CHAPTER 3

Differentiated Services Working Group

This chapter provides an overview of the evolution of Differentiated Services. A short history is provided to illustrate the main ideas that formed the basis for the Differentiated Services Working Group.

Most of this chapter discusses the principle achievements of the first year of the working group. The first two documents reached the RFC status in December 1998: RFC 2474, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers," and RFC 2475, "An Architecture for Differentiated Services." Those documents describe the theoretical foundation of a Differentiated Services network. The third document reviewed in this chapter is "A Framework for Differentiated Services" (Bernet *et al.* 1998), which supplements the view of the two RFCs by giving more practical guidance for the application Differentiated Services.

3.1 A Short History of Differentiated Services

This brief historical overview is based primarily on the mailing list discussion of the Integrated Services working group. The mailing list archive offers useful documentation of the Differentiated Services effort before the establishment of the Differentiated Service Working Group and the corresponding mailing list in February 1998. (These mailing lists are available at `http://www.ietf.org/html.charters/diffserv-charter.html` and `http://www-nrg.ee.lbl.gov/diff-serv-arch/`, respectively.) It can even be safely claimed that that the Differentiated Services activity is a direct extension of the effort made by the Integrated Services and RSVP Working Groups, because the goal of all these efforts is basically the same: the provision of service differentiation in IP networks.

It should be stressed, however, that this introduction provides only one view of the complex development process of Differentiated Services. The End-to-End research group discussed some of the same issues several years before the establishment of Differentiated

Services, for instance (The Internet Research Task Force). This overview tries to disclose the chief motivation behind the effort to develop Differentiated Services by picking some occurrences during the evolutionary process that finally led to the establishment of the Differentiated Services Working Group.

The first mention of differentiation took place in a mailing in November 1995; Mark Garrett mentioned a claim that ATM was the first network technology to have meaningfully differentiated QoS (in contrast to X.25 and ISDN). Note that although the verity of this statement is arguable, the main point here is that until then there was no mention at all of differentiation on the mailing list. The prevalent view seemed to be that real QoS required hard guarantees, and without them there was no QoS at all. (Not everyone held this view, however.)

After that, the general discussion about service differentiation seems to have been buried in the standardization effort of Integrated Services and RSVP. The concerns about the scalability of RSVP gradually increased, however. Consequently, preliminary arguments that something simpler was needed to obtain better scalability began to be bantered about.

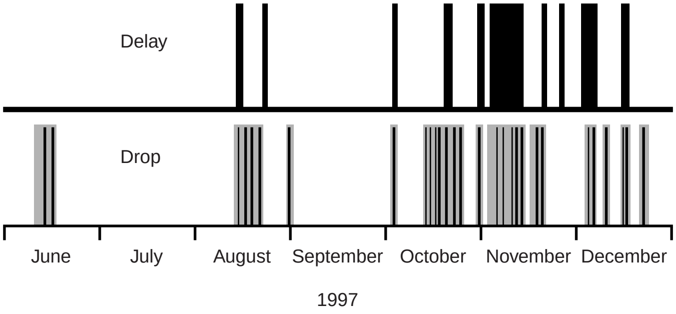
The first culmination point was the "Future Directions for Differential Services" BOF session arranged at the IETF meeting in April 1997. The final remarks concluded that the taxonomy of Differentiated Services could include three options (Mankin 1998):

- Service with extremely high predictability
- Two services, one cheap and one expensive; interoperability issues integral
- A lot of finely differentiated services (although apparently this was not a popular idea at the meeting)

A couple of weeks after the meeting, the real discussion started about the basis for a new Differentiated Services model. Figure 3.1 depicts the nature of the discussion. The figure shows the days in which there were messages in the Integrated Services mailing list related to two topics: the need of drop precedence levels (the terms *drop preference* and *drop priority* have also been used) and delay classes (or *delay categories*).

The main reason for presenting Figure 3.1 is to illustrate the emergence of these fundamental issues. Better delay characteristics are required in the future Internet because of real-time applications, and Differentiated Services must provide those characteristics. The concept of drop precedence seems to be more controversial. Drop precedence basically provides a tool with which you can inform the network about the relative importance of packets. It is still somewhat unclear, however, how drop precedence bits should be actually applied. Differentiated Services attempts to give one answer to this issue.

Figure 3.1 Discussion activity on the Integrated Services mailing list.



The following three general observations can be made:

- This *traffic* process is very similar to that of any Internet traffic process, with intermittent bursts of activity and relatively long idle periods.
- Traffic streams can be highly correlated. In the case of Figure 3.1, a possible reason for correlation is that both threads of discussion have basically the same objective to specify the fundamental architecture of Differentiated Services. Some of the aspects of both issues could be discussed separately, but several aspects of the issues do indeed overlap.
- The third observation relates to the general behavior of mailing list discussions. The same threads of discussions emerge repeatedly at arbitrary points of time, which makes following the discussion somewhat difficult.

Moreover, by calculating the number of mailings within one day and the length of the mailings, a *self-similar* process results. This correlation between discussions and Internet traffic is not so far-fetched as it might seem at first sight: Internet traffic reflects complex human processes, largely in the same way as does a discussion in a mailing list. The lesson of this brief exercise is that controlling Internet traffic could be as difficult as controlling the discussion in the mailing lists.

The seven months from June to December were crucial for the emergence of Differentiated Services. In August, IETF had a meeting in which the subject of service differentiation was discussed in the Integrated Services Working Group session (Wroclawski 1997). Several presentations related to different service models were made. The presentations and ensuing discussion revealed the basic arguments as well as the basic factions. The following principle factions were obvious:

- The scalability of RSVP is the problem to be solved.
- Highly reliable IP service is the key target.

- The main tool is several drop priority levels.
- Low-delay service is also important.
- The core network mechanism should allow the implementation of any imaginable service.

This list may appear somewhat disordered: The list includes some target services, tools to meet the targets, and more general objectives. This list illustrates the reality of the mailing list—with diverse objectives and viewpoints. A systematic and coherent list would be "spin", because the reality has been unsystematic.

A lively discussion emerged after the IETF meeting in August, as shown in Figure 3.1. The most controversial issue among the many topics was whether Differentiated Services should be based on drop preferences. This particular controversy crystallized into two opinions expressed in the mailing list [Int-Serv]:

> …I really like the idea of using the IP precedence bit field, allowing eight (0–7) levels of distinction, so that something similar to WRED can provide for differentiated drop in the core.
>
> —Paul Ferguson, 17 August 1997

> Support for drop preference creates an undesirable incentive for applications to send packets that will not reach their destinations (knowing that the routers will discard the "less important" or "out of profile" packets at a point of congestion).
>
> —Steve Deering, 15 October 1997

This division was detectable also in the five Internet drafts that were submitted in November 1997.

Despite the diversity of the proposals, the activity clearly indicated that there was significant support for Differentiated Services in the Internet community, although the service structure was still an open issue. (Another important thing to note is that every IETF session related to Differentiated Services has been very crowded.) After the public discussions at the IETF meeting in December (Wroclawski 1997), a smaller group took over the reins of the Differentiated Services effort, mainly to attain a compromise that would satisfy the various needs of the different parties. The IESG approved the working group on 26 February 1998.

3.2 The Position of the Differentiated Services Working Group

The goal of Differentiated Services is clearly explained in the description of the working group that was published on March 1, 1998 (available at `http://www.ietf.org/html.charters/diffserv-charter.html`). It emphasizes the need for simple, but versatile methods of providing service differentiation. The essence of the approach is that *a small set of building blocks is defined and services are built from those blocks.* Two building blocks are explicitly mentioned: DS byte and Per-Hop Behavior (PHB). Further, interoperability is required to enable the provision of reasonable end-to-end services.

The original description identifies the need for two documents: a standard track document and an informational document. The standard track document defines the general use of the DS byte. In addition, it standardizes a couple of *codepoints* that are applied to Per-Hop Behaviors that are commonly used in the current Internet. (See the section "Terminology" later in this chapter for definitions of key terms.) The objective of the informational framework document is to define architecture and common language for Differentiated Services.

An additional goal of the working group was to experiment with different Per-Hop Behaviors. After successful experiments, these Per-Hop Behaviors could be specified in experimental RFCs, or they could become standardized. Although the main building blocks were clearly DS byte and PHBs, one goal of the working group was to investigate other components required to build services, such as traffic shapers and packet markers. Moreover, as a general requirement for any IETF standard, security issues had to be analyzed. Finally, the working description declared two issues to be beyond of scope of the working group: mechanisms for the identification of individual traffic flows within the network, and signaling mechanisms to support the marking of packets.

The description defined a tight timetable for the working group. It gave 10 months to prepare documents for the basic DS standard, framework, boundary mechanisms, and traffic conditioners. Although the working group was not able to totally comply with the timetable, it made good progress in 1998.

3.3 Basic Working Group Documents

The Differentiated Services Working Group prepared three main documents during 1998: RFC 2474, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers," RFC 2475, "An Architecture for Differentiated Services," and an Internet draft, "A Framework for Differentiated Services" (Bernet *et al.* 1998). There is a clear difference between the status of the first two and the last of these documents.

Internet Drafts Versus Request for Comments

Internet drafts are working documents, and therefore do not have any official status at all—in fact, they can be removed or replaced at any time by a more recent version of the same document. Consequently, a mere Internet draft is not a published document, which makes it questionable to use them as references.

Unfortunately, some significant documents are not available in any other form than Internet drafts, and therefore are used in this book as a reference—or rather as a pointer to a text (but this is avoided when possible).

Request for Comments (RFCs) are much more stable. There are two special subseries within the RFCs (Bradner 1996):

- *Standard track*: These RFCs may reach an Internet standard status after there is enough evidence that the standard is appropriate for practical use.

- *Nonstandard track*: These RFCs are either experimental or informational. They are supposed to provide useful information about the application of an Internet standard, but do not contain any strict requirements for implementation.

The DS Field document (RFC 2474) is a standard track document; as a result, it is the most important document made by the Differentiated Services working group (Baker *et al*. 1998). However, it provides only limited guidance as to how to apply the building blocks defined in the document. Therefore, the architecture document, RFC 2475, is also essential even though it is not a standard track RFC. It is fair to say that RFC 2475 reflects the opinion of a large group within the Differentiated Services Working Group. The viewpoints of this book and the architecture document are mostly congruent, although there are a couple of differences. (See the section "Provisioning and Configuration" later in this chapter for some of the differences in viewpoint.)

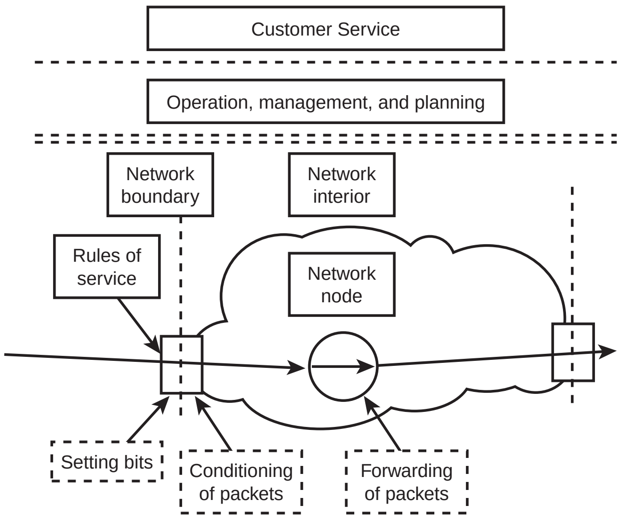
The framework document did not reach a RFC status by the end of 1998 (Bernet *et al*. 1998). A revised version will be submitted to the RFC editor as a Proposed Standard for the Internet Community probably in 1999. Nevertheless, the document was prepared by some of the key authors in this field, and it is also a kind of working group draft, which means that it has a certain level of acceptance among the working group. This book and the framework document disagree on a couple of fundamental issues. Those differences are explained in Chapter 4, "General Framework for Differentiated Services."

3.3.1 Introduction to Differentiated Services Model

The abstract of the RFC 2474 clearly states one of the most important characteristics of the Differentiated Services model: There is a definite distinction between boundary functions and interior functions, or boundary nodes and interior nodes. Boundary nodes are responsible of the setting of DS bits in the packet, and of the conditioning of packets; interior nodes, on the other hand, forward packets in different ways based on the value of the DS field.

Service providers can build different network services from the main building blocks: setting of bits, conditioning of packets, and forwarding of packets. Without any common rules, however, there is no consistent service. The rules define how bits are set and how packets are conditioned at the boundaries and forwarded inside the network. The actual substance of the service is the system, consisting of the rules and of the proper design of the building blocks. Moreover, the importance of operation, management, and planning of the system cannot be overemphasized. Figure 3.2 shows these elements of service provision.

Figure 3.2 Construction of Differentiated Services.



Because the service construction as such is beyond the scope of IETF, IETF documents do not usually give detailed examples of how to apply the specification for business purposes. This book attempts to satisfy the evident need for examples.

The following example, "Real-Time Service Versus Best-Effort Service," illustrates a realistic scenario in which the fictional ISP Fairprofit can improve the network service with a quite simple system that is in accordance with RFC 2474. It should be stressed, however, that the example is only for illustration purposes, and any of the details mentioned here can be realized in different ways. (Note that Fairprofit was used in the section titled "Fairness and Service Provision" in Chapter 1, "The Target of Differentiated Services," to assess the fairness issues related to overloads.)

Real-Time Service Versus Best-Effort Service

The service provider Fairprofit wants to offer two types of service: real-time service for interactive applications and best-effort service for data applications. First of all, Fairprofit must define the basic rules to be applied:

- Customers are allowed to send IP packets into the network, basically with any bit rate, limited only by the physical access rate.
- At the network boundary, packets are marked as real-time packets based on the information in the IP header; that is, the marking divides the traffic into real-time and data substreams.
- If the data traffic sent by the customer exceeds a certain limit, packets are marked as lower importance, but are sent into the network.
- Every user is allowed to send real-time traffic at a certain bit rate, and all excessive packets are discarded at the network boundary.

The system for setting bits requires that the boundary node make a proper decision about what packets should be marked as real-time packets. Traffic conditioning means that the boundary node measures the bit rate of both real-time and data traffic streams, and then according to the result can either re-mark a data packet with lower importance or discard a real-time packet.

The forwarding system inside the network then treats the packet according to the marking. There are three groups of packets: real-time, data with normal importance, and data with low importance. Real-time packets have their own queue that is served before the queue for data packets, to achieve delay differentiation. During overload situations—that is, when the occupancy level of the data queue (or real-time queue) is very high—data packets with low importance are discarded.

The main task of the operation and management system is to keep the real-time traffic low enough to guarantee that real-time queue is not filled, and that there is enough bandwidth for data traffic. In addition, the traffic level of data packets with normal importance should be low enough to keep the packet-loss ratio for those packets very low.

Finally, even though this service model does not necessarily require any changes made by the customers, the Fairprofit Corporation has to be able to explain why the service model is better than the current best-effort model, and how the customer can use the new service offering in the best manner. Their brochure could include such statements as better suitability for interactive applications and relatively high assurance that certain minimum bandwidth is always available for data services. The professionals of Fairprofit can then expect that customers will be ready to pay for these service characteristics.

3.3.2 The Differentiated Services Field in IPv4 and IPv6 Headers

The title of RFC 2474, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers," indicates that the main issue of the document is to specify the contents and meaning of the DS field. In addition, it provides useful information about the terminology and basic structure of Differentiated Services.

Terminology

RFC 2474 defines the basic terminology of Differentiated Services. Although it is impractical to explain all the terms in different words, some additional explanations of the key terms can be useful:

- Per-Hop Behavior
- Customer service
- Network service
- PHB class
- Codepoint
- Mechanism

The term *Per-Hop Behavior (PHB)* is both difficult to comprehend and important for understanding the whole idea of Differentiated Services. Technically speaking, PHB denotes a combination of forwarding, classification, scheduling, and drop behaviors at each hop. However, PHB is not only a technical concept; instead, the main purpose of PHB is to make a comprehensible connection between packet-level implementations and service models. PHB is, in a way, an intermediary term.

Based on the wording of RFC 2474, it is possible to derive the following guidelines for designing a Per-Hop Behavior. (Note that this is only an interpretation of the formal standard from the viewpoint of this book.)

- PHB is primarily a description of desired behavior on a relatively high abstraction level; in particular, a PHB must have a comprehensible motivation.
- PHB should allow the construction of predictable services.
- The desired behavior should be externally observable—for instance, the description of behavior should not use any internal terms, such as it.
- The desired behavior should be local—that is, it should concern behavior within one node rather than the whole network.
- The description of behavior is related to an aggregate that consists of all packets belonging to the same PHB in a certain point of the network.
- The packets belonging to a PHB should experience the same treatment independent of other information in the packet and independent of the traffic process of individual flow inside the aggregate.
- The PHB description should not suppose any particular conditioning function at the network boundary.

Consequently, the first two items together specify the general target of PHBs: They should provide meaningful basis for understanding the behavior of the Differentiated Services system. The other items limit the terms that can be used to describe the desired behavior. Figure 3.3 shows a simplified model for PHB specification that concerns the treatment of an aggregate stream inside a black box—that is, an interior node in a Differentiated Services network.

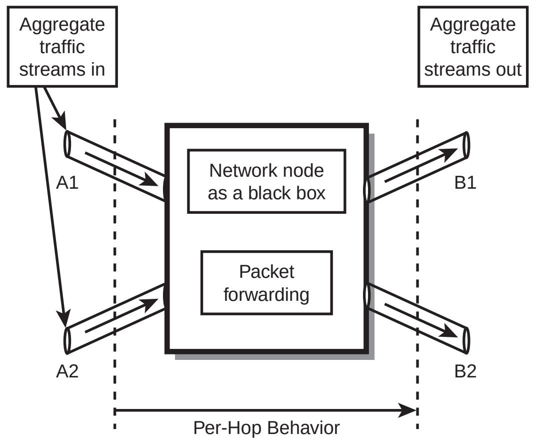
As to the last, somewhat arguable item, there are two opposite needs:

- To keep each PHB as multipurpose as possible
- To make it possible to design predictable end-to-end services

The view adapted in this book is that the last instruction item should always be applied when there is not a compelling reason to bind a certain PHB to a specific traffic-conditioning function. Note that the architecture document (RFC 2475) declares explicitly that Differentiated Services architecture should *decouple* traffic conditioning and service provisioning functions from forwarding behaviors.

The main reasoning behind this decoupling, or separation, of traffic conditioning and forwarding behaviors is flexibility. After the core network system is specified and the applications of core network functions have been established, the service evolution can continue by inventing new traffic-conditioning functions for boundary nodes.

Figure 3.3 Per-Hop Behavior (PHB).



Even though not directly stated in RFC 2474, one permissible approach is to divide the whole network capacity into several parts in a static manner, in a way that each part is operated separately as an independent network although all packets use the same transmission resources. In that case, the PHB description concerns one individual part of the network resources rather than the whole network.

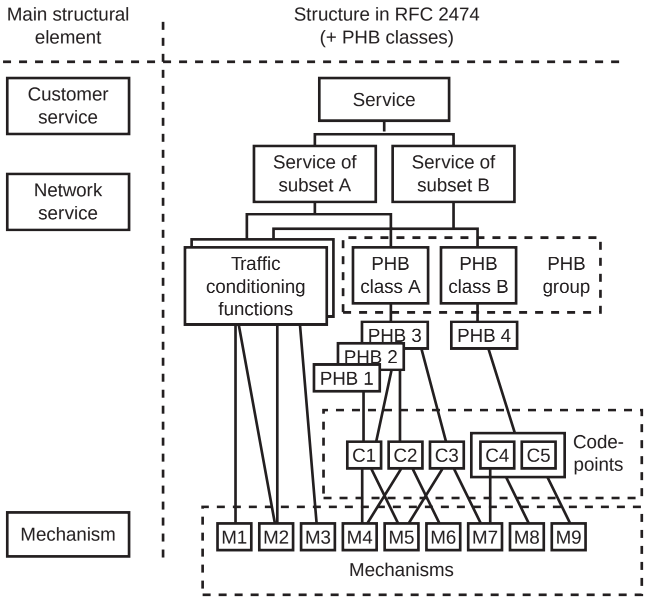
Chapter 1 introduced two service concepts: customer service and network service. The definition of service in RFC 2474 can refer to either of them. *Customer service* is a description of the overall treatment of a customer's traffic (including with other aspects such as pricing); *network service* refers mainly to a subset of a customer's traffic.

Now it's time to introduce one supplementary term not used in RFC 2474—a *class*, or more accurately, a *PHB class*. This term has been used in some important PHB specifications (Baker *et al.* 1998). A PHB class is a collection of PHBs intended to be applicable for transmitting packets of one application. Technically this means that the service provider is allowed to re-mark packets within a PHB class, but not from one class to another class. The main requirement for a PHB class is that packets should not be reordered inside the network. A PHB class with proper traffic-conditioning functions at the network boundary is the nearest equivalent for the network services in connection-oriented networks, such as ATM networks.

Codepoints are the handles used to inform inside nodes about the PHB of the packet. The fundamental requirement is that the codepoint of the packet unambiguously define the PHB. On the contrary, several different codepoints can map to the same PHB, which means that an aggregate can consists of packets with different codepoints. In that case, the treatment should be the same within one PHB regardless of the actual codepoint used in the packet.

According to RFC 2474, a *mechanism* is the implementation of one or more Per-Hop Behaviors according to a particular algorithm. A mechanism can be used for implementing several PHBs, and several mechanisms are usually needed to implement a PHB. Figure 3.4 depicts the total picture of this Differentiated Services structure. (See Chapter 1 for a description of the main structure of Differentiated Services.) The following example, "Implementing Real-Time Service and Data Service," illustrates using PHBs to effect services.

Figure 3.4 The main building blocks of Differentiated Services.



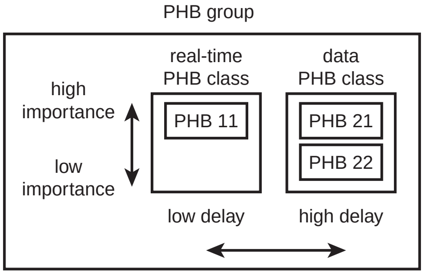
Implementing Real-Time Service and Data Service

The fictional service provider Fairprofit can implement the services introduced in the preceding example, "Real-Time Service Versus Best-Effort Service," by means of one PHB group consisting of two PHB classes: one for real-time service and another one for data service. The real-time service consists of one PHB (11); whereas the data PHB class consists of two PHBs (21 and 22) with different importance. Figure 3.5 shows this PHB structure.

We may ask whether the two classes should belong to a PHB group. This issue is not totally clear; for management purposes, however, it is certainly useful to define the relationship between the two classes and effectively, to make a PHB group.

Hence, as to the PHB description, it could be enough to say that real-time aggregate is served with shorter delay than the data aggregate and that packets marked with PHB21 encounter smaller discarding probability than PHB22 packets. The actual performance of end-to-end service depends then crucially on traffic-conditioning functions at the network boundaries. In general, it is not feasible to define exact performance values for PHB because PHB as such does not cover traffic-conditioning functions at network boundaries.

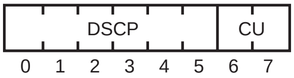
Figure 3.5 Per-Hop Behavior group with two PHB classes.



DS Field Definition

The PHB information is transmitted in the Differentiated Services field (DS field) that consists of an octet IP header. Section 3 in RFC 2474 defines the structure of this field, shown in Figure 3.6. This new definition replaces the older definitions of the IPv4 TOS octet and the IPv6 Traffic Class octet. The DS field is divided into two parts: the first six bits (DSCP field) are used as a codepoint to select the PHB, and the two last bits (CU bits) are reserved for future use.

Figure 3.6 The structure of the Differentiated Services field.



A DS-compliant node uses all the six bits of the DSCP field, but not any other bits to select the PHB. Further, RFC 2474 requires that the implementation of codepoint mapping should be very flexible: The field should be treated as an index without any internal structure, and the operator should be able to map any codepoint to any PHB.

Historical Codepoint Definitions

If a packet with an unrecognized codepoint is received, it should be forwarded according to the default behavior. Therefore, the default PHB must be available in every DS node. This default behavior corresponds to the best-effort service provided in current networks—that is, the network tries to deliver as many packets as possible and as soon as possible. The other part of the service (although not formally a part of the PHB) is the traffic conditioning functions. Although according to the conventional best-effort paradigm there is no particular traffic conditioning at the network boundary, in a Differentiated Services network some conditioning is possible, but not mandatory, for default PHB traffic.

The main issue related to the default PHB is the relationship between it and other PHBs used in the network. In general, the operator may apply any kind of policy in this respect. It is recommended that the service policy ensure that default PHB always gets reasonable amount of resources, however, regardless of the other PHBs. This allows a smooth coexistence of both Differentiated Services–aware and non-aware traffic streams in the same network. Because of this backward compatibility issue, the codepoint 000000 must map to the default PHB (or to another PHB with similar characteristics).

In addition to the default behavior, RFC 2474 defines a PHB group called *class selector PHB* with defined codepoints. The reason for this standardization is that the IP Precedence field defined in RFC 791 has been used in some real networks. The specification of this PHB group is described further in RFC 2474 and also in Chapter 7, “Per-Hop Behavior Groups,” in section 7.2, “Class Selector PHB Group,” of this book.

Per-Hop Behavior Standardization Guidelines

RFC 2474 provides some guidelines for those who write PHB specifications. First, according to the common rule applicable to any IETF specification, implementation, deployment, and proven usefulness are prerequisites for any PHB to be standardized. Because no mechanisms are standardized, vendors can use any appropriate mechanisms that together satisfy the definition of a PHB, PHB class, or PHB group.

It is assumed that certain common Per-Hop Behaviors will evolve in such a way that an established set of services will emerge. Although this is a likely scenario, it is premature to predict how the PHB field will evolve and which one of the PHB proposals will be commonly used and which will vanish.

Although the DSCP field is in principle unstructured, three different pools are introduced in Chapter 6 of RFC 2474, mainly for codepoint-management purposes. The first pool with codepoints xxxxx0 is for standard action, the second pool with codepoints xxxx11 is for experimental and local use, and the third pool with codepoints xxxx10 is initially for experimental use, but may be later used for standard PHBs.

3.3.3 Architecture for Differentiated Services

RFC 2475, “An Architecture for Differentiated Services,” defines the architecture for implementing scalable service differentiation on the basis of the DS field specification (Black *et al.* 1998). Because the architecture itself is not a matter of standardization, this architecture document is an informational RFC. This document both further clarifies some issues addressed in the standard track RFC 2474 and, of course, discusses architectural issues.

Terms and Targets

The introduction section of RFC 2475 further illuminates the terminology of Differentiated Services. It states that service characteristics may be specified in terms of *throughput*, *delay*, *jitter*, *loss*, or *relative priority* of access to network resources. This list is actually a central tool for the development of PHBs: The differentiation made possible by a PHB should concern some of the essential service characteristics.

The introduction section also provides a quite comprehensive and useful list of terms. The basic architecture of Differentiated Services is described by a list of requirements. The main points of the list are as follows:

- *Versatility*: A wide variety of end-to-end services should be possible to realize; network services should be independent of applications, and they should be directly applicable with current applications and with current network services.
- *Simplicity*: The overall system or parts of it should not depend on signaling for individual flows; only a small set of forwarding behaviors should be necessary.
- *Cost efficiency*: Information about individual flows or customers should not be used in core nodes, but only states of aggregate streams should be used in core nodes.

As you can see, part of the terminology presented in Chapter 1 is also used here. A similar list of terminology also appears in Chapter 4.

Comparison with Other Approaches

The introduction section of RFC 2475 also provides a concise overview of other approaches for service differentiation. Although the basic arguments are similar to those in Chapter 2, "Traffic Management Before Differentiated Services," it is worthwhile to make an overview to explain the target of the Differentiated Services Working Group (which might differ slightly from that of this book).

The categories applied in the comparison are relative priority marking, service marking, label switching, Integrated Services / RSVP, and static per-hop classification. In the priority-marking approach, the application or some other entity selects a relative priority for each packet, and the network nodes use it to decide which kind of forwarding behavior should be applied to the packet. Differentiated Services can be considered a refinement of this model.

The difference between priority marking and service marking is subtle. In the service-marking approach, the required end-to-end service is more definitely expressed—for instance, "minimize delay" or "maximize reliability." This information is not only used to select the forwarding behavior, but also the route; the Differentiated Services field, on the other hand, is not particularly intended for route selection. Because the Differentiated Services approach leaves as much space as possible for further evolution, it is considered that the possible services are not built-in parts of the Differentiated Services structure. (You may deem this self-contradictory; Differentiated Services really does not consider services, but only building blocks for services.)

The label-switching (or virtual-circuit) model includes Frame Relay, ATM, and MPLS. In this model, the granularity of resource allocation can vary from individual flows to large aggregate streams. The cost of the fine granularity is the complex management and configuration needed to establish and maintain all the information related to the large number of flows. The main difference of Integrated Services / RSVP model compared to label switching is that it relies on traditional packet forwarding as the underlying technology. The additional element of RSVP is that it allows sources and receivers to inform network nodes about the needs of applications and make reservation through the network.

The main problem of the RSVP model is scalability in high-speed core routers. Therefore, it is supposed that by using Differentiated Services in the core network, the scalability problems of RSVP can be avoided; in the access network, however, RSVP can be used to make definite reservations. Furthermore, different combinations of technologies and service models are needed in practical implementations. If ATM is used as the underlying technology for Differentiated Services, for example, the end-to-end service could be a compromise between the ATM service model and Differentiated Services model.

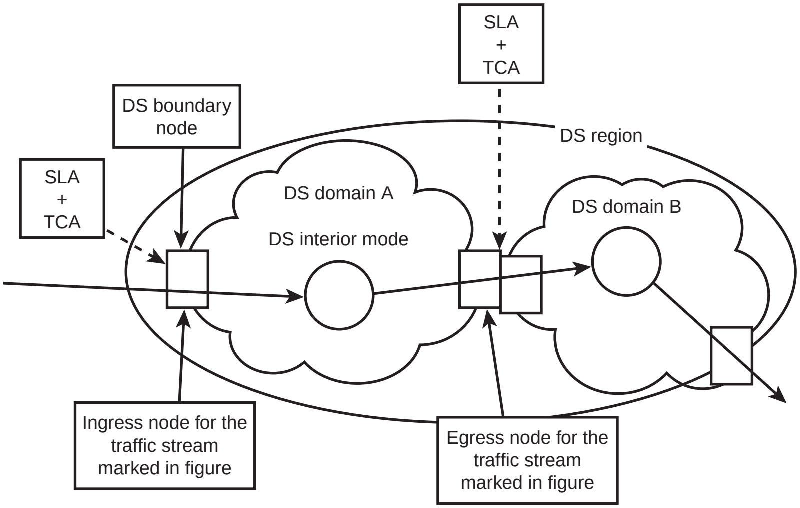
Architecture Model

The basic elements of the architecture model are explained thoroughly in the second chapter of RFC 2475, “Architecture for Differentiated Service.” The key elements for building Differentiated Services are DS boundary nodes, DS interior nodes, ingress nodes, and egress nodes, as shown in Figure 3.7. These elements are virtual in the sense that one physical node may contain all characteristics of all node types. It can be said that each type of node is a collection of characteristics:

- *Boundary node*: A collection of functions needed to interconnect a DS domain to another DS domain or to a non-DS–capable domain
- *Interior node*: A collection of functions needed if a node is connected only to other DS-capable nodes
- *Ingress node*: A collection of functions needed to handle incoming traffic streams to a DS domain
- *Egress node*: A collection of functions needed to handle outgoing traffic streams from a DS domain

In particular, a real boundary node can (and typically does) contain all these functions: The same node can be a boundary node for some traffic stream and an interior node for some other streams. Moreover, any interior node can have part of the functions of boundary nodes—for instance, the interior node may have a limited capacity of traffic conditioning.

Figure 3.7 Basic elements of a Differentiated Services network.



Further, it should be noted that there are two levels of agreements:

- *Service-level agreement (SLA)*: A contract between a customer and a service provider that specifies the forwarding service
- *Traffic-conditioning agreement (TCA)*: Defines the rules used to realize the service, such as metering, marking, and discarding

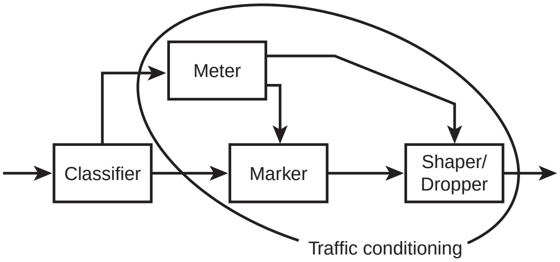
The same SLA and TCA concepts are also applicable between two network domains. SLA seems to be rather a customer service and / or network service concept; TCA, on the other hand, should be defined by terms that belong to the traffic-handling level (compare to Figure 1.1 in Chapter 1).

A DS domain is a part of a network in which nodes are DS compliant and in which a set of PHB groups are applied based on the same service-provisioning policy. At interfaces where PHB structure and / or the service policy is otherwise significantly changed, a boundary node is needed to make appropriate mappings between PHBs. A contiguous set of DS domains forms a DS region. Differentiated Services can be provided over a DS region, although significant differences in service structure and PHBs make it difficult to design and provide useful end-to-end services.

Traffic Classification and Conditioning

Figure 3.8 presents the logical structure of traffic classification and conditioning functions. This structure is based on the assumption that classification is made according to the information in the packet header (such as source address and destination addresses and DS field) and the incoming interface (RFC 2474). This model seems to exclude the possibility that traffic metering has any effect on the classification. This is a feasible approach if you suppose that the classification means the selection of a PHB class rather than an individual PHB. In such a case, marking is something done within the class, but does not cross the boundaries of PHB classes.

Figure 3.8 Packet classifier and traffic conditioning according to the architecture document.



A traffic profile is one way to present the traffic-conditioning rules. In the simplest model, each packet is either *in-profile* or *out-of-profile* based on the metering result at the arrival time of the packet. In-profile packets obtain better traffic-conditioning and forwarding treatment than out-of-profile packets. This model is further evaluated in this book in section 7.4, “Assured Forwarding PHB Group,” of Chapter 7.

According to the architecture document, the *meter element* measures each traffic stream. (*Stream* seems to correspond here to the concept of a PHB class, although the term is not used in this document.) Meter then informs the marker, shaper, and dropper mechanisms about the state of the stream:

- *Marker*: This sets an appropriate codepoint to the DS field of the packet. Actually this means that the marker is allowed to change the original value of the DS field. To avoid re-ordering within the network, the marker should comply with certain rules when re-marking packets.
- *Shapers*: These can be used to smooth the traffic process of particular aggregate streams.
- *Dropper mechanisms*: Based on the content of SLA and TCA, some packets can be discarded at the traffic-conditioning element.

Traffic conditioners are usually located in DS boundary nodes. The basic assumption is that all complicated classification and conditioning procedures are made near the source, while inside the network only some straightforward actions are necessary. Consequently, the most advisable approach could be that shaping is done in the private network, marking is done in boundary nodes, and dropping in interior nodes.

PHB as a Tool for Resource Allocation

According to RFC 2475, Per-Hop Behavior is the means by which a node allocates resources to aggregate streams; the given example is that a PHB gets *x*% of the capacity of a link. This definition satisfies the technical requirements of PHB specification (see the list in the beginning of this chapter); it is more difficult to comprehend how this PHB definition allows constructing predictable services, however. The core of the problem is that a larger capacity for an aggregate stream than for another one does not automatically produce better service because the final result depends crucially on the traffic load of each aggregate stream.

Further, the document states that PHB groups shall effectively partition the link and node resources between aggregates. However, it is not clear how this requirement can be deduced from the general requirements for PHB specification presented in the standard track document (RFC 2474) or from the desired service characteristics presented earlier in the architecture document (throughput, delay, jitter, loss, or relative priority; RFC 2475). Even though the aggregate bandwidth and throughput are similar terms, you can expect that end users are interested in the throughput of their own flow rather than the bandwidth of an aggregate stream.

Although the overall assessment depends on the interpretation of the word *effectively*, the idea that PHB essentially means resource allocation seems to be somewhat restricting and does not necessarily cover all Differentiated Services systems. Chapter 7, in section 7.4, “Assured Forwarding PHB Group,” examines this issue in more detail.

Per-Hop Behavior Guidelines

The architecture document, RFC 2475, gives 15 additional instructions for PHB designers. According to the first item, a PHB must include recommendation about the DS codepoint. Furthermore, PHB specification should include the following:

- An overview of the general purpose of the PHB group
- Specification of interactions between individual PHBs within a PHB group
- Provisioning restrictions if necessary (for example, whether the proper function of the PHB depends on the traffic-conditioning actions)

- A statement as to whether the PHB group is considered for general or local use
- A statement about the circumstances under which a packet can be re-marked within the PHB group from one PHB to another PHB

In addition, a PHB specification should discuss various issues such as interactions with previously defined PHBs, tunneling, conformance requirements, security, impacts on higher-layer protocols and link-layer mechanisms. A PHB specification that fulfils all the mentioned recommendations would certainly be very useful and comprehensible (unfortunately this not the case with all proposals).

Interoperability Issues

A non-DS–compliant node is a node that does not appropriately interpret the PHBs used within the DS domain. A specific case of a non-DS–compliant node is a so called legacy node that uses the first three precedence bits defined in RFC 791. Chapter 4 of RFC 2475 briefly considers interoperability issues, mainly in two specific cases. In the first case, a non-DS–compliant node is situated within a DS domain. One possibility to solve the apparent problems of this situation is to keep the overall load level so low that quality characteristics (delay and loss) are good enough for all aggregates. If that is not possible, one solution could be that the DS domain uses only class selector codepoints defined in RFC 2474.

An even more difficult case is when Differentiated Services traffic is sent to a non-DS–capable domain. The main alternatives in this case seems to be that all packets are marked with default PHB, or packets are mapped into class selector codepoints. The latter approach may provide a limited version of Differentiated Services, although the overall result is difficult to predict.

Multicast Streams

RFC 2475 briefly discusses two issues related to multicast streams. First, an incoming multicast packet can consume much more network resources than an incoming unicast packet. Note that multicast packets usually consume fewer resources than unicast packets given the number of recipients. In addition, the amount of resources needed inside the network is difficult to predict at the ingress node. This may cause fairness problems between unicast and multicast streams. One approach to alleviate this problem involves the understanding that multicast packets use different codepoints and different PHBs than unicast packets.

The second issue relates to a situation in which a multicast packet coming through an ingress node may be transmitted to several different network domains. This makes if difficult to select a DS codepoint that conforms to all the service agreements of the different domains.

Security and Tunneling issues

The main security issues considered both in RFC 2474 and RFC 2475 are the so-called denial-of-service attacks and the interactions with security protocols. Because certain PHBs and the corresponding codepoints provide better service compared to majority of traffic, some users may try to modify codepoints in their packets to try to obtain better service. In the worst case, this kind of behavior can yield to a denial-of-service attack, in which the modified packets exhaust the resources available for other traffic streams.

Moreover, because interior nodes are allowed to rely purely on the codepoint's value set by the boundary nodes, it is of great importance to design boundary nodes in such a way that every packet gets an appropriate codepoint value and properly traffic-conditioning actions are applied to all traffic streams marked with any preferential PHB.

3.3.4 A Framework for Differentiated Services

The framework document, RFC 2475, addresses basically the same issues as this book. In general, it provides a lot of helpful ideas, concepts, and recommendations for building networks and services based on the Differentiated Services approach: The document is one step toward real service differentiation in the Internet. Because practically all the issues considered in RFC 2475 are discussed in Part II, "Building a Network Domain Based on Differentiated Services," it is not worthwhile to review the whole document. Nevertheless, certain issues may need clarification, particularly in cases where the viewpoint of this book differs from that of the framework document.

Service Models

According to the framework document, the service provider forms the service by combining PHBs, traffic conditioners, provisioning strategies, and billing mechanisms (Bernet *et al*. 1998). Without doubt, this is a reasonable statement. Two additional qualifiers presented in the framework document are as follows:

- DS services are for unidirectional traffic.
- DS services are for traffic aggregates.

On certain levels of the system, these seem to be reasonable and valid statements.

Remember, however, that because customers are surely interested in traffic in both directions (most applications need a transmission channel in both directions), both directions must be somehow taken into account on the customer service level. Another issue is the technical implementation that can be based on unidirectional transmission channels. Moreover, it is somewhat questionable to equate the treatment of aggregate and service, because service depends crucially on the traffic-conditioning actions at boundary nodes, and these actions can be customer specific or even flow specific.

The fundamental difference between the Differentiated Services of this book and the framework document is the *basic model of service*. Part of the framework is generally applicable with various situations and with various service models. A good example of this is the taxonomy of services that discerns three basic categories: qualitative, quantitative, and relative. The examples presented in the framework document are illustrative. The service definitions related to packet-loss ratio could be as follows:

- *Qualitative service*: Low loss ratio.
- *Quantitative service*: Less than 5% packet-loss ratio.
- *Relative service*: Packet-loss ratio on service level A is smaller than that on service level B.

This practical classification will be used in the Chapter 4 of this book as well. Nonetheless, part of the framework document seems to be based on the assumption that service differentiation definitely means different levels of quantitative service (Bernet *et al*. 1998). Section 4.2 of the framework document states that TCA is an important subset of the SLA, for example, and it specifies detailed service parameters for each service level. Examples of service parameters are *throughput*, *drop probability*, and *latency*. Moreover, there are several references to traffic profile with conforming and nonconforming traffic; without quantitative parameters, the meaning of conformance is a somewhat unclear concept. Obviously, the assumption is that customer service is always based on a strict control regardless of the nature of the network service.

The reasoning seems to be that the customer requires certain service that is defined by a detailed TCA including traffic parameters, such as bit rate. The service provider offers a service with a certain price that likely depends on bit rate and quality requirements of the connection. Based on the TCA, the service provider measures the traffic sent by the customer at the boundary node and marks each packet either in-profile or out-of-profile. This is a very similar model to that of VBR service in an ATM network when CLP bit is applied. The main difference between ATM / VBR service and this service model is the implementation of traffic-control functions inside the network.

Without doubt, this is one possible approach, but not the only one. If the service provider wants to adopt a significantly different service model, some of the ideas presented in the document are difficult to apply. It should be noted, however, that the framework document clearly states that static SLAs are the norm at the present time. Therefore, changes in TCA can occur on the order of days or weeks rather than seconds or minutes. The framework document seems, in a way, to hover between the traditional connection-oriented service model and a new Differentiated Services model that is still to be defined. (This book endeavors to fill this gap.)

Despite this minor criticism, the framework document offers many useful instructions that provide additional tools for effective service differentiation. The fourth section, for instance, discusses problems related to the requirement of controlling received traffic and to the provision of dynamic SLAs.

Further, section 5 introduces three possible service models: better than best-effort service, leased-line emulation service, and quantitative assured media playback service. Better than best-effort gives higher priority than the normal best-effort service. By that means, the content provider can transmit packets with a higher rate than other content providers can. Leased-line service can be used by corporations to transmit, for instance, IP telephony calls between network sites of a corporation. Media playback service provides similar characteristics as the leased-line service, but with lower level of assurances. All these services can be implemented by using the EF and AF PHB groups. (See sections 7.3, "Expedited Forwarding PHB," and 7.4, "Assured Forwarding PHB Group," in Chapter 7.)

Provisioning and Configuration

Provisioning and configuration issues are discussed thoroughly in Chapter 4. It could be, however, helpful to illustrate the main points in the framework document, as well as the main differences between the view of the framework document and this book.

The boundary provisioning is the easy part of the issue, in particular in the direction from the boundary node to the core network. The much harder question is the interior provisioning—that is, the dimensioning (to use the term applied in Chapter 4 of this book) of link or service class capacity inside the network. As appropriately noticed in the framework document, a good understanding and control of traffic is necessary for efficient provision (Bernet *et al.* 1998). The statement is that although traffic volumes cannot be anticipated with 100% accuracy, the internal provision is still a tractable problem. Although this statement could be partly true, the whole issue is so complicated that it could be overoptimistic to rely on the predictability of the traffic process unless both the route and the traffic sent by the customer are strictly controlled.

One approach to alleviate the provisioning problems is to make certain that quantitative and qualitative services are isolated by using different PHBs. In general, quantitative services should have higher priority than qualitative services, although this definitely depends on the level of quantitative assurance. In this respect, relative services can be considered as a system consisting of several qualitative services. Moreover, it is supposed that only a small fraction of traffic uses the quantitative services.

In section 6.2 of the framework document, it is said that dynamic provision techniques are desirable because traffic volumes are likely to change dynamically, even if TCA is static (Bernet *et al.* 1998). Dynamic provision means in this connection either that capacity requests are signaled through the network, or that the nodes adjust resources based on measurement results. Dynamic provision is in principle an apprehensible idea, and in some cases certainly necessary. It should be noted, however, that signaling mechanisms do not belong to the main tools used by the Differentiated Services Working Group.

Viewpoint of This Book

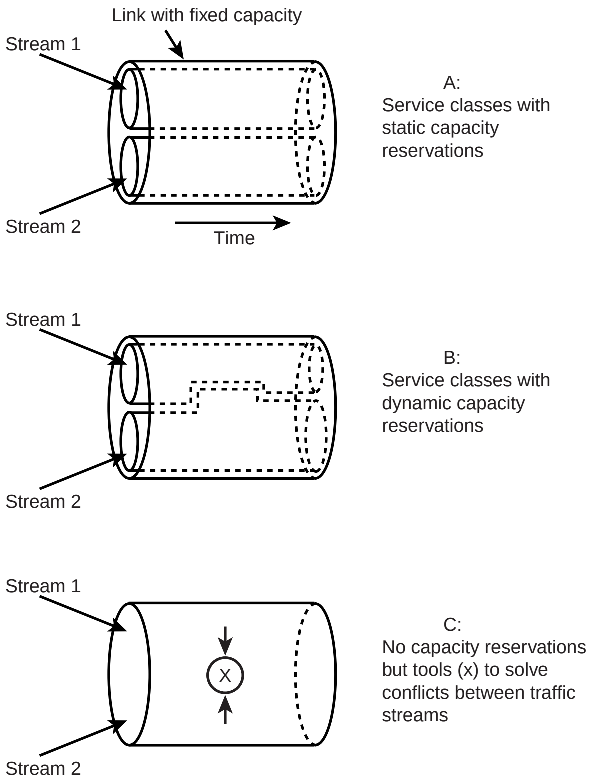
It is not clear whether the dynamic provision truly is the best approach to optimize the use of network resources. Figure 3.9 illustrates the basic dilemma of dynamic provision. There are two traffic streams sharing a link with a fixed-capacity reservation for both; the streams can represent two service classes with different quality requirements. The starting point could be that a fixed capacity is reserved for both traffic classes. (Case A in Figure 3.9 shows this.) This is probably not a very efficient approach if traffic streams are highly variable.

One possible way to improve the situation is to dynamically adjust the capacity of service classes based on the momentary traffic load of the streams. (Case B in Figure 3.9 shows this.) Some issues limit the usefulness of dynamic provision, however. First, there is usually an ultimate limit for the total reservation because link capacity seldom can be adjusted dynamically. Therefore, an increase of reservation for stream 1 must be taken from somewhere, either from another reservation or from a free pool. Second, reservations as such consume some resources, and moreover they can never follow every change in traffic demand.

A third alternative could be the key to better use of network resources. It is based on solving a conflicting situation when it really occurs rather than on proactive reservations. (Case C in Figure 3.9 shows this.) As long as there is enough capacity for all traffic streams, no special action is needed: The system merely transmits packets forward. Only when there are momentary overloads is a decision system activated in a way that makes possible an efficient use of the network and fair service for all traffic streams.

In this model, no explicit capacity reservation is used to avoid conflict situations, and no explicit mechanism is used to warn senders about possible overloads. This is a fundamental idea of Differentiated Services presented in this book—but note that there still are differing opinions about the basic principles of Differentiated Services.

Figure 3.9 Three approaches to sharing link resources.



Summary

The evolution of Differentiated Services has been eventful and fast. Although part of the frenetic activity could be classified as hype rather than progress, the establishment of a dedicated working group has systematized the effort (thanks to the co-chairs of the working group, Brian Carpenter and Kathleen Nichols).

This chapter provided an historical overview that illustrated the many various goals for Differentiated Services. This variety has had a significant effect on the development process so far, and it is likely that the diversity of proposals will be even greater in the future. Whether all proposals truly comply with the principles of Differentiated Services is a matter of continuous debate.

The most concrete results of the first year are two RFCs that specify the structure of the DS field in IP packets and the basic architecture for Differentiated Services. This chapter provided an overview on these documents, as well as an overview of a framework document that contains instructions about the application of Differentiated Services models in real networks.

The most important concept of Differentiated Services is Per-Hop Behavior (PHB). Although PHB is primarily a technical term, PHB should not specify mechanisms. Moreover, PHB is also used to depict the purpose of the whole system; it should not be a service specification, however. It will take a lot of expertise and effort to define the PHB proposal in an appropriate manner.

The final part of the chapter illustrated the main differences between the working group documents and the later chapters in this book. The main reason for developing Differentiated Services, from the perspective of this book, is not merely to provide scalable implementation of the integrated services model. In technical terms, this book resolutely promotes an approach based on flexible sharing of network resources rather than capacity reservations. This starting point makes it necessary to reconsider the whole service model of the Internet, which is the topic of Chapter 4.