra bandwidth will be the first means to protect the vulnerable applications but prioritization of traffic might give the desired result with very small (or even zero) increase of capacity.

Intelligent switch technology offers some means to protect mission-critical applications from the congestion in LAN. These intelligent switches can differentiate traffic generated by different applications using prioritization in LAN. Traffic prioritization can be achieved on the second layer switches using the IEEE 802.1p technology. Layer 2 switches group incoming LAN packets into separate traffic classes. 802.1p defines eight classes of priority. Intelligent switches may map layer 2 priority to layer 3 DSCP values to enable traffic differentiation in the network layer. Table 5.1 shows one possible scenario to map 802.1p priorities to DSCP values.

---

<sup>16</sup> CAIDA (Co-operative Association for Internet Data Analysis) provides tools and analyses promoting the engineering and maintenance of a robust, scalable global Internet infrastructure.

<sup>17</sup> SLAC (Stanford Linear Accelerator Center) designs, constructs and operates state-of-the-art electron accelerators and related experimental facilities for use in high-energy physics and synchrotron radiation research.

Table 5.1 Mapping 802.1p priorities to DiffServ PHBs

IEEE 802.1p	PHB
7	Network
6	EF PHB
5	AF PHB
4	AF PHB
3	AF PHB
2	BE PHB
1	BE PHB
0	BE PHB

The need for prioritization will arise among customers when the business critical applications suffer from congestion in LAN although there should be capacity available. Most often existing LAN equipment are then replaced with intelligent switches. However, prioritization in layer 2 is only meaningful when all the components in LAN can take part in prioritization. Otherwise customers need to be satisfied with layer 3 prioritization.

Mostly switches operate at layer 2 (link layer) but some of them also offer mapping between layer 2 and layer 3. Many vendors (e.g. Allot Communications, Packeteer, Sitara Networks, Kentrox, NetReality, Hewlett-Packard and Cisco) offer intelligent switches that have QoS characteristic [62].

Enterprise might also find it difficult to deliver the bandwidth fairly between the sites or from a site to the Internet. For example, the performance of business critical traffic may deteriorate when interacting with bandwidth greedy traffic. DiffServ offers a valuable means to handle this kind of situation. Customer may also use equipment that shapes the traffic traveling in the LAN before it enters to the WAN router as shown in Figure 5.1.

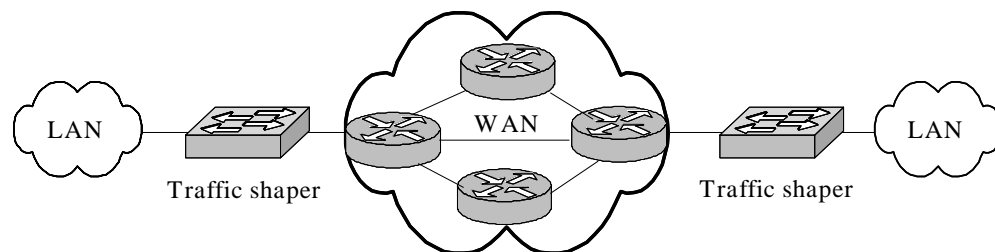


Figure 5.1. Traffic shaping with independent equipment

Traffic shaper (such as Packeteer's PacketShaper [63] or Staselog's Network Equalizer [64]) offers an application traffic management system that monitors, controls, and improves application performance over the link between the company LAN and the Internet. The enterprise may shape the existing traffic and effect on the

bandwidth usage. Unnecessary traffic may be blocked out or its bandwidth may be reduced to minimal value. The use of bandwidth in WAN link will become more efficient and business-oriented.

### 5.3 Fairness and the end-to-end QoS with TCP and UDP

Fairness is an important subarea of QoS. Customers will not directly see the fairness issues, but fairness may have a remarkable effect on performance of individual flows. It is shown in many studies that DiffServ has problems in terms of fairness. Problems may occur inside the aggregate or among the aggregates. In this context fairness is discussed from intra-class (fairness within an aggregate) and inter-class (among the aggregates) perspectives. Intra-class fairness deals with fairness between responsive (TCP-like) and unresponsive (UDP-like) traffic and bandwidth sharing among flows<sup>18</sup> in an aggregate. Inter-class fairness discusses mostly the sharing of excess bandwidth among aggregates and behavior of different aggregates sharing the given bandwidth.

Fairness is considered to mean the right for a flow or an aggregate to receive the QoS characteristics stated in the SLA. In this thesis aggregate means a set of packets marked with the same DSCP value traversing the same direction in the DS domain.

Intra-class fairness has been the subject of the most research done in the area of fairness. A lot of interest has been given to the co-operation between responsive (TCP-like) and unresponsive (UDP-like) flows as well as behavior of TCP flows with different characteristics.

The majority of Internet traffic today is carried by TCP flows. However, the amount of unresponsive UDP traffic will increase as the multimedia and VoIP applications increase their favour. TCP and UDP differ from each other mainly by their characteristics during congestion. TCP uses a congestion control scheme to react to congestion [65]. Congestion control scheme consists of TCP slow start mechanism with congestion avoidance mechanism. Slow start phase is very aggressive way to increase the sending rate after the congestion situation in the network. Congestion control mechanism offers a moderate way to increase the sending rate after the congestion. UDP has no congestion control mechanism at all. In case of congestion, TCP flows will back off and UDP flows increase their share of bandwidth. Because of the different nature of TCP and UDP flows some mechanism is needed for fair delivery of these flows.

Mostly, it is seen that TCP traffic should be protected from the misbehaving UDP traffic. It has been proposed in [66] that TCP and UDP should be served with different drop precedence inside the service class. Similar conclusion is drawn in [67]. In this approach TCP flows get lower drop precedence than UDP flows. Seddigh and Pieda showed in their study [66] that TCP can be protected from UDP using different drop precedence but the fairness for both TCP and UDP could not be met using different drop precedences only.

UDP flows are more often used to carry the critical applications of customers and therefore these flows can not be penalized for TCP flows. UDP in-profile and TCP in-profile flows should be treated according to the needs of customers. The separation of UDP and TCP to different classes and different queues is proposed in [66, 68] to offer fairness for both TCP and UDP traffic. The separation will realize

---

<sup>18</sup> A flow means in this context a sequence of packets with a common source and destination.

fairness issues as well as restrict the delay that UDP packets experience in the queue. This will make the AF PHB Group more suitable for real-time service [66].

Seddigh et al. in [68] studied five factors and their effect on bandwidth assurance issues. These factors were RTT, number of micro-flows in a target aggregate, size of the target rate, packet size and existence of non-responsive flows. The results show that all these factors have an effect on the throughput the flows experience. In overprovisioned networks, all the target rates are achieved, but the distribution of the excess bandwidth will depend on these factors. The results show that TCP flows with short RTTs will use greater share of excess bandwidth than the TCP flows with longer RTT. In under-provisioned networks the target rates will not be achieved. These above mentioned five factors have an effect on degradation of target rates.

The interaction with TCP flows that have different size has been in the interest of many researchers. TCP flows are either long-lived (like FTP) or short-lived (like HTTP) and differ from each other by their size or lifetime. Only a small percentage of Internet traffic is long-lived flows, but these flows carry a large percentage of the traffic [45]. It is shown in many studies [46, 69, 70] that short connections lose out to long connections in competition of available bandwidth. This is consequence of TCP congestion control mechanism. Short-lived connections end their transmission, before the congestion window grows large and spend most of the time in slow start phase. Long-lived connections increase their congestion window during the slow start phase aggressively and spend most of the time in congestion avoidance phase. It is proposed that short-lived and long-lived connections should be classified to different classes to improve the fairness and service predictability.

Brownlee and Claffy have studied in [71] the Internet traffic flows. They separated streams into four different categories based on their size and lifetime. Flows are categorized to be long-lived, short-lived, high-volume or small-volume flows. The authors made clear distinction with size and lifetime of traffic streams. They also pointed out that size and lifetime are independent dimensions. They found that most of the streams (about 45%) last less than 2 seconds. However, the long-lived streams (hours to days) carry the most of the bytes in the Internet. The rapid growth of the Internet line speeds and improvement of computer hardware have given rise to growth of Web object size. The rate of the Web object size up to 50 kB is common today. This has to be noticed when ISPs design and dimension their networks.

There are many issues that have influence on inter-class fairness such as number of flows in the aggregate, target rate of an aggregate or network tuning in the core. Seddigh et al. showed in their study [68] that the number of flows in an aggregate can have notable effect on the excess bandwidth the aggregate will have. An aggregate that has more flows will be allocated more of the excess capacity in the overprovisioned case while in the underprovisioned case the difference is not remarkable. This has effect on cases where a customer will have thousands of flows in the aggregates, while an other customer will have only hundreds of flows [68].

Tuning in the core effects on the fairness between the aggregates. It is essential that core routers have a mechanism to control the aggregates in a fair manner, because edge routers do not have information about the available core resources. The excess bandwidth in case of overprovisioning as well as the degradation of bandwidth in case of underprovisioning should be given further research to meet the fairness issues.

## 6. Conclusions

The Internet has moved from an experimental and research-oriented network to a commercial network, where best effort service no longer fulfills the expectations of customers. An increasing number of performance-critical applications require some kind of QoS guarantees from the Internet. This thesis concentrates on DiffServ mechanisms and the feasibility of DiffServ to realize the requirements of QoS in intra- and interdomain networks. The QoS is determined in the SLA and this thesis focused also on the shortages of today's SLAs.

DiffServ is a respectable QoS concept that offers the basic tools to build the QoS. The IETF has standardized DiffServ building blocks during years 1998-2003. DiffServ Working Group was closed in spring 2003 when it had finished all of the tasks that had been set. Many vendors have implemented DiffServ into their own equipment and some ISPs already use DiffServ in their networks. DiffServ provides a valuable way to classify the traffic according to the strategy of the customer. Issues, such as SLAs, policy, other techniques, metrics, fairness, have an effect on the realization on DiffServ QoS and they have to be taken into account.

SLAs are the essential tools between the customer and the ISP as well as between backbone operators. SLAs and the DiffServ-based QoS have to correspond with each other to serve the needs of a customer and an ISP. Today's SLAs are mostly static and serve only the traditional customers that need to update the network characteristics very seldom. In the future SLAs are expected to be dynamic. Dynamic SLAs will make the negotiation more effective and improve the utilization of resources. SLAs and SLSs should be standardized to make the negotiation easier. Customers and ISPs as well as NSPs could negotiate using parameters that are understandable for both sides. Standardized SLAs could help also the customer to compare the SLAs with each other. SLAs are an important part in the chain of developing the end-to-end QoS.

SLAs have to be verified to a customer to ensure that the promised QoS has been reached. Customers may monitor the QoS on-line, if the provider has that kind of service or customers may use verifiers of their own. It is important that also metrics between the NSPs will develop and offer means to measure the end-to-end QoS. Another way is to use a third trusted party to provide the end-to-end QoS metrics. Today many tools are available to monitor the QoS for operators and customers.

Fairness is a remarkable area of DiffServ concept. Although customer will not see fairness directly, it has an effect on the performance of applications. A lot of research has been done about the behavior of flows and aggregates inside the DiffServ network. In this thesis it has been given some guidelines to plan the classification of traffic in the DiffServ. TCP and UDP flows should be classified to different classes as well as short-lived (like HTTP) and long-lived (like FTP) TCP flows should be served through different classes because of the nature of TCP behavior. The classification based on the applications and their behavior is a key component as ISPs construct the QoS services. Standardized rules might help ISPs and NSPs to use the DiffServ and would help to interconnect with each other. This thesis gives some guidelines to map the applications to different traffic classes (PHBs). Especially the use of the AF PHB will be in important role, because of the specific nature of drop precedence levels and different classes the AF PHB offers. The AF PHB offers quite unlimited possibilities to build a DiffServ network.

DiffServ can be combined with other techniques such as IntServ and MPLS in order to get better predictability of service. MPLS technique is a very promising technique and lot of research has been done for the co-operation of DiffServ and MPLS technology. MPLS offers effective traffic engineering to the core of the DiffServ network, but expects careful network resource configuring and provisioning to meet the requirements that are set to the DiffServ-capable network. DiffServ evolution will continue in future and researchers will develop DiffServ models using more effective and fair marking, queuing and scheduling mechanisms. It might also be that whole new QoS techniques will emerge. The Internet will not hardly ever become a one big DiffServ environment but it will be composed from different QoS mechanisms. It is essential that there is a standard that offers rules for different QoS mechanisms to build the ultimate QoS in the Internet.

The disadvantage of DiffServ is its incapability to be aware of the utilization of the resources in the core. The PBN is the area to develop a way to handle the resource utilization in an effective way. Bandwidth brokers turned out to be too complicated and vulnerable and more development is needed in that area. More effective traffic classification would be possible using viable PBNs. PBNs should be developed bearing in mind that the idea of DiffServ is to keep the core simple and bring the complexity to the edge, not forgetting the scalability of DiffServ network.

An end-to-end QoS is the most demanding challenge the researchers, vendors, ISPs and NSPs will have in near future. Using the commercial Internet, it is expected that there are means to offer reliable on-net and off-net SLAs and operators interconnect with each other in an advanced way. IXPs are in major role as the most of the Internet traffic traverses IXPs before it reaches the ultimate goal. Today most IXPs will have no QoS services, therefore no real interdomain QoS cannot be established through the IXPs networks until QoS characteristics are supported in the domain of IXPs. Explicit standardized rules are needed to determine the way to carry out the QoS information like DSCP value. A PDB was an effort to build the basis for the edge-to-edge QoS and to develop the mechanism towards the end-to-end QoS. Due to its informational nature the PDB Internet-draft document did not raise enough interest to make it an experimental or standard RFC document. The end-to-end issue is complicated not only from the technical point of view but also the commercial models need to develop to meet the needs of the customers as well as ISPs and NSPs. Yet there is plenty of work to do before the end-to-end QoS becomes reality.

This work has mostly emphasized DiffServ. However, in all likelihood the Internet will be composed of many different QoS techniques and DiffServ is sure one of them. The primary goal is to develop end-to-end mechanisms and SLA standards that would contain a standardized way to implement QoS and would not depend on the techniques used.

**REFERENCES**

- [1] R. Braden, D. Clark, S. Shenker, "Integrated Services in the Internet Architecture: An Overview", *Network Working Group, RFC 1633*, June 1994.
- [2] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, W. Weiss, "An Architecture for Differentiated Services", *Network Working Group, RFC 2475*, December 1998.
- [3] E. Rosen, A. Viswanathan, R. Callon, " Multiprotocol Label Switching Architecture", *Network Working Group, RFC 3031*, January 2001.
- [4] B. Carpentier, K. Nichols, "Differentiated Services in the Internet", *Proceedings of the IEEE*, Vol. 90, No. 9, pp. 1479-1494, September 2002.
- [5] S. Floyd, V. Jacobson, "Random Early Detection Gateways for Congestion Avoidance", *IEEE/ACM Transactions on Networking*, Vol. 1, No. 4, pp. 397-413, August 1993.
- [6] D. Clark, W. Fang, "Explicit Allocation of Best Effort Packet Delivery Service", *IEEE/ACM Transactions on Networking*, Vol. 6, No. 4, pp. 362-373, August 1998.
- [7] K. Nichols, S. Blake, F. Baker, D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", *Network Working Group, RFC 2474*, December 1998.
- [8] K. Ramakrishnan, S. Floyd, "A Proposal to add Explicit Congestion Notification (ECN) to IP", *Network Working Group, RFC 2481*, January 1999.
- [9] J. Heinanen, R. Guerin, "A Single Rate Three Color Marker", *Network Working Group, RFC 2697*, September 1999.
- [10] J. Heinanen, R. Guerin, "A Two Rate Three Color Marker", *Network Working Group, RFC 2698*, September 1999.
- [11] I. Andrikopoulos, L. Wood, G. Pavlou, "A Fair Traffic Conditioner for the Assured Service in a Differentiated Service Internet", *Proceedings of the IEEE International Conference on Communications (ICC2000)*, Vol 2, pp. 806-810, New Orleans, June 2000.
- [12] B. Davie, A. Charny, J.C.R. Bennett, K. Benson, J.Y. Le Boudec, W. Courtney, S. Davari, V. Firoiu, D. Stiliadis, "An Expedited Forwarding PHB (Per-Hop Behavior)", *Network Working Group, RFC 3246*, March 2002.
- [13] A. Charny, J.C.R. Bennett, K. Benson, J.Y. Le Boudec, A. Chiu, S. Davari, W. Courtney, V. Firoiu, C. Kalmanek, K.K. Ramakrishnan, "Supplemental Information for the New Definition of the EF PHB (Expedited Forwarding Per-Hop Behavior)", *Network Working Group, RFC 3247*, March 2002.
- [14] G. Armitage, B. Carpenter, A. Casati, J. Crowcroft, J. Halpern, B. Kumal, J. Schnizlein, "A Delay Bound alternative revision of RFC 2598", *Network Working Group, RFC 3248*, March 2002.

- [15] J. Heinanen, F. Barker, W. Weiss, J. Wroclawski, "Assured Forwarding PHB Group", *Network Working Group, RFC 2597*, June 1999.
- [16] W. Fang, N. Seddigh, B. Nandy, "A Time Sliding Window Three Colour Marker (TSWTCM)", *Network Working Group, RFC 2859*, June 2000.
- [17] M. Goyal, A. Durrezi, P. Misra, C. Liv, R. Jain, "Effect of Number of Drop Precedences in Assured Forwarding", *Proceedings of the IEEE Global Telecommunications Conference (Globecom'99)*, Vol. 1(A), pp. 188-193, Rio De Janeiro, December 1999.
- [18] E. Marilly, O. Martinot, H. Papini, D. Goderis, "Service Level Agreement: A main challenge for Next Generation Networks", *Proceedings of the 2<sup>nd</sup> European Conference on Universal Multiservice Networks (ECUMN'2002)*, pp. 297-304, Colmar, France, April 2002.
- [19] I. Constantiou, N. Mylonopoulos, "Towards sustainable quality of service in interconnection agreements: Implications from information asymmetry", *Proceedings of the 9<sup>th</sup> European Conference on Information Systems (ESIC)*, Bled, Slovenia, June 2001.
- [20] D. Verma, "Supporting Service Level Agreements on IP Networks", Macmillan Technical Publishing, pp. 19-20, 1999.
- [21] E. Marilly, O. Martinot, S. Betgé-Brezetz, G. Delégue, "Requirements for SLA Management", *Proceedings of the IEEE Workshop on IP Operation and Management (IPOM)*, pp. 57-62, Dallas, USA, October 2002.
- [22] D. Pappalardo, "WorldCom touts real-time apps monitoring", [www.nwfusion.com](http://www.nwfusion.com/news/2003/0127carrprioip.html), <http://www.nwfusion.com/news/2003/0127carrprioip.html>, 1/27/03.
- [23] D. Grossman, "New Terminology and Clarifications for DiffServ", *Network Working Group, RFC 3260*, April 2002.
- [24] D. Goderis, Y. T'joens, C. Jacquenet, G. Memenious, G. Pavlou, R. Egan, D. Griffin, P. Georgiadis, P. Van Heuven, "Service Level Specification Semantics, Parameters and negotiation requirements", *Internet draft*, draft-tequila-sls-01.txt, Expired December 2001.
- [25] S. Salsano, F. Ricciato, M. Winter, G. Eichler, A. Thomas, F. Fuenfstueck, T. Ziegler, C. Brandauer, "Definition and usage of SLs in the AQUILA consortium", *Internet draft*, draft-salsano-aquila-sls-00.txt, Expired May 2001.
- [26] Y. Bernet, P. Ford, R. Yavatkar, F. Baker, L. Zhang, M. Speer, R. Braden, B. Davie, J. Wroclawski, E. Felstaine, "A Framework for Integrated Services Operation over Diffserv Networks", *Network Working Group, RFC 2998*, November 2000.
- [27] R. Braden, D. Clark, S. Shenker, "Integrated Services in the Internet Architecture: an Overview", *Network Working Group, RFC 1633*, June 1994.
- [28] B. Braden, L. Zhang, S. Berson, S. Herzog, S. Jamin, "Resource ReSerVation Protocol (RSVP) – Version 1 Functional Specification", *Network Working Group, RFC 2205*, September 1997.

- [29] J. Wroclawski, "The Use of RSVP with IETF Integrated Services", *Network Working Group, RFC 2210*, September 1997.
- [30] H. de Meer, Piers o'Hanlon, G. Feher, N. Blefari-Melazzi, H. Tschofenig, G. Karagiannis, D. Partain, V. Rexhepi, L. Westberg, "Analysis of Extending QoS Solutions", *Internet draft*, draft-demeer-nsis-analysis-03.txt, Expired May 2003.
- [31] V. Fineberg, "A Practical Architecture for Implementing End-to-End QoS in an IP Network", *IEEE Communications Magazine*, pp. 122-130, January 2002.
- [32] F. Le Faucheur, L. Wu, B. Davie, S. Davari, P. Vaananen, R. Krishnan, P. Cheval, J. Heinanen, "Multi-Protocol Label Switching (MPLS)", *Network Working Group, RFC 3270*, May 2002.
- [33] V. Alwayn, "Advanced MPLS Design and Implementation", Cisco Press, pp. 313-315, 2002.
- [34] V. Fineberg, C. Chen, X. Xiao, "An End-to-End QoS Architecture with the MPLS-Based Core", *Proceedings of IP Operation and Management (IPOM)*, pp. 26-30, Dallas, USA, October 2002.
- [35] T. Ferrari, T. Chown, N. Simar, R. Sabatino, S. Venaas, S. Leinen, "Experiments with Less than Best Effort (LBE) Quality of Service", *Proceedings of the 9<sup>th</sup> GEANT Task Force on Next Generation Networking (TF-NGN)*, October 2002.
- [36] R. Bless, K. Wehrle, "A Lower Than Best-Effort Per-Hop Behavior", *Internet draft*, draft-bless-diffserv-lbe-phb-00, expired September 1999.
- [37] R. Bless, K. Nichols, K. Wehrle, "A Lower Effort Per-Domain Behavior for Differentiated Services", *Internet draft*, draft-bless-diffserv-pdb-le-01, Expired May 2002.
- [38] TF-NGN LBE WG, <http://www.cnaf.infn.it/~ferrari/tfngn/lbe>, January 2003
- [39] QBone Scavenger Service, <http://qbone.internet2.edu/qbss>, January 2003
- [40] Funet-verkko, <http://www.funet.fi>, April 2003.
- [41] Abilene network, <http://abilene.internet2.edu>, April 2003.
- [42] D. Verma, "Supporting Service Level Agreements on IP Networks", Macmillan Technical Publishing, pp. 21-28, 1999.
- [43] I. Cselenyi, N. Borg, C. Hatch, H. Gharib, D. Milham, P. Schmid, A. Sumesgutner, P. Haraszti, L. Fiard, H. Wang, V. Cruz, S. Escudero, "Inter-operator interfaces for ensuring end to end IP QoS: Measurements of Performance Metrics and Service Events", *EURESCOM Technical Report*, P1008, Part 2 of Deliverable 3, May 2001.
- [44] *ITU-T recommendation G.114*, "One-way transmission time", February 1966.
- [45] I. Matta, L. Guo, "Differentiated Predictive Fair Service for TCP Flows", *Proceedings of the 8<sup>th</sup> IEEE International Conference on Network Protocols (ICNP)*, pp. 49-58, Osaka, Japan, November 2000.
- [46] K. Tokuda, G. Hasegawa, M. Murata, "Analysis and Improvement of the Fairness between Long-lived and Short-lived TCP Connections",

- Proceedings of the 7<sup>th</sup> the International Workshop on Protocols for High-Speed Networks (PfHSN'02), Berlin, April 2002
- [47] D. Allen, "Qwest Introduces On- and Off-Net SLAs", *Network Magazine*, <http://www.networkmagazine.com/article/NMG20020104S00060>, 1/07/02.
- [48] D. Pappalardo, M. Martin, "Broadwing raises the bar on SLA's", *Network World*, <http://www.nwfusion.com/news/2002/0218broadwing.html>, 02/18/02.
- [49] K. Nichols, B. Carpentier, "Definition of Differentiated Services Per Domain Behaviors and Rules for their Specification", *Network Working Group, RFC3086*, April 2001.
- [50] V. Jacobson, K. Nichols, K. Poduri, "The 'Virtual Wire' Per-Domain Behavior", *Internet draft*, draft-ietf-diffserv-pdb-vw-00.txt, Expired in January 2001.
- [51] N. Seddigh, B. Nandy, J. Heinanen, "An Assured Rate Per-Domain Behavior for Differentiated Services", *Internet draft*, draft-ietf-diffserv-pdb-ar-00.txt, Expired August 2001.
- [52] M. Brunner, A. Banchs, S. Tartarelli, H. Pan, "An one-to-any Assured Rate Per-Domain Behavior for Differentiated Services", *Internet draft*, draft-brunner-diffserv-pdb-one2any-ar-00.txt, Expired August 2001.
- [53] B. Moore, E. Ellesson, J. Strassner, A. Westerinen, "Policy Core Information Model", *Network Working Group, RFC 3060*, February 2001.
- [54] A. Westerinen, J. Schnizlein, J. Strassner, M. Scherling, B. Quinn, S. Herzog, A. Huynh, M. Carlson, J. Perry, S. Waldbusser, "Terminology for Policy-Based Management", *Network Working Group, RFC 3198*, November 2001.
- [55] D. Durham, J. Boyle, R. Cohen, S. Herzog, R. Rajan, A. Sastry, "The COPS (Common Open Policy Service) Protocol", *Network Working Group, RFC 2748*, January 2000.
- [56] K. Nichols, V. Jacobson, L. Zhang, "A Two-bit Differentiated Services Architecture for the Internet", *Network Working Group, RFC 2638*, July 1999.
- [57] P. Appan, Y. Lingjia, P. Radhakrishna, "A Policy Based QoS Management System for the IntServ/DiffServ Based Internet", *Proceedings of the 3<sup>rd</sup> International Workshop on Policies for Distributed Systems and Networks, (POLICY'02)*, 2002.
- [58] S. Giordano, S. Salsano, S. Van den Berghe, G. Ventre, D. Giannakopoulos, "Advanced QoS Procioning in IP Networks: The European Premium IP Projects", *IEEE Communications Magazine*, pp. 30-36, January 2003.
- [59] E. Mier, V. Battistelli, A. Miner, "Sizing up VoIP listening tools", *Network World*, <http://www.nwfusion.com/research/2002/1209voipfeat.html>, 12/09/02.
- [60] Performance Tools Taxonomy, <http://www.caida.org/tools/taxonomy/performance.xml>, 3/5/03.

- [61] Network monitoring tools, <http://www.slac.stanford.edu/xorg/nmtf/nmtf-tools.html>, 3/5/03.
- [62] C. Griffin, G. Goddard, "Searching for the QoS Holy Grail", *Network World*, <http://www.nwfusion.com/reviews/2002/0603rev.html>, 06/03/02
- [63] Packeteer's homepage, <http://www.packeteer.com/products/packetshaper.cfm>, 5/14/03.
- [64] Staselog's homepage, <http://www.staselog.com> 5/7/03.
- [65] R. Braden, "Requirements for Internet Hosts - Communication Layers", *Network Working Group, RFC 1122*, October 1989.
- [66] N. Seddigh, P. Piedad, "Study of TCP and UDP Interaction for the AF PHB", *Internet draft*, draft-nsbnpp-diffserv-tcpudpaf-01.pdf, September 1999.
- [67] B. Nandy, N. Seddigh, P. Piedad, J. Ethridge, "Intelligent Traffic Conditioners for Assured Forwarding Based Differentiated Services Networks", *Proceedings of IFIP High Performance Networking*, Paris, June 2000.
- [68] N. Seddigh, B. Nandy, P. Piedad, "Bandwidth Assurance Issues for TCP flows in a Differentiated Services Network", *Proceedings of Global Internet Symposium, (GLOBECOM'99)*, Vol. 3, pp. 1792-1798, Rio De Janeiro, December 1999.
- [69] I. Guo, I. Matta, "The War Between Mice and Elephants", *Proceedings of the 9<sup>th</sup> IEEE International Conference on Network Protocols (ICNP)*, Riverside, California, November 2001.
- [70] V. Laatu, J. Harju, P. Loula, "Evaluating Performance Among Different TCP Flows in a Differentiated Services Enabled Network", *Proceedings of the 10th International Conference on Telecommunications (ICT'2003)*, Vol. 1, pp. 709-715, Tahiti, Papeete, French Polynesia, February 2003.
- [71] N. Brownlee, KC Claffy, "Understanding Internet Traffic Streams: Dragonflies and Tortoises", *IEEE Communications Magazine*, pp. 110-117, October 2002.
- [72] R. Serban, C. Barakat, W. Dabbous, "Dynamic Resource Allocation in Core Routers of a DiffServ Network", *Proceedings of the Asian Computing Science Conference (ASIAN'02)*, pp. 153-167, Hanoi, Vietnam, December 2002.