

Per-Hop Behavior Groups

The traits and characteristics that define a Per-Hop Behavior (PHB) consist of the rules used to treat packets in specific ways inside the network. The rules should be indisputable in the sense that they can be used for practical (human) discussions among different parties, particularly service providers, backbone operators, and vendors. Otherwise, the whole concept of PHB can be harmful rather than useful—PHB itself is situated somewhere in the nebulous space between mechanisms and services.

The term *hop* also needs some clarification. A path between two end systems consists of a number of hops. From a Differentiated Services viewpoint, a hop is a subpath between two network nodes that does not have any significant effect on the characteristics of traffic flows. An ATM virtual path (VP) between two routers is one hop, for instance, although the VP goes through several ATM nodes, if those nodes keep the traffic process virtually unaltered. If a node makes decisions related to statistical multiplexing (such as which packets or cells should be discarded), the incoming and outgoing links evidently belong to different hops.

One of the fundamental ideas of DiffServ is that the treatment of flows should be similar in every node and over every hop to provide reasonable network service. This chapter makes an earnest effort to define the following four PHB groups:

- Class selector PHB
- Expedited forwarding PHB
- Assured forwarding PHB
- Dynamic RT/NRT PHB

The search is based on information covered in the previous chapters, which provided various perspectives on Differentiated Services from service and quality models to packet marking

and queuing. (Refer to Chapter 4, “General Framework for Differentiated Services,” and Chapter 5, “Differentiation of Customer Service.”)

7.1 *Systematic Basis for PHB Evaluation*

This section introduces the system used to evaluate PHB groups. First, I give a clear insight into each PHB group. The description of PHB should be something that can be used for practical discussions among different parties. The description does not, however, make any reference to either any particular implementation or to any particular end-to-end service.

Although Per-Hop Behavior is not a service in the traditional sense, PHBs cannot be analyzed without considering services. Service is the ultimate purpose of defining and implementing PHBs. One of the main difficulties of designing PHBs is that there are various different ideas on what the service build from the PHBs should be. Table 7.1 summarizes the main service aspects from fairness criteria to VPN type.

Note

Although the importance of each individual aspect of service depends on the business model of the service provider, in general, all aspects are relevant and should be considered. Each PHB group differs on both the relevance of different aspects and the applicability to different service models.

Table 7.1 Service Models

Aspect of Service	Service Models	Chapter Section	Figure
Fairness criteria	Application Customer Organization	4.2.2	4.3
Dynamics and degree of guarantee	Guaranteed connections Leased-line service Dynamic importance Resource sharing	5.1.1	5.1
Control of load and destination	Constant bit-rate (CBR) connection “VPN with VBR service” Best effort	4.4.1	4.4
VPN type (multiplexing)	No capacity sharing Limited sharing Total sharing	4.5.3	4.13

One key aspect of a PHB group is how it establishes a feasible quality structure. Without a consistent and comprehensible quality structure service, providers cannot build reasonable services within their own domain (and even fewer services extending several domains). Urgency (or delay), importance (packet-loss ratio), and bandwidth (bit rate) are the main quality aspects. Every extensive PHB should somehow cover all these aspects, and a confined PHB should define how it can be located in a broader framework covering all aspects. Predictability of quality is an aspect that can have an autonomous importance in some cases. Table 7.2 shows the main quality options and references to earlier chapters.

Table 7.2 Quality Models

Aspect	Main Options	Chapter Section	Figure
Urgency and importance	Fixed order of import. Flexible order of import.	4.6.2	4.14 4.15
Bandwidth and importance	Bandwidth partition Proportional bandwidth	4.6.3	4.16
Predictability	Marking depending on duration of flow Guaranteed service	4.3.2, 4.6.4	n/a

Service differentiation is defective without sophisticated pricing. Although there is not always any direct relationship between a PHB group and a pricing model, you can often assess which quality aspects are easy to realize and which ones are either difficult to realize or unrealistic. Table 7.3 presents the main relationships between quality aspects and price.

Table 7.3 Relationships Related to Pricing

Aspect	Versus	Chapter Section	Figure
Bandwidth	Price	5.1.3	5.8
	Availability		5.11
	Quality		5.13
Availability	Price	5.1.3	5.10
	Bandwidth		5.11
	Quality		5.12
Quality (delay)	Price	5.1.3	5.9
	Availability		5.12
	Bandwidth		5.13

Technical implementation is the other main part of network service in addition to the general structure. Tables 7.4, 7.5 and 7.6 summarize the main implementation alternatives. The most central part of any PHB implementation is related to the interior nodes—remember that PHB refers to the treatment of the packets inside the network. Nonetheless, because boundary node functions are essential parts of service provision, they have to be addressed when evaluating the whole system build by a given PHB group.

Table 7.4 Implementation Options for Boundary Nodes

Aspect	Main Options	Chapter Section	Figure or Table
Service request	RSVP message DSCP indication Quality inquire	5.1.2	Table 5.1 Table 5.2
Classifying	Based on service-level agreement (SLA) Based on application	5.2.1	n/a
Metering	Recognition of excess load Metering of load level	5.2.2	Figure 5.17
Marking	Only excessive packets All packets based on momentary load	5.2.3	Figure 5.19
Shaping	Before or after other conditioning functions	5.2.4	Figure 5.20 Figure 5.21 Figure 5.22
Discarding	No/yes	5.2.5	n/a

Table 7.5 Implementation Options for Interior Nodes

Aspect	Main Options	Chapter Section	Figure
Buffering	FIFO, Priority queuing, CBQ, WFQ	5.3.1	5.23 5.24 5.25
Discarding	Hard thresholds RED	5.3.2	5.26
Feedback information	Congestion indication Available bandwidth	5.3.3	n/a

Table 7.6 Implementation Options Related to Network Domain

Aspect	Main Options	Chapter Section	Figure
Routing	Static, adaptive	4.5.4	5.4
Resource reservation	Static or dynamic per flow or aggregate	4.4.4 5.4.2	4.9
Network dimensioning	Mathematical model Practical experience	5.4.3	n/a
Contracts between domains	Based on charging Based on marking	5.4.4	n/a

7.2 Class Selector PHB Group

According to the RFC 2474, the main motivation of Class Selector PHBs (CS PHB) is backward compatibility with present uses of bits 0–2 of the IPv4 TOS octet. The history of these bits is somewhat convoluted. According to the original Internet Protocol specification, the precedence can mean that only traffic above a certain precedence at the time of high load is accepted (Postel 1981).

In 1989, “The Requirements for Internet Hosts—Communication Layers,” stated that the precedence field (three first bits of TOS field) was intended for the Department of Defense applications of the Internet protocols (Braden 1989). The RFC even recommended that vendors should consult the Defense Communication Agency for guidance on the precedence field. It seems that the main use of the bits has been to separate packet related to network control from best-effort packets.

Because the use of these bits is similar to that of PHBs, it is reasonable to adapt the deployed uses of the bits into the Differentiated Services framework. Therefore, the perspective of the Class Selector PHB group is special: Although the other PHB groups are used to build new services, the definition of the CS-PHB group tries to encompass various, in worst case contradictory uses, in existing networks. Because of this background, the specification of CS PHBs primarily gives limits of acceptable use of the bits, but not any unambiguous exact rules.

7.2.1 Description

RFC 2474 states the following:

A Class Selector PHB should give packets a probability of timely forwarding that is not lower than that given to packets marked with a lower Class Selector PHB, under reasonable operating conditions and traffic loads.

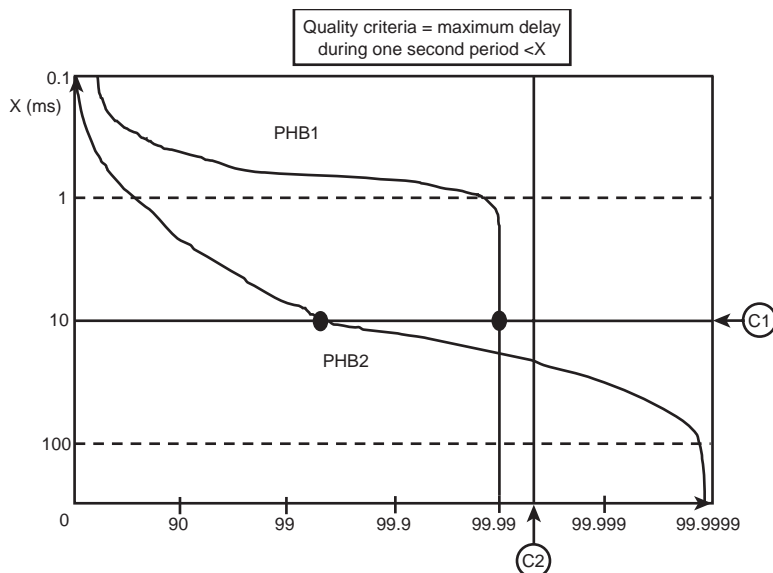
This can be called the *timely forwarding* requirement. A further clarification is that a discarded packet is considered to be an extreme case of untimely forwarding. Moreover, a network node can limit the maximum resources used by each PHB. CS PHBs might be used to provide delay, importance, or bandwidth differentiation.

All three aspects can be obtained at the same time only if the traffic load within each PHB is tightly controlled—which is an unrealistic assumption in Differentiated Services networks. Particularly, *timely forwarding* and *bandwidth enforcement* objectives can conflict during a congestion situation. If a higher PHB class is using an excessive amount of resources, the node cannot comply with both the timely forwarding requirement and the bandwidth enforcement requirement. The interpretation used in this chapter is that bandwidth enforcement is an acceptable tool to realize the timely forwarding behavior, but it is not an independent target of the CS-PHB group itself. This leaves two aspects, delay and importance.

The fundamental question is, what does higher probability of timely forwarding primarily mean? Figure 7.1 illustrates this dilemma. The figure shows two PHB behaviors:

- *PHB1*: Provides excellent delay characteristics most of the time—that is, less than 1-millisecond delay per node. However, the probability that a packet has to be discarded (0.01%) could be unsatisfactory for some applications.
- *PHB2*: Provides worse delay characteristics for forwarded packets, but the packet-loss ratio is significantly better than that of PHB1.

Figure 7.1 Criteria for higher probability of timely forwarding.



As a result, the requirement of *higher probability of timely forwarding* is ambiguous. Only in cases where the delay characteristics are systematically better for one PHB than for another one should it be stated that the probability of timely forwarding is higher for the first PHB. Basically, there are four approaches to solve this problem:

- Both the delay and loss ratio of a CS PHB must be at least as good as any of the lower CS PHBs.
- The probability of timely forwarding is defined for fixed-delay characteristics—for instance, for a 10-millisecond delay (criterion C1 in Figure 7.1).
- The availability value is fixed. For instance, the “timeliness of forwarding” is defined for availability of 99.995% in Figure 7.1 (criterion C2).
- The question is intentionally left vague; each operator is allowed to define its own criterion (if any).

Although the two middle approaches are somewhat attractive because they give, in principle, a tool for systematic assessment of any PHBs, they could be hard to apply because the final conclusion depends essentially on the traffic processes within each PHB. Therefore, either the first tight approach or the last loose approach should be applied. The compromise used in this chapter is that the requirement of consistent ordering based on packet-loss ratio should be met, and there could be small deviations in the ordering for different delay criteria.

Returning to the original goal of CS PHB, backward compatibility, the application of CS PHBs is limited by the mechanisms available in the interior nodes. Therefore, the main advantage of CS PHB is that the service provider can avoid upgrading interior nodes but still attain some level of differentiation by upgrading boundary nodes.

If the interior nodes of a service provider have only a couple of importance (or delay) levels without any supporting tools, however, the application of the CS PHB system remains inevitably limited. Actually that kind of system deserves only a brief discussion in this book. On the contrary, the approach of this section is to consider the full possibilities of the CS-PHB system with all eight individual PHBs, although it is not certain that current equipment supports this system.

7.2.2 Position in the Framework

This section identifies the location of the CS-PHB group in the framework(s) introduced in the first section of this chapter. The extensive framework consists of service models and quality models, whereas *pricing* is an issue that is hard to assess in case of CS PHB because boundary behaviors are a totally open issue.

Service Models

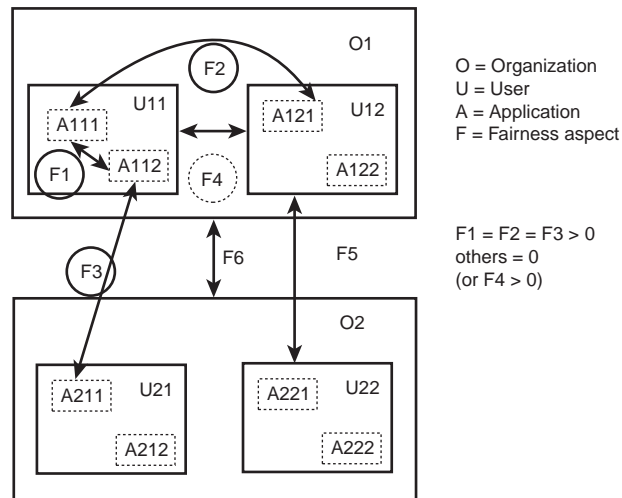
Let's start with delineating the three main service models according to the most relevant fairness targets:

- In the application model, fairness issues are considered from an application viewpoint.
- In the customer model, they are considered from the viewpoint of the paying customer.
- In the organization model, the viewpoint is an organization that wants to provide an appropriate service to a large number of end users.

Because of its special background, the CS-PHB structure is primarily an application service model rather than a customer or organization model. Another possible approach is that each person or other entity has a right to use specific PHBs. In that sense, CS PHB could be a tool to regulate the use of network resources among end users.

Yet, the basic nature of CS-PHB is a pronounced lack of any traffic-conditioning specifications. Therefore, the main fairness issues to be considered are those among applications: F1, F2, and F3 in Figure 7.2. Of course, the CS PHBs can be used to build customer services with appropriate traffic conditioning at boundary nodes. Although that is possible, it appears more reasonable to use other PHB groups for that purpose, because there is no common understanding of the proper condition for the CS-PHB group.

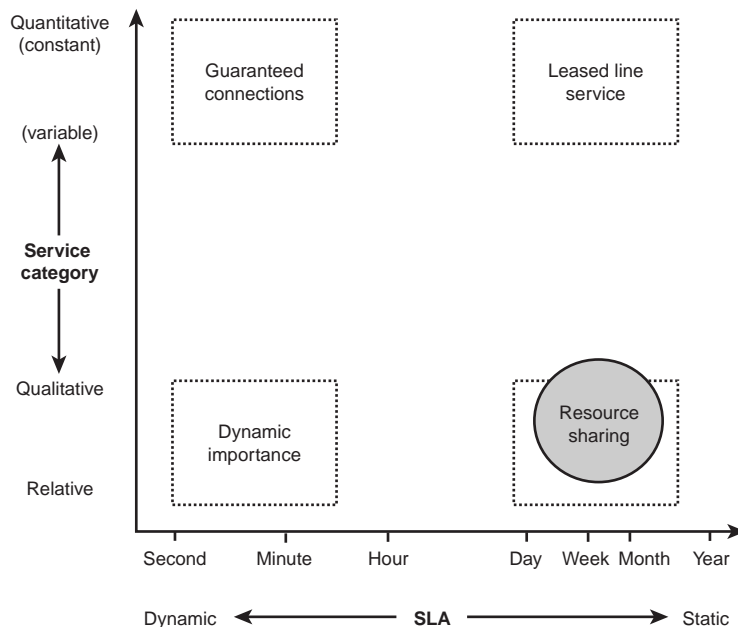
Figure 7.2 Fairness criteria for Class Selector PHBs.



The issue of service categories is somewhat irrelevant to CS PHBs, because the original target of the bits was not to build services. Nonetheless, in a future environment with specific services for customers, even CS PHBs will have a proper location—they cannot merely float over other PHBs and services without any relationship with them.

One way to define a *class selector service* is to eliminate those service models that are not reasonable. The guaranteed-connections model is apparently excluded because it does not have traffic-control functions in boundary nodes. Dynamic importance is excluded because it doesn't have a signaling system to transmit information about changes in packet classification. Leased-line service is excluded because it seems irrational to define different leased lines in the order of probability of timely forwarding. (Leased lines are more about bandwidth separation.) Therefore, as for the service classification presented in Figure 7.3, the only relevant model is the resource-sharing model. In other words, the system behavior is static and the service description is either qualitative or relative.

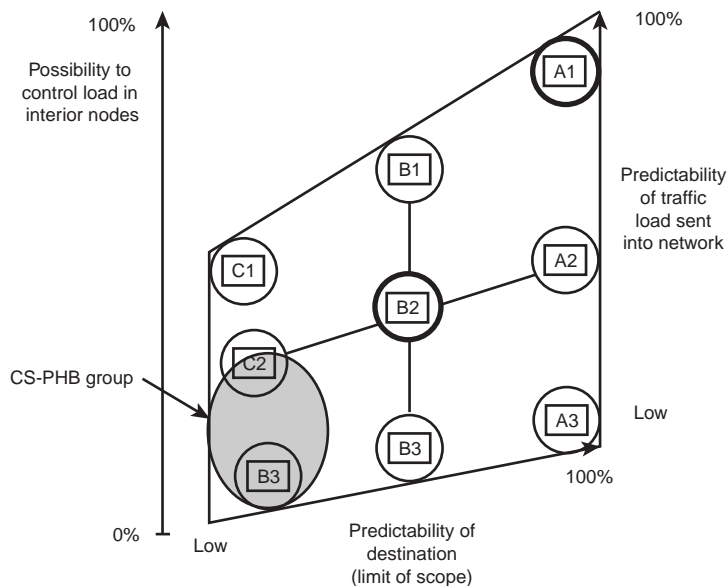
Figure 7.3 Most relevant service category for Class Selector PHB group.



The origin of Class Selector PHBs is deep in the traditional IP networks. The fundamental assumption in IP networks has been that there is only limited traffic-control mechanisms inside the network. The best-effort model includes the principle of not controlling the destination or the traffic sent into the network. That statement concerns mainly the lowest PHB, not the other levels, because several levels of importance can hardly be used without

any clear limits of use. Although this issue is difficult to assess without any systematic mechanism in boundary nodes, it seems that the CS-PHB group does not fix the destination of packets. Because some level of load control is inevitable, the region of CS PHB is extended in Figure 7.4 somewhat in the moderate level of traffic control.

Figure 7.4 Predictability of load and destination for the Class Selector PHB group.



No fundamental obstacle stands in the way of combining the Class Selector system with appropriate mechanisms at boundary nodes, and thereby extending the application of CS PHBs into the areas of more advanced control of load and destination. This remark is valid also with the last customer-service aspect in Table 7.1, virtual private networks; although the Class Selector PHB group is perhaps not directly applicable to building VPNs, the CS PHBs might be used to control traffic within a VPN.

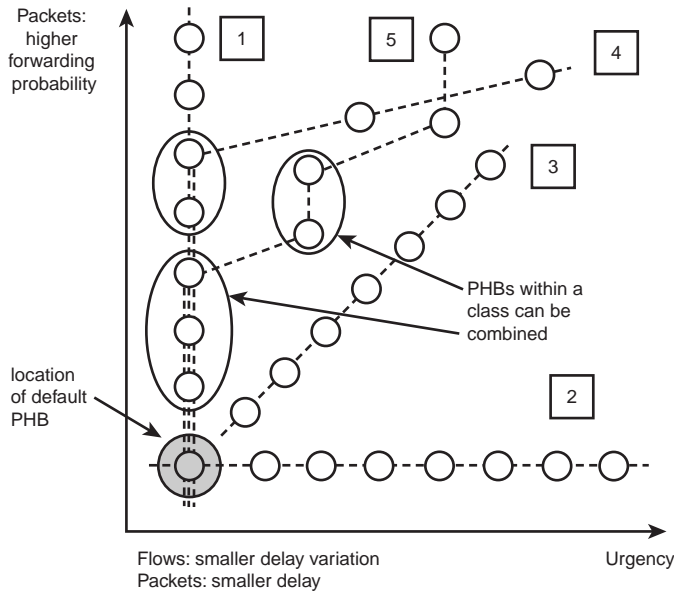
Quality Models

As discussed in the preceding section, “Service Models,” the requirement of higher probability of timely forwarding can be realized in different ways. Figure 7.5 shows some of the possibilities. The first option is that only the importance, but not urgency, is a significant factor. In this case, there are several (at most eight) importance levels within one PHB class. This straightforward scheme is unacceptable, however, because there has to be two independently forwarded classes of traffic according to RFC 2474, “Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers” (Baker *et al.* 1998).

The alternative 2 in Figure 7.5 means that there are eight different delay classes with the *same* importance. This seems technically acceptable and realizable, but apparently this is not the original purpose of the Class Selector PHBs. A pure delay differentiation without any differences in packet-loss ratios is not a viable approach if the packet-loss ratio is as high as can be expected in the Internet.

The third option in Figure 7.5 is that every CS PHB differs both in delay and importance. Although this is definitely acceptable according to the Class Selector PHB specification, a system with eight discernible delay classes is not necessarily a practical goal. Approach 4 could therefore be more realistic: six (or perhaps five or seven) lowest importance levels share the same delay class, but both of the two highest importance levels have their own delay class. The highest class can be used for transmitting packets related to internal traffic such as traffic-control messages and routing messages. The second highest class can be used for demanding real-time applications. Finally, a more complicated structure 5 is possible as well.

Figure 7.5 Delay and importance relations of Class Selector PHBs.



The other fundamental quality-control target presented in Table 7.2 is bandwidth management. Although there is no obstacle to dividing the network into separate parts and to use CS PHBs for quality management, the main purpose of the Class Selector PHBs is most likely in other quality aspects.

What, therefore, is the role of *predictability* in the case of the Class Selector PHB group? This is a somewhat artificial question—there is no direct relationship between the duration of flow and packet-loss ratio with any of the CS PHBs. (Note that predictability of quality means here that a flow with high quality in the beginning gets the same quality for the duration of the flow). The main aspect of predictability in case of CS PHB seems to be that, for instance, routing information *always* gets high probability of timely forwarding. (In the terminology used here, however, that is an importance aspect rather than a predictability aspect.)

7.2.3 *Required Technical Tools*

This section mainly addresses approach 4 in Figure 7.5. The assumption is that the operator wants to implement three different classes: Each of the two highest classes consists of one PHB, and the lowest six PHBs share the same PHB class. Although there surely are other reasonable approaches as well, the implementation requirements do not differ significantly from those presented in this section.

Implementation Options for Boundary Nodes

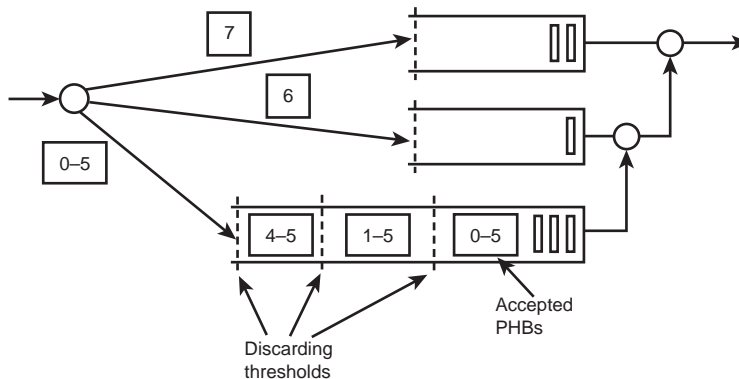
Again, the problem of obscurity arises. When no clear idea about the general objective of the system is available, it is impossible to say much about the mechanisms needed in boundary nodes. The only evident issue is that a logical marking system is required, in the minimum, based on the rights of use of certain PHBs. Packets with inappropriate marking can be either dropped immediately or marked with the default PHB. In addition, a logical scheme is to provide classification based on applications.

Implementation Options for Interior Nodes

Option 4 in Figure 7.5 requires three queues. The highest queue has strict priority over all other queues, the second queue has strict priority over the lowest queue, and the lowest queue is served only if the other queues are empty. The two higher queues accept packets as long as empty space exists. The lowest queue is divided into three regions. If the occupancy level is low enough, all packets are accepted; in the middle region, default packets (PHB=0) are discarded; and in the highest region, only packets with PHB=4 or PHB=5 are accepted.

As usual, the default PHB, which is used to implement the best-effort service, clearly benefits from RED. The other thresholds for the lowest queue can also use RED, whereas the higher queues most likely apply a hard threshold. Figure 7.6 presents a possible implementation of Class Selector PHB.

Figure 7.6 An example of Class Selector PHB implementation.



Implementation Options Related to the Network Domain

It is difficult to identify any specific requirements for routing, network dimensioning, and contracts between domains related to the Class Selector PHB group. Appropriate routing and network dimensioning are surely necessary components. Then contracts between different domains need to define the rules related to the use of CS PHBs.

Finally, the issue of explicit reservation for some of the CS PHBs must be addressed. Considering all the options presented in Figure 7.5, it seems that explicit bandwidth reservations are harmful rather than useful. In general, reservation leads to a situation in which the packet-discarding decision may depend more on the momentary load level of the class than on the importance level of the packet. Why should this be done? Two reasons can be identified:

- *Increased robustness against faulty situations:* If a defective device sends a large amount of packets with a certain PHB, for example, other PHBs still obtain some network resources.
- *Avoiding starvation of the lowest PHBs:* In particular, the default PHB should obtain a moderate amount of resources even during overload situations.

It could be useful, therefore, to have some level reservations that have no significant effect on the network behavior under normal load conditions but that may improve the network performance, robustness, or fairness under high load conditions.

7.2.4 Evaluation of Attributes

Now it is time to apply the list of attributes to check whether all relevant aspects have been considered. It is fair to assume that the main reason to apply Class Selector PHBs is

enhanced traffic management within a domain without any significant effect on customer services. In this case, the main fairness requirement is that the default service should provide a moderate service quality, at least without considerable interruptions. If the network can provide this characteristics, the better quality of other PHBs could be fair and acceptable.

Because the implementation and management of CS PHBs seem to be relative straightforward, cost efficiency is probably not a major concern. One of the primary ideas behind the use of CS PHB is improved robustness. If the classification of packets is appropriate, the prioritization of packets could significantly better the network performance under abnormal situations. The main deficiency of the Class Selector model is versatility. There is one clear customer service, best effort, but not any other commonly accepted service level. If service differentiation is necessary, other PHB groups should be used instead of, or in addition to, the Class Selector PHB group.

7.3 *Expedited Forwarding PHB*

Expedited Forwarding PHB (EF PHB) has a clearly defined target similar to the guaranteed-services model addressed in Chapter 2, “Traffic Management Before Differentiated Services.” The main merit of EF PHB is the adaptation of the guaranteed-service model into the framework of Differentiated Services.

7.3.1 *Description*

The objective of *Expedited Forwarding Per-Hop Behavior* (EF PHB) is to provide tools to build a low loss, low latency, low jitter, assured bandwidth, end-to-end service through DS domains. The essence of EF seems to be in the following requirements for the treatment of an EF PHB traffic stream (Jacobson, Nichols, and Poduri 1998):

The departure rate *must* equal or exceed a configurable rate. EF traffic *should* average at least the configured rate when measured over any time interval equal to or longer than a packet time at the configured rate.

What does this actually imply? Apparently, if packets of the same size come exactly according to the configured rate, the node should send the packets forward with exactly the same rate. Although the situation is not as clear with variable size packets, the basic idea is to avoid any excessive delay and jitter if possible. The other side of this model is that it requires tight traffic control in boundary nodes.

Figure 7.7 illustrates the principle of EF PHB. Note that marking is not available in case of EF PHB because there is only one importance level. Consequently, if a packet arrives before its scheduled time, there are three options both in boundary and interior nodes:

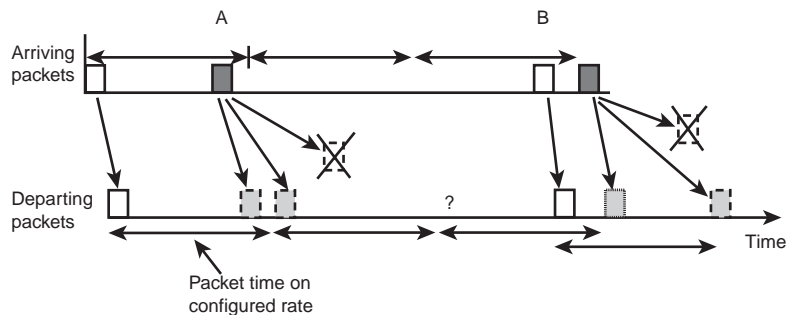
- To forward the packet immediately
- To forward the packet at the scheduled time
- To discard the packet

Before scheduled time means here that the stream would exceed its configured rate if the packet were sent immediately.

There is a clear difference between boundary and interior nodes. Traffic control in a boundary node should apply either the second or third option to prevent the source from exploiting excessive bandwidth. In contrast, the wording of the RFC indicates that the first option is strongly recommended inside the network. The second, shaping option, also appears to be possible inside the network. Packet B in Figure 7.7 depicts why shaping is not necessarily a reasonable alternative. If packet B has encountered additional delay inside the network, the next packet (C) may come too soon after packet B compared to the configured bit rate. The shaping option means that packet C, and possibly all the following packets, will encounter excessive delay because of the problems of one packet in the middle of the stream. That behavior seems to be against the principles of EF PHB.

In summary, EF PHB means a strict bit-rate control in the boundary node and as quick forwarding in the interior nodes as possible. The fundamental assumption is that traffic is controlled in boundary nodes in a way that each packet inside the network can be delivered immediately without risking that the flows get excessive bandwidth.

Figure 7.7 Shaping and discarding in EF PHB.



Moreover, one principle of EF seems to be that an excess rate of EF aggregate inside the network is an erroneous condition; and for protecting other services, EF packets can be discarded rather than packets belonging to the other PHBs (usually with lower importance). This might or might not be a rational rule. Nevertheless, if the excess is a result of a configuration error inside the network, packets belonging to other PHBs should apparently be

discarded first. On the contrary, if the reason is that a traffic source can send an excessive amount of EF packets because of a defective traffic control in the boundary nodes, it could be justifiable to discard EF packets before some other packets. Note in this case, however, that several innocent EF flows may perceive deteriorated quality.

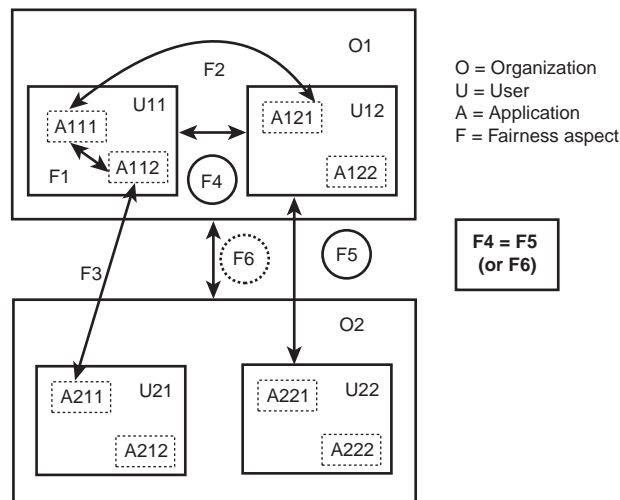
7.3.2 Position in the Framework

Expedited forwarding is one of the clearest specifications within Differentiated Services. It has a comprehensible target and a large part of the implementation is relatively easy. This section attempts to further clarify the model, and in particular, to address those issues not profoundly addressed in the EF specification.

Service Models

The primary EF service model is based on the customer contract. Therefore, the main fairness issue considered by EF PHB is the fairness between different customers, as shown in Figure 7.8. In addition, it is possible to apply EF PHB to separate organizations. The usefulness of EF PHB for that purpose is still somewhat questionable, because it does not offer any tool for traffic control inside the bit pipe (because there is only one importance class).

Figure 7.8 Relevant fairness issues for expedited forwarding.



EF PHB specification mentions one service, virtual leased line (Jacobson, Nichols, and Poduri 1998). Therefore, there is no problem in defining the main region for EF in Figure 7.8. With a static service-level agreement, EF PHB can provide a leased-line service. In

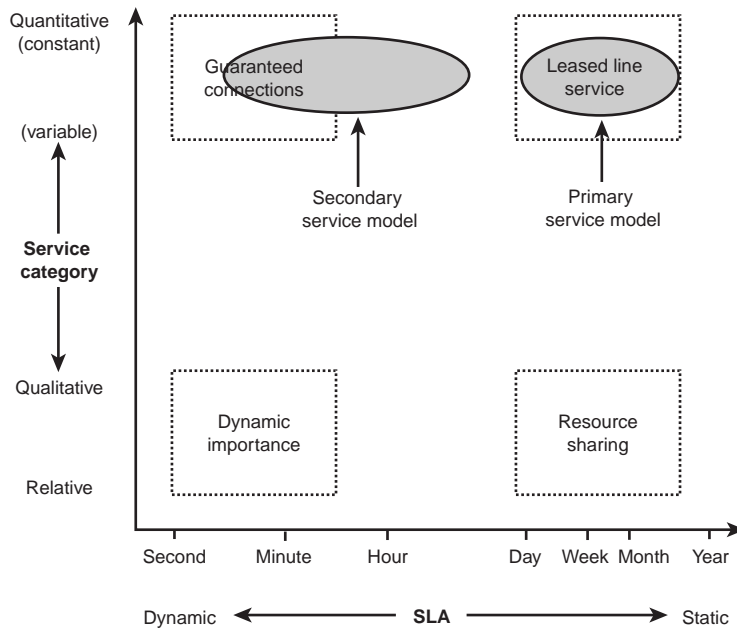
addition, if the network can support dynamic SLAs and dynamic capacity reservations, the EF region can be extended to the area of *guaranteed connections*.

Note

The term *guaranteed connection* in this book refers to a service used by an individual application with specific quality requirements and limited duration.

It should be further noticed that EF does not support variable bit-rate reservations, although the real bit rate may, of course, be variable provided that the peak rate remains below the policed rate. Figure 7.9 shows the principal service models that can realized by Expedited Forwarding PHB.

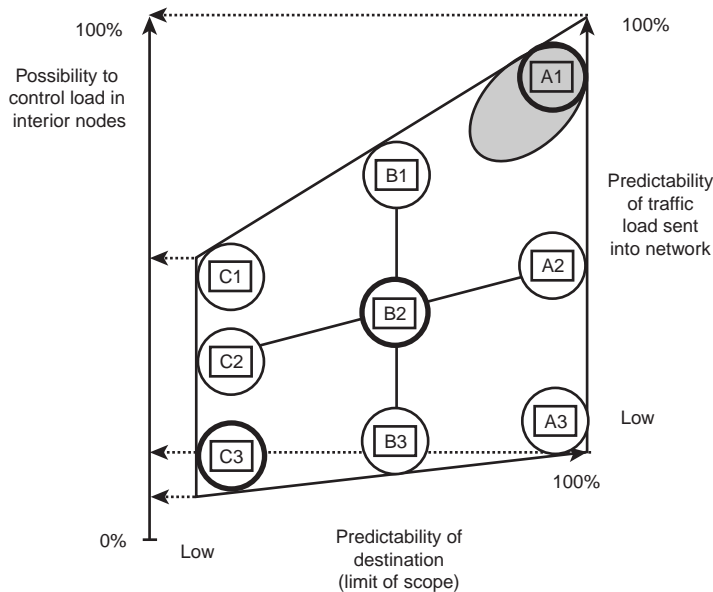
Figure 7.9 Primary and secondary service models for EF PHB.



The fundamental principle of EF PHB is to provide very high quality, related to both delay and loss characteristics. Essentially, the aim of EF PHB is to provide a connection that appears as a physical link with very low delay and bandwidth variations. As extensively discussed in section 4.4.1, “Predictability of Load and Destination,” in Chapter 4, real assurances can be given only if traffic is controllable, and controllability requires that load be controlled in boundary nodes and that the destination be fixed.

Although EF PHB provides a strict traffic control in boundary nodes, fixing the destination or route is not explicitly required. Yet, a successful provision of EF type of service likely requires that the destination be known, because an EF service to an arbitrary destination appears to be an impractical approach. Consequently, the applicable region of EF PHB is in the upper-right corner of Figure 7.10.

Figure 7.10 Predictability of load and destination for Expedited Forwarding PHB.

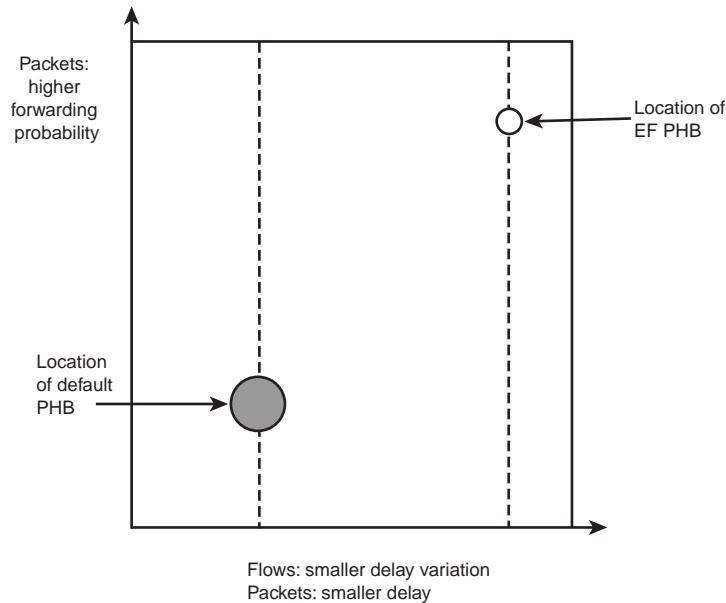


As mentioned earlier, the primary target of EF is to provide a reserved connection between two locations. Those connections can be used to build virtual private networks, but primarily without any capacity sharing between VPNs. The main problem of EF as a VPN technology is that a full mesh of reserved connections could be a complex and inefficient way to build VPNs, if the number of sites is large. This issue is discussed further in section 9.3, “Virtual Private Networks by Using an EF-PHB,” in Chapter 9, “Implementing Differentiated Services.”

Quality Models

The location of EF PHB in the scale defined by delay and importance is apparent, at least in relation to the best-effort service (default PHB). EF PHB should provide much better delay as well as much better loss characteristics, as shown in Figure 7.11.

Figure 7.11 Expedited Forwarding versus default PHB in delay-importance scale.



Predictability is the essence of the whole EF service model. A high level of assurance means largely a predictability of quality realizable only with the following:

- High-quality standard
- Constant bit-rate traffic model
- Fixed destination

These characteristics are exactly what EF PHB provides. If *real* predictability of quality is needed, therefore, the model adopted by EF PHB appears appropriate.

The service provider must also consider several other issues. Pricing is one of the most important issues. Table 7.7 summarizes the main approaches supported by EF PHB. Some kind of pricing related to the amount of bandwidth is necessary to avoid uncontrollable use of EF PHB. Another relevant aspect is the relationship between availability and price, in the sense that price should likely depend on the time of day and on the destination. Because there is only one quality class, there is no true quality differentiation within the EF PHB. (The situation changes when the operator simultaneously uses several PHB groups.)

Table 7.7 Charging Options for Expedited Forwarding

Differentiation of	As a Function of	Remark
Price	Bandwidth	Necessary
Price	Quality	Only one quality class
Price	Availability	Charging may depend on the time of day and day of the week
Availability	Bandwidth	Possible in limited sense (course granularity)
Availability	Quality	Not a probable approach
Quality	Bandwidth	Not a probable approach

7.3.3 Required Technical Tools

Because the EF PHB, as such without anything else, is quite a trivial issue on the packet-handling level, this section considers a combination of EF PHB and default PHB. The assumption is that default PHB is used to transmit all packets using the best-effort service, and all flows with higher-quality requirements use EF PHB.

Implementation Options for Boundary Nodes

The first issue is how the user should request an EF PHB, or actually a virtual leased-line service. If the primary model is based on static reservation, the request could be implemented in almost any imaginable way—for instance, by calling the service provider that makes the required configurations in network nodes through the management system. If a more dynamic system is needed, RSVP could be a suitable tool for capacity reservations. Bandwidth brokers, discussed in section 6.3.2, “Resource Reservation,” in Chapter 6, “Traffic Handling and Network Management,” can also be a part of the solution.

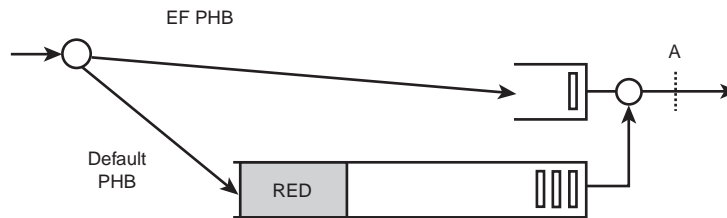
One possible scenario is for the user to inform the service provider that packets using certain port numbers and/or with certain destination addresses should be classified into the EF PHB, and all other packets use the default service. This seems to be necessary because users will probably want to use their network connection for several applications at the same time, and buying a separate EF connection for all flows is an impractical approach.

The workability of EF PHB requires that the boundary node strictly control the bit rates of each incoming flow. A token bucket type of mechanism is one possibility to identify excessive packets that should be dropped immediately instead of being marked as lower importance. Traffic shaping can be used to improve the loss characteristics perceived by end applications.

Implementation Options for Interior Nodes

The main principle of EF seems to be that EF packets deserve the best possible service. EF-PHB specification mentions as possible alternatives a simple priority queue, weighted round-robin, and CBQ (Jacobson, Nichols, and Poduri 1998). Figure 7.12 illustrates one simple realization with two queues: a relatively small EF queue with strict priority, and a larger default queue with the RED mechanism. The main additional requirement is that the available bit rate be high enough at point A regardless of other possible traffic streams.

Figure 7.12 One possible implementation of Expedited Forwarding with default PHB.



Implementation Options Related to the Network Domain

All the matters so far related to EF PHB have been clear and reasonable. The most critical area of EF implementation has not yet been discussed, however. The fundamental issue is how to make the resource reservations inside the network with possibly hundreds of nodes and millions of flows, without generating the same scalability problems that are the main concern of RSVP.

On a principle level, the answer provided by Differentiated Services and explicit forwarding is that reservations are not done for individual flows but for traffic aggregates. Unfortunately, this is not a perfect solution: If there are not per-flow reservations, boundary nodes cannot accurately decide whether to accept new EF connections. Therefore, without per-flow reservation, the destination is not fixed from the viewpoint of the interior nodes. That essentially deteriorates the controllability of EF traffic inside the network. (Note that if the aggregate reservation is always updated when a new flow is established, that *is* per-flow reservation.)

Network dimensioning is an alternative tool to manage the service. That is, the overall load level of the EF connection is kept low enough to virtually guarantee that there is enough capacity on every interior link within the network domain. In the extreme approach, each interior link would transmit all existing EF connections. As long as there is no robust and clear solution to this problem, the surplus value of EF PHB compared to the more traditional guaranteed-service approaches is questionable.

7.3.4 *Evaluation of Attributes*

Expedited forwarding has the same advantages and disadvantages as all other CBR services. Cost efficiency is a somewhat arguable issue in a large network with a limited number of EF connections; however, the additional mechanisms could be acceptable from an implementation and management perspective.

Fairness depends significantly on the pricing structure. The main issues are the relationship between price and reserved bandwidth and the pricing difference between services provided by EF PHB and default PHB. If the price gap is too large, the demand for and use of high-quality service could be negligible, whereas a very small difference may complicate the network dimensioning.

EF PHB by itself is not versatile. Note, however, that the purpose of EF PHB is not to provide the only basis for Internet services but rather to complement service provision. If EF PHB is built properly and the number of EF connections is limited, the robustness is not a big concern. In contrast, despite the intrinsic robustness of the EF model, the management of a large number of connections using a centralized OAM system could be prone to human errors.

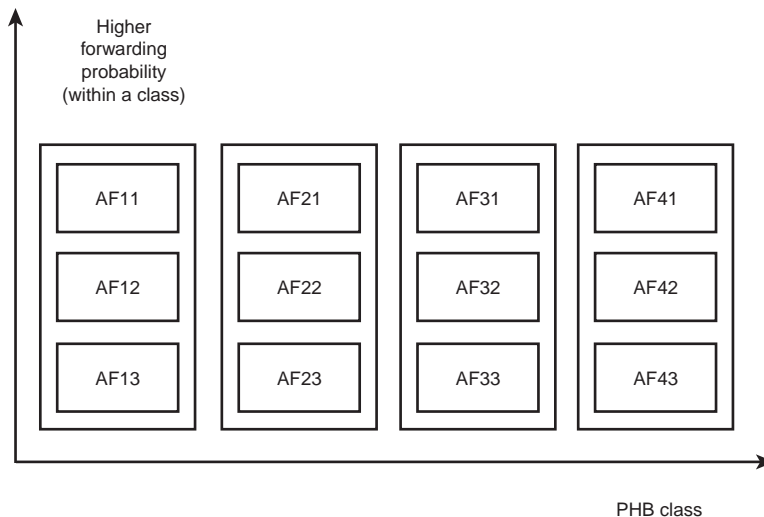
7.4 *Assured Forwarding PHB Group*

Assured Forwarding PHB (AF PHB) group is an integral part of Differentiated Services. AF PHB seems to reflect some of the inmost ideas of the Differentiated Services model. In particular, there is much freedom to implement different behaviors within the AF specification (although not necessary all at the same time). The other side of freedom is that the inherent logic of the AF system appears somewhat weak. This section discusses the diverse possibilities of the AF-PHB group as well as the difficulties in achieving consistent end-to-end services.

7.4.1 *Description*

In principle, the AF-PHB group can have a number of PHB classes (N), each with a number of drop precedence levels (M). The current specification defines that $N = 4$ and $M = 3$, but more classes or levels can be defined for local use (Baker *et al.* 1999). This section discusses only the basic AF structure with four classes and three drop precedence levels, as shown in Figure 7.13. Note that the wording of the AF specification indicates that it is not obligatory to implement all four AF classes. In addition, if the operator expects that congestion situations will be rare, an implementation with only two drop precedence levels is acceptable.

Figure 7.13 Structure of AF-PHB group.



The additional essential characteristics of the AF model are that a Differentiated Services node

- *Must* allocate a configurable, minimum amount of buffer space and bandwidth to each AF class
- *Must not* aggregate two or more AF classes together
- *Must not* reorder AF packets of the same flow when they belong to the same AF class, regardless of their drop precedence

It should be stressed that the target of AF is not any end-to-end service model, but an assemblage of tools to build services.

7.4.2 Position of AF-PHB Group in the Framework

The evaluation of the AF-PHB group is based on this author's best understanding, and tries to be as realistic as possible. Nevertheless, AF PHB, and particularly its applications, is such an ambiguous area that it is hard to make absolutely certain or clear statements.

Service Models

The inherent freedom of AF makes it possible to realize different service models based on applications, individual customers, or an organization. However, it seems that the relevance order is as follows:

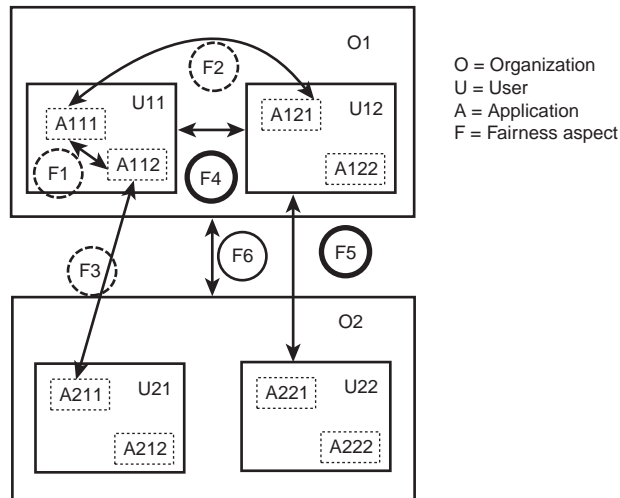
1. Customer model, in which the fairness considerations are based on the service-level agreement between customer and service provider
2. Organization model, in which organization makes the contract with a service provider, but the end user is basically free to use the bought bandwidth
3. Application model, in which flows are classified into different PHB classes based on the information about applications

The corresponding fairness issues presented in Figure 7.14 are summarized in Table 7.8; where *X* means primary fairness issue, *Y* means secondary fairness issue, and the hyphen (-) means inapplicable fairness issue.

Table 7.8 Relevance of Different Fairness Issues for AF-PHB Group

Service Model	F1	F2	F3	F4	F5	F6
Customer	-	-	-	X	X	-
Organization	Y1	Y1	-	Y2	-	X
Application	X	X	X	Y1	Y1	Y2

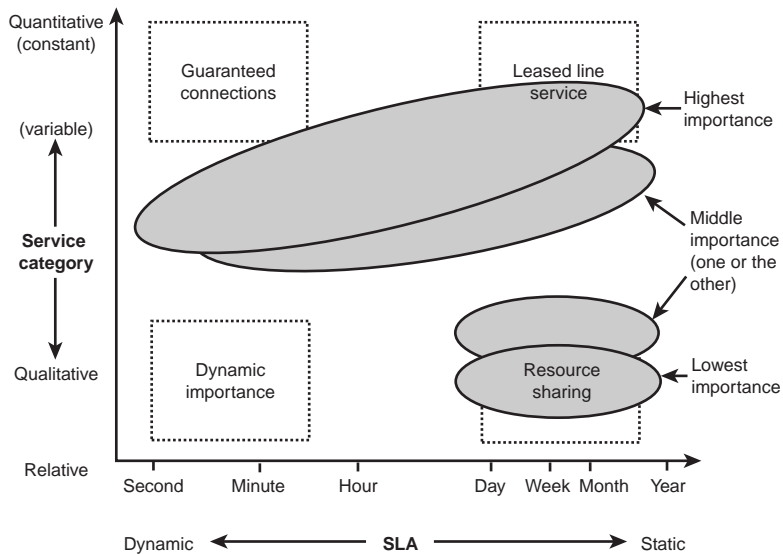
Figure 7.14 Relevant fairness issues for Assured Forwarding PHB group.



One of the practical problems of the AF-PHB system is that the fairness targets positively depend on the service model, and because different service models may coexist, even within one network domain, the overall system tends to become either complex or ambiguous (or both). Hopefully, this evaluation will help to clarify the situation.

The next issue is the relevance of different service models. This issue is, once again, open to various interpretations; Figure 7.15 shows two of them. In both options, the highest importance level provides a relatively high assurance of quality, and the lowest level is basically a best-effort service. In the first option, the middle level provides a somewhat lower assurance, but still the expectation is that the packet-loss ratio is small most of the time. In the second option, the two lowest importance levels together provide a better than best-effort service (for instance, with higher fairness than a basic best-effort service). Actually, “Assured Forwarding PHB Group” mainly promotes the second option because it requires that if a node supports only two drop precedence levels, two lower importance levels must be combined (Baker *et al.* 1999).

Figure 7.15 Possible service model for an AF-PHB class.



Supposing that the weight of each AF-PHB class is permanent, the level of assurance (vertical axis in Figure 7.15) depends on the dynamics of the SLA. The more permanent the SLA, the more likely that the network operator will have to provide truly assured service. In contrast, if customers can request short-lived assured connections, the network must be prepared for rapid and significant changes of load levels. If the weights cannot be adjusted according to the variable demand, the only possibility is to permanently reserve capacity of more than the average demand.

The lowest importance level of each PHB class forms a best-effort service that has essentially the same characteristics all the time, although the available capacity can change. It is hard to identify any need to request a change of best-effort service. As a result, the

dynamic importance of an AF-PHB system could mean that some parameters of the marking algorithm change based on user request. In other words, the threshold between best effort and better service can be changed for an individual customer, although the service itself remains the same. It is not clear, however, whether there are enough importance levels in an AF-PHB system for this purpose.

Moreover, note that Figure 7.15 primarily depicts the situation of one AF-PHB class; a natural question is how PHB classes may differ from each other. The main answer seems to be that PHB classes differ in the following ways:

- In the assurance level of highest importance level
- In the dynamics of the highest importance level
- In the role of the middle importance level
- In a quality aspect not presentable in Figure 7.15

Figure 7.16 shows one possible approach with three AFD classes. PHB1, PHB2, and PHB3 differ in each of the three first aspects. This figure mainly illustrates the possibilities; however, a system covering all possible aspects could be impractical to manage in real networks.

Figure 7.16 Possible service model for three AF-PHB classes.

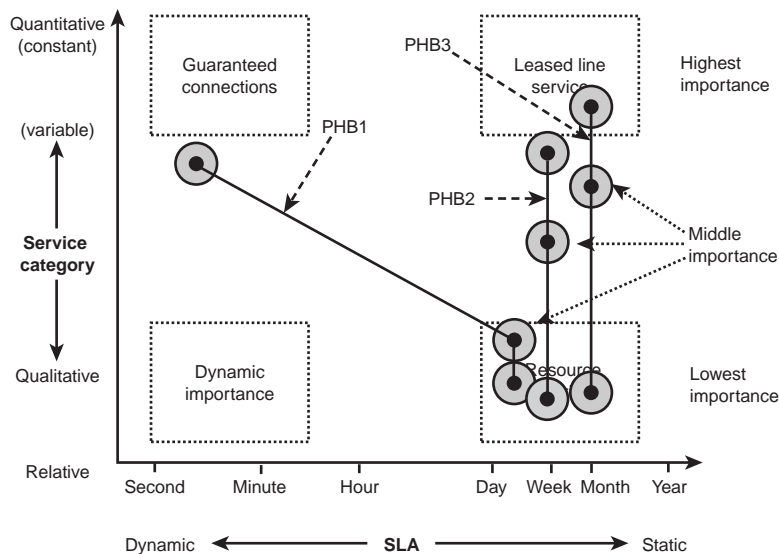
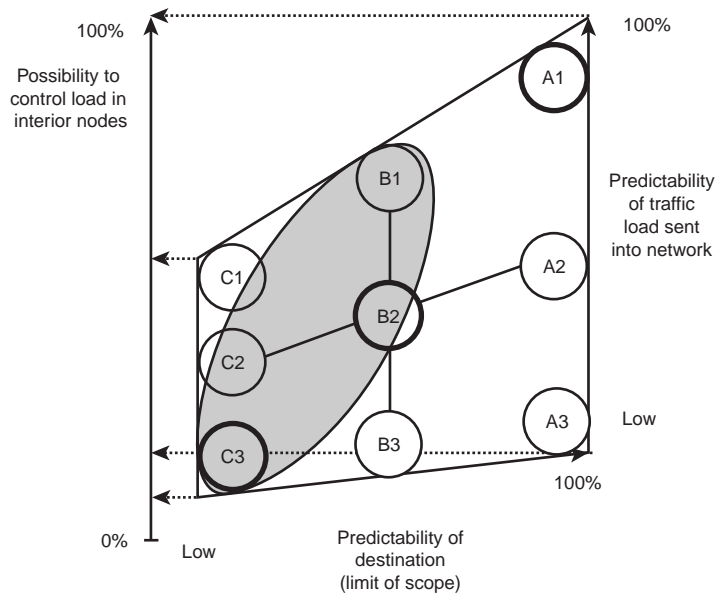


figure 7.17 further evaluates the same issue. In contrast to EF PHB, Assured Forwarding is likely operated without per-flow reservations and without taking explicitly into account the route through the network domain. This assumption limits the scope of AF to cases where the interior nodes cannot have strong control over traffic process even on the highest importance level. Consequently, the AF area in Figure 7.17 extends from best effort (C3 in Figure 7.17) to strictly controlled load with limited destination (B1).

Strictly controlled incoming load without any destination limits (C1) appears to be somewhat impractical, as does the limitation of destination without any load limitations (A3 and B3). Moderate destination and/or load control (C2 and B2) could be a suitable area to make experiments related to the applicability of the middle importance level. In practice, model B2 may mean a network domain with a small number of nodes and traffic control that is not as loose as with best-effort service but which still allows relatively large traffic variations.

Figure 7.17 Predictability of load and destination for AF-PHB group.



The last service-oriented issue is the applicability of AF-PHB groups to realize VPNs. Table 7.9 presents some preliminary considerations about this issue. The most relevant approach seems to be a combination of VPN-A2 and VPN-C, which means that the load level of the highest importance level is kept very low, and a limited capacity sharing is provided by using the other two importance levels.

Table 7.9 Suitability of AF-PHB Group for Different VPN Types

VPN Type	Main Tool to Manage Congestion Inside Network	Remark Considering AF-PHB Group
VPN-A1	Full mesh of CBR connections	Technically possible, but somewhat opposed to the principles of AF PHB.
VPN-A2	Strict limit for traffic sent to the network ($< C/(N-1)$)	Possible approach, but relevant only with the highest importance level.
VPN-A3	Alternative routing for load balancing	Not a primary approach.
VPN-B	Limited capacity sharing between VPNs	Perhaps the most realistic in principle, although the exact rules are unclear.
VPN-C	Total capacity sharing between VPNs	Three levels of importance could be insufficient for this approach.

Quality Models

Without defining the goal of the AF PHB, it is practically impossible to say anything about any quality issue. Therefore, it is necessary to select some service goals to evaluate the possibilities of an AF-PHB scheme, as follows:

- Delay differentiation
- Bit-rate differentiation
- Importance differentiation
- Improved predictability
- Pricing differentiation

The following sections discuss these service goals.

AF PHBs for Delay Differentiation The first idea is to suppose that each AF class represents a delay class with its own importance scale. Unfortunately, it is extremely difficult in practice to attain four discrete delay classes merely by regulating weights for aggregate streams. Why? Although any real proof is hard to provide, several issues support this statement in the case of Differentiated Services:

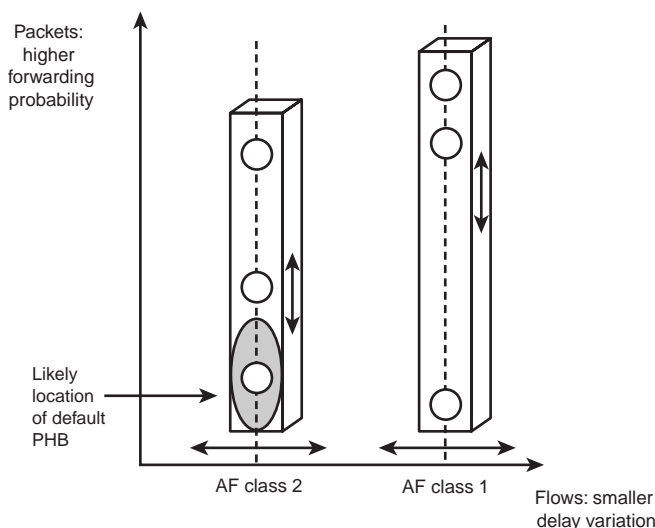
- Delay, especially maximum delay, is very sensitive to the traffic process. Even if two traffic processes have exactly the same average bit rate, the difference in maximum delay could be several orders of magnitude.
- The traffic process inside the network depends essentially on the intrinsic control mechanisms that regulate the incoming traffic. These mechanisms are largely (although not totally) out of the control of interior network nodes.
- The traffic process on different importance levels could be totally different. If this is the case, because the delay characteristics are common to all importance levels, the overall result could be an intractable problem.
- If the traffic on a certain class is tightly controlled (that is, smooth even on short timescales), there are significant delay differences only on high load levels. Consequently, if delays are controlled by regulating load levels, the process will be very accurate because tiny changes in load level—for instance, from 98% to 101%—may change the delay and loss characteristics abruptly.

In summary, even if the operator has advanced tools to control the weights, the characteristics of the four delay classes most probably will overlap each other most of the time. In contrast, it seems possible to realize two PHB classes with consistently different delay characteristics. Figure 7.18 illustrates this structure. But even in this case, it is somewhat difficult to identify any simple tool for determining the proper values for weights, or proper control mechanism to regulate the weights.

As a general rule, the relative load level compared to the bandwidth defined by the weight should be lower for the better importance class. This approach probably leads to a situation in which the packet-loss ratio of both the highest and the middle importance levels of the low-level PHB class are negligible. On the contrary, the packet-loss ratio of the lowest importance level could even be higher for the low delay class (PHB class 1 in Figure 7.18) than for the normal delay class 2.

The reason for this somewhat counter-intuitive result is that low delay class has to use relatively small buffers that probably cannot cope with the large burst of IP packets. Yet, it should be stressed that this reasoning is based on the assumption that the traffic on the lowest level is not tightly controlled; whereas with a proper traffic control, the situation could be totally different. If the traffic control at the boundary nodes keeps the packet size in PHB class 1 small and the bit-rate variations minimal, the packet-loss ratio can be insignificant even with small buffers. (Actually, the CBR service category in ATM networks demonstrates this property quite well.)

Figure 7.18 Two delay classes made by two AF PHBs.



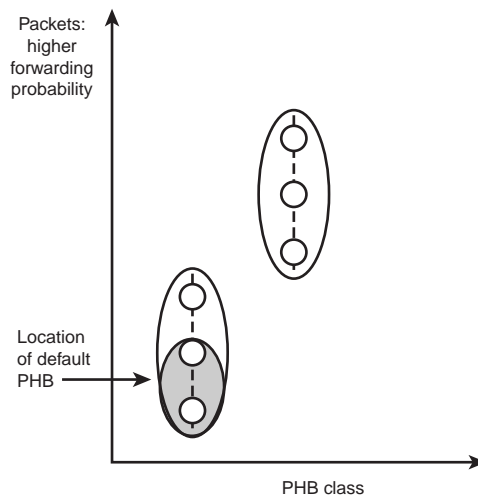
AF PHBs for Bit-Rate Differentiation The second idea for using AF PHBs meaningfully is to provide bit-rate differentiation. In a way, this is an attractive idea because bandwidth is one of the most important aspects of service differentiation.

For simplicity's sake, assume only two classes, one for customers with an expected bit rate of 50kbps and another class for customers with an expected bit rate of 250kbps. So far fine, but what do these bit rates really mean? The answer cannot be the bit rate of the highest importance level, because that bit rate is controlled in the boundary node and there is no evident reason to restrict the bit rate into one value per PHB class. The same reasoning is valid for the middle importance level as well. The only remaining class is the lowest one.

Then what is the purpose of knowing the expected bit rate for a flow inside the network? The primary purpose cannot be that the capacity of the lowest importance is divided proportional to the expected bit rate because that requires per-flow measurements in interior nodes. One, at least partly reasonable answer is that the information about the expected bit rate is used for management purposes. The network operator attempts to regulate the load levels and weights in a way that customers on average attain bandwidth proportional to the expected bit rates. If all four PHB classes are used for this purpose, it means four different expected bit rates for management purposes at the expense of nine extra code-points. (Note that an arbitrary amount of assured bit rates can be achieved merely by one PHB class.)

AF PHBs for Importance Differentiation The third imaginable option is to apply the PHB classes for additional importance levels. Figure 7.19 shows the target system. Even though the target itself can be justifiable, however, the implementation of this system could be a tricky task if you comply with the basic rules of AF—that is, the relationships among AF classes are regulated by the weights. Besides, because it is allowed to change the order of packets belonging to different AF classes, the whole range of importance values are not properly available for one individual flow. Therefore, it seems apparent that if more than three importance levels are needed, an AF class should be extended instead of using several AF classes.

Figure 7.19 Six importance levels made by two AF PHBs.



AF PHBs for Improved Predictability The former three approaches related to delay, bandwidth, and importance did not provide any evident purpose for all four AF classes. The fourth option is predictability. Perhaps one AF class is reserved for high predictability. What does that really mean? The target could be to keep the quality changes as slow as possible in such a way that the following occurs:

- The highest importance level permanently provides high quality (but that is expected anyway).
- The middle importance level permanently provides moderate quality (but that will be really difficult to realize in practice).
- The lowest importance level permanently provides the best-effort quality without starvation (but that is expected anyway).

The main conclusion of this preliminary assessment is that some level of predictability is a rational target for any AF class, although a useful differentiation seems difficult to realize. In particular, the predictability on the middle level likely requires an artificial deterioration of quality (excessive packet discarding or additional delays).

AF PHBs for Pricing Differentiation Finally, maybe it is necessary to omit all the previous considerations and just declare that AF classes are for pricing differentiation.

Unfortunately, it is not possible to totally avoid the quality or bandwidth issues, because there must be a reason for customers to pay more for one service than for another one. As section 5.3, “Pricing as a Tool for Controlling Traffic,” in Chapter 5 discussed, there is no prominent differentiation aspect that can be used as a basis for pricing. Even so, Table 7.10 presents some preliminary ideas about the possibilities of the AF-PHB group being used for pricing differentiation.

Table 7.10 Preliminary Pricing Considerations for AF-PHB Group

Differentiation of	As a Function of	One AF Class	Four AF Classes
Price	Bandwidth	Necessary at least for the highest importance level and perhaps for the middle level.	It does not seem reasonable to have different prices for purebandwidth in different AF classes if all other aspects are constant.
Price	Quality (delay)	Not relevant.	Two delay classes with different prices is a reasonable approach.
Price	Availability	Pricing may depend on the time of day and day of the week.	Different AF classes may have different pricing structures (for instance, because class 1 is for residential users and class 2 is for business users).
Availability	Bandwidth	Possible in a limited sense (more bandwidth with the same price during idle hours).	As above.

Differentiation of	As a Function of	One AF Class	Four AF Classes
Availability	Quality	Not relevant.	Possible to a limited extent because customers are allowed to change AF class.
Quality	Bandwidth	Not relevant	As above.

7.4.3 *Required Technical Tools*

Once again, the problem of uncertainty arises. When no clear idea about the general objective of the system is available, it is impossible to strictly evaluate the required mechanisms. The following evaluation is based on certain assumptions that are deemed relevant, but that may be proved erroneous in real networks.

Implementation Options for Boundary Nodes

Suppose that there are two reference bit rates: a lower one for controlling the traffic on the highest importance level, and a higher one for controlling the traffic on the middle importance level. Another possible difference between the two importance levels is that the measuring period for the middle class is longer than the measuring period for the highest importance class. If there is no premarking made by the customer, the boundary nodes mark as many packets as possible into the highest importance level according to the rules defined in the SLA. Correspondingly, if the packets cannot attain the highest importance level, the middle importance level is used, if possible. Only if both of these are impossible is the packet marked on the lowest importance level.

In summary, the philosophy of AF seems to be that only excessive packets are marked, which means that regardless of the used bit rate on the highest importance level, the user is allowed to send as many packets on the lowest importance level as wanted. However, the SLA may also include restrictions to send an excessive amount of packets on the lowest importance level. All these properties can be implemented, for instance, by three token buckets with different bucket rates and bucket depths.

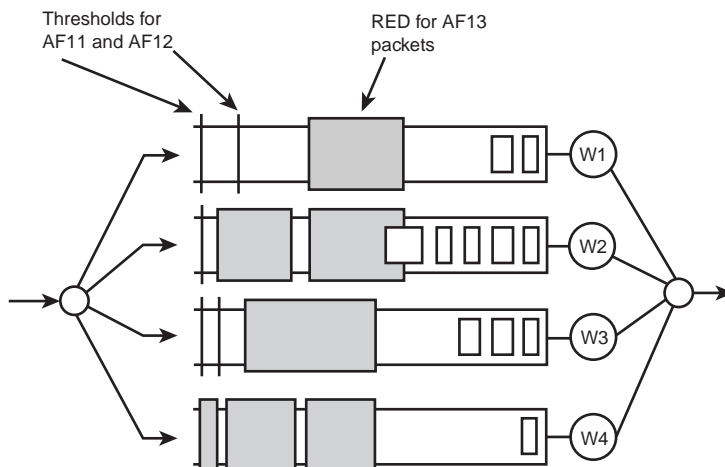
Implementation Options for Interior Nodes

Although the AF specification “Assured Forwarding PHB Group” does not mention any specific queuing system, the AF-PHB structure resembles the class-based queuing (CBQ) system (Baker *et al.* 1999). In particular, a configurable amount of buffer space and bandwidth must be allocated to each of the four AF classes—this can hardly be done without

four real or virtual queues. Bandwidth allocation for different queues is exactly what CBQ is supposed to do. Nonetheless, vendors are allowed to use any queuing system that satisfies the AF-PHB requirements.

The other integral part of AF-PHB implementation is the algorithm used to decide which packets have to be dropped during overload situations. The specification states that an AF implementation must attempt to minimize long-term congestion within each class. This seems to require an active queue-management algorithm, such as *Random Early Drop (RED)*. An open issue is whether RED or a similar algorithm should be used for all thresholds or only for the lowest importance levels. Figure 7.20 shows an implementation with four queues and different combinations of hard thresholds and RED.

Figure 7.20 AF implementation based on four queues and three importance levels.



Implementation Options Related to Network Domain

Table 7.11 presents a preliminary assessment of the management aspect related to the AF-PHB group. The main concern is related to the network dimensioning in general. If the use of AF-PHB classes is more or less unclear, it is extremely difficult for the vendors to develop appropriate supporting systems for network operation and management.

Table 7.11 Preliminary Management Consideration for AF-PHB Group

Aspect	Main Options with AF
Routing	Primarily static routing.
Resource Reservation	The primary idea of AF is to use permanent or semipermanent reservation of traffic aggregates.

Aspect	Main Options with AF
Network dimensioning	Because mathematical modeling seems to be very difficult, network dimensioning will probably be based instead on practical experience.
Contracts between domains	The primary option is to apply appropriate packet marking, but some level of compensation could be needed if the load is not balanced.

7.4.4 *Evaluation of Attributes*

The traffic-handling part of the AF-PHB system is relatively simple, and hardly results in high implementation costs. Moreover, if properly managed the statistical multiplexing could be efficiently utilized with a couple of clearly defined AF-PHB classes. However, the costs related to network management and customer care are difficult to predict when the service model is not known. The same concern is valid also with fairness—an AF system may provide tools to offer good fairness, but that is not an evident situation.

In a sense, the AF system provides considerable versatility because a large variety of service models can be implemented. It is even fair to say that there is too much versatility, because there is no commonly accepted service model for AF that makes it possible to design useful end-to-end services. Robustness also depends on the general effectiveness of the AF system; a systematic approach is definitely needed to avoid ambiguous situations that endanger the robustness of the system.

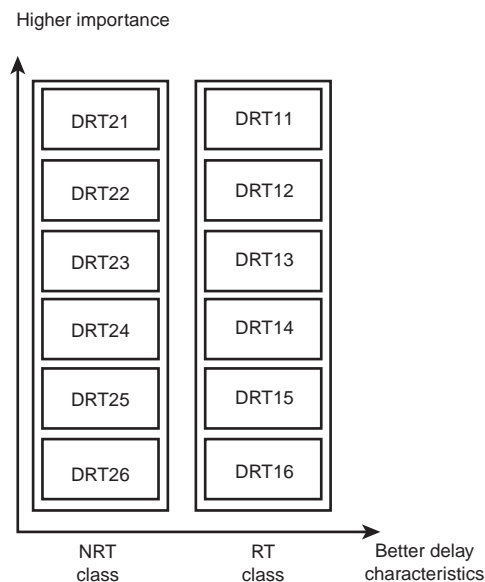
7.5 *Dynamic RT/NRT PHB Group*

The main target of Dynamic RT/NRT (DRT) PHB group is to provide a consistent and clear framework for building Differentiated Services (Kilikki, Loukola, and Ruutu 1998). Although the DRT-PHB specification is an independent document, the roots of the specification are in the more extensive Simple Integrated Media Access (SIMA) model (which can be found at <http://www-nrc.nokia.com/sima/>).

7.5.1 *Description*

The DRT-PHB group defines a system with two PHB classes and six PHBs in each of the classes depicted in Figure 7.21. PHB classes offer two distinctly different delay characteristics: RT class is for flows needing real-time service, and NRT is for flows without strict delay requirements. Six importance levels offer wide dynamics for various traffic-control and pricing schemes. The framework can be extended in both directions if necessary; there could be more than two delay classes and more than six importance levels. The values 2 and 6 are considered as the minimum to provide appropriate characteristics for various needs.

Figure 7.21 Structure of DRT-PHB group.



7.5.2 Position of DRT-PHB Group in the Framework

The evaluation of the DRT-PHB group is based on this author's opinion and on the original idea behind the DRT-PHB group. That basis is both an advantage and disadvantage. On the one hand it makes the assessment, in a way, easy; on the other hand, there is risk that the assessment of the PHB group is somehow partial (or subjective). In particular, it is comprehensible that the DRT-PHB group fits better in the framework presented in earlier chapters than the other PHB groups, because the PHB model and the framework are based on the same insight of Differentiated Services.

This section tries to give, after all, a relatively impartial assessment of the DRT-PHB model. Nonetheless, personal preferences always have substantial effects on the conclusions, unless something can be strictly proven (and unfortunately, mathematical proofs are not possible in this case). Therefore, the conclusive comparison of different PHB models remains open as far as there is no extensive experiments of different PHB models.

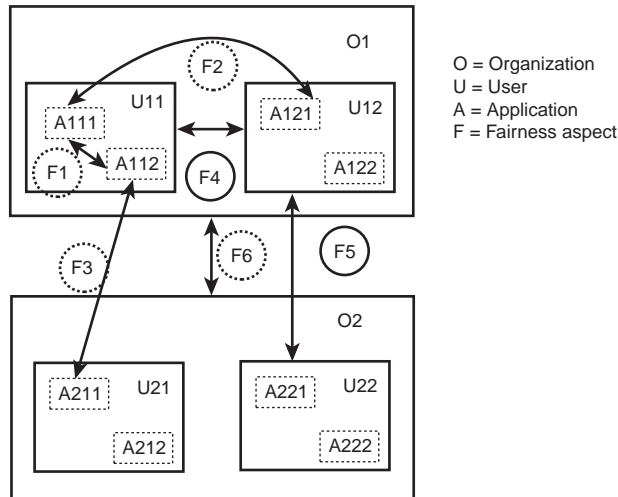
Service Models

The DRT-PHB group is essentially, as all PHB groups should be, a specification of packet treatment inside the Differentiated Services network. Yet, to assess the possible services, it is necessary to specify the boundary functions to some extent. The original system behind

the DRT-PHB model—that is, Simple Integrated Media Access (SIMA)—includes specifications for both the boundary functions and the interior functions, and even for the pricing structure (found at <http://www-nrc.nokia.com/sima/>). Although the DRT-PHB group is not limited to the SIMA system model, it is used to assess the possible services build on the basis of the DRT-PHB group, because the SIMA model comprises in a systematic manner all the aspects from customer services to packet-level mechanisms.

One of the key concepts of the SIMA model is *nominal bit rate (NBR)*. NBR defines the relative amount of resources that a certain entity is supposed to achieve from the network. The entity can be anything: a part of a flow, flow, customer, group of customers, or a large organization. In that sense, the DRT-PHB group can be applied to any of the three service models: application, customer, or organization model. Yet, the basic philosophy of the DRT-PHB group mainly supports the customer model (in which each customer buys as big a share of the network as she wants and is ready to pay for). Therefore, the primary fairness relations addressed by the DRT-PHB group are those between users—that is, F4 and F5 in Figure 7.22.

Figure 7.22 Relevant fairness issues for the DRT-PHB group.

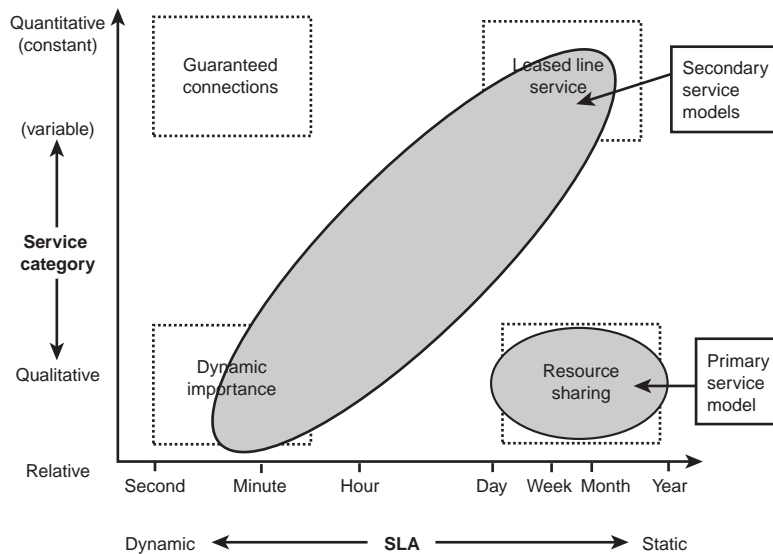


The next issue is the SIMA service in relation to dynamics and level of assurance. The primary model of SIMA—and this is generally valid with the DRT-PHB group as well—is resource sharing, as illustrated in Figure 7.23. NBRs (that is, the shares of customers) should usually be as static as possible to facilitate the network management.

Nevertheless, the DRT-PHB framework can support other models as well: Leased-line service basically needs a high enough importance level (perhaps beyond the six importance

levels of DRT PHB). A constant bit-rate connection could have a special traffic control in the boundary node similar to that of EF PHB. Furthermore, if NBRs are dynamic rather than static, the service provider can use a dynamic importance model. Note, however, that increased dynamics make load prediction and network dimensioning more difficult. Guaranteed connections need additional mechanisms that are not covered by the specification of the DRT-PHB group. It is, therefore, likely that increased dynamics will be used in real networks only after significant experience is gathered about static quality differentiation.

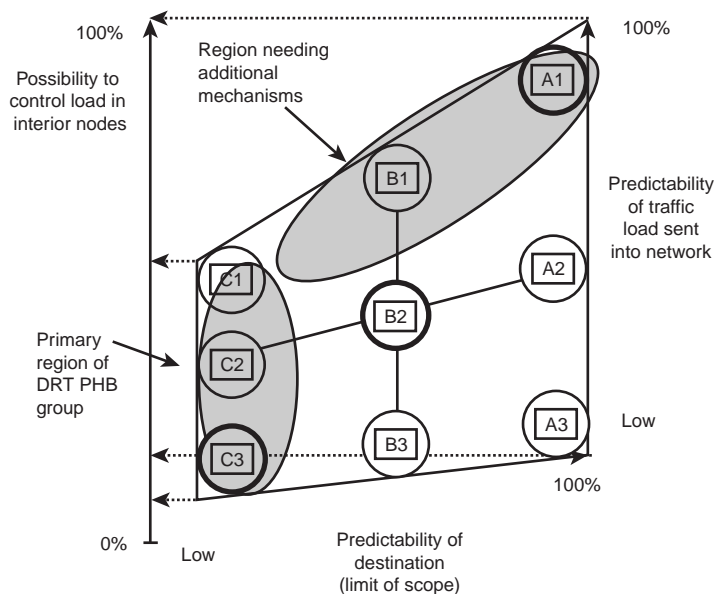
Figure 7.23 Primary and secondary service models for DRT-PHB group.



One of the basic assumptions of the DRT-PHB group is that most of the traffic on the Internet consists of short-lived flows to variable destinations. The predictability of destination is intrinsically low for that kind of traffic, and there is not much to be done to improve the situation. The architecture of DRT PHB is adapted to that situation. There is no obstacle to using DRT PHB with long-lived flows, however, such as IP telephony and video streams.

The six (or more) levels of importance cover the scale from very low predictability of load to high predictability. Even on the highest level, the assumption is not a constant bit rate, but rather the reasoning is that the higher price keeps the use of the highest level limited (particularly if the NBR is static). On the contrary, the lowest importance level could be available without any restrictions, which makes it difficult to predict load level and traffic process inside the network. Figure 7.24 shows both primary service models and the most important region that is not covered directly by DRT-PHB group.

Figure 7.24 Predictability of load and destination for DRT-PHB group.



The main VPN model for the DRT-PHB group is VPN-C in Table 7.12. That is, all relevant marking is done at boundary nodes, and the interior nodes are not aware of the VPN itself. The actual result depends on the marking rules implemented in the boundary nodes, and in some cases the result could be similar to VPN-B. Moreover, some simple additional rules can be used in interior nodes to increase the separation between different VPNs (model VPN-B in Table 7.12). VPN-A is also possible, although it clearly requires additional tools that are beyond the scope of the DRT-PHB group.

Table 7.12 Suitability of the DRT-PHB group for Different VPN Types

VPN Type	Main Tool to Manage Congestion Inside Network	Remark Considering DRT-PHB Group
VPN-A1	Full mesh of CBR connections	Not relevant with DRT-PHB group.
VPN-A2	Strict limit for traffic sent to the network ($< C/(N-1)$)	Could be used, but presumably not a model.
VPN-A3	Alternative routing for load balancing	Not a likely model.

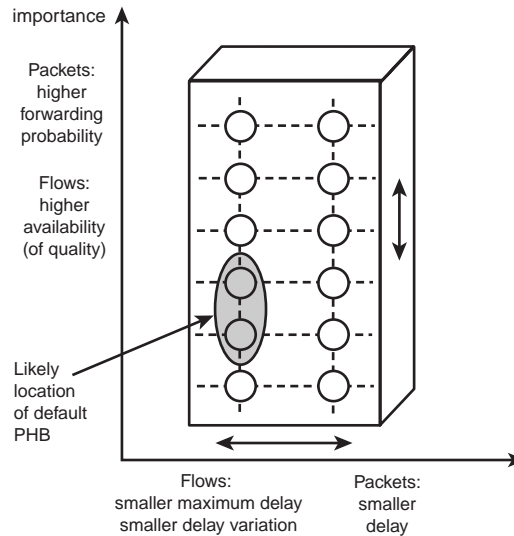
Table 7.12 Continued

VPN Type	Main Tool to Manage Congestion Inside Network	Remark Considering DRT-PHB Group
VPN-B	Limited capacity sharing between VPNs	Can likely be done by combining NBR management in boundary nodes and weights for VPNs in interior nodes (but implementation is not trivial).
VPN-C	Total capacity sharing between VPNs	Can be done by appropriate managing of NBRs in boundary nodes. (This model is in accordance with the basic philosophy of DRT PHB.)

Quality Models

If there is something really original in the DRT-PHB group it is the idea to fasten the importance levels of two PHB classes together. Figure 7.25 illustrates this. Although the actual packet-loss ratio of an importance level depends significantly on the load level and traffic processes, the instantaneous value for an importance level is independent of the service class. Consequently, the whole package of 12 PHBs moves in the scale defined by delay and loss ratio in a way that relations of each PHB pair remain the same regardless of the load situation in the network. This is the target model of the DRT-PHB group.

Figure 7.25 Quality model of the DRT-PHB group.



Predictability with DRT-PHB Group

Because of the basic principles of the DRT-PHB group, the predictability of quality can be limited, in particular on all intermediate importance levels. If high predictability is positively needed, it seems that some additional mechanisms are necessary. On the highest quality level, that could mean an EF type of PHB model.

The other relevant predictability target is that default PHB should not be starved even in the case of a high load level. This issue can be divided into two parts: the treatment of traffic flows *with* and *without* the possibility to obtain a higher importance level by reducing the bit rate of the flow. The first case with the possibility is trivial: Every flow can avoid starvation just by reducing its bit rate. If the application cannot adjust the bit rate, it should use a higher importance level rather than suppose that the default PHB provides permanently high enough quality.

The other case without the possibility is more difficult. The primary approach is to avoid mapping those streams into PHBs with a high possibility of starvation. (For this reason, default PHB is not the lowest importance level in Figure 7.25). Another approach is to guarantee that the importance level used by the default best-effort service always attains a minimum bandwidth. Because this kind of system violates the fundamental logic of the DRT-PHB system, however, it should be avoided whenever possible. Finally, default service

may use its own queue with a more or less permanent weight. In that case, the importance relationship between packets belonging to default PHB and DRT PHB can be defined only approximately.

Pricing Models

One of the potential advantages of a consistent PHB architecture is that it makes it possible to design comprehensible relationships between different service aspects and price. Table 7.13 gives a short overview of some of possibilities provided by the DRT-PHB group.

Table 7.13 Preliminary Pricing Considerations for the DRT-PHB Group

Differentiation of	As a function of	DRT-PHB Group
Price	Bandwidth	The price of NBR defines the relationship between bandwidth and price. (Note that the price of NBR has no effect on the mechanisms inside the network.)
Price	Quality (delay)	For a constant packet flow, there is no explicit pricing difference between the two delay classes. On the contrary, for a given variable bit-rate flow, higher NBR (that is, higher price) is needed to attain the same importance level if the customer uses the RT class rather than the NRT class.
Price	Availability	This is one of the inherent relationships of the DRT-PHB system: The availability of given quality and bandwidth depends automatically on the load situation. That means that what the system provides is different between busy and idle hours without any additional management effort.
Availability	Bandwidth	Basically the same relationship as above: If the price and quality are constant, a lower bit rate automatically means higher availability, and vice versa.

Availability	Quality	Possible in a limited sense: For a variable bit-rate flow, a change from RT PHB to non-RT PHB may improve the availability.
Quality	Bandwidth	Possible in a limited sense: For a variable bit-rate flow, a change from RT PHB to NRT PHB may allow the use of a higher average bit rate.

7.5.3 *Required Technical Tools*

The following investigation of required technical tools is based on the assumption that the DRT-PHB group is mainly used to implement a resource-sharing model. If the service model is different, the requirements could be somewhat (but probably not much) different.

Implementation Options for Boundary Nodes

The main, and almost the only requirement for boundary node functions is that the bit rate of each entity with an NBR be measured; and based on the measurement result, the boundary node determines the importance level for each packet. It is also possible to use some other, totally different approaches. For instance, the boundary node can mark the packets merely based on the application. That kind of systems violates the fundamental resource-sharing principle of the system, however. If the marking does not depend systematically on the bit rate used by the customer, the interior nodes cannot divide the capacity in a fair manner among customers.

The primary marking system for DRT PHB is that the instantaneous metering result defines entirely the packet marking of every packet—that is, as long as the metering result remains the same, all packets obtain the same importance marking. Nevertheless, it seems possible to develop a more complicated system in which the marking depends also on other issues, such as the packet marking made by the customer. This kind of system makes it possible to support, for instance, a layered video coding with more and less important frames.

Finally, it should be noted that because the RT-PHB class with relatively small buffers is more sensitive to bursts of packets (and large packets as well), the marking system should somehow depend on the PHB class selected by the customer. That can be achieved in practice by a short measuring period that punishes real-time flows with high burstiness or very large packet size.

Implementation Options for Interior Nodes

Figure 7.26 presents one possible implementation for the DRT-PHB group. Two queues and six thresholds are basically inevitable system components. In addition, two important principles largely define the system behavior:

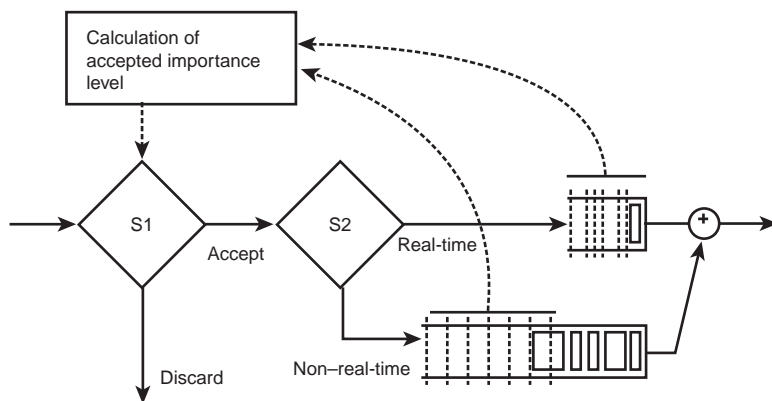
- The RT-PHB queue has a strict priority over the other queue.
- The packet-discarding decision is made independent of the PHB class.

The first principle provides an easy method to guarantee a maximum delay for the RT PHB in a way described in the section “Priority Queuing” in Chapter 6, “Traffic Handling and Network Management.” A mere strict priority, however, may result in a situation in which the RT-PHB queue can use the whole link capacity while the NRT-PHB queue becomes totally starved. This is a critical problem and should be solved somehow.

One possibility is to define a minimum amount of link capacity for the NRT queue—that is, a finite weight. The main difficulty of this approach is that there is no clear relationship between the importance levels of the two queues. Depending on the load situation of the two queues, at an arbitrary point in time the system may discard RT packets of importance level 4, and at the same time accept NRT packets of importance level 6. When the load levels of the two queues change, the situation could be totally reverse after a while.

This kind of behavior is not desired in a DRT-PHB system. Therefore, the DRT-PHB system uses another alternative to avoid starvation of the NRT queue. In DRT PHB, the discarding of a packet is independent of the PHB class. In consequence, the packet-loss probability at a given point in time depends on the importance level, but it is virtually independent of the PHB class. If there is any starvation problem, it is common for a certain importance level of both RT and NRT classes.

Figure 7.26 An implementation of the DRT-PHB group.



Implementation Options Related to the Network Domain

Table 7.14 presents a preliminary assessment of the management aspect related to the DRT-PHB group. The main concern is related to the network dimensioning. Without appropriate network dimensioning, the services build on the basis of the DRT-PHB group cannot be consistent. If the load level in one part of the network is much higher than that of another part, for instance, the average share attained by the same price could vary significantly. That could deteriorate the fairness of the service among customers in different parts of the network. Moreover, customers will probably expect that the available share be at least somehow predictable, if not constant. Large and quick variations can also be defective for customer services.

Table 7.14 Management Considerations for the DRT-PHB Group

Aspect	Main Options with DRT-PHB Group
Routing	Primarily static routing. If adaptive routing is used, it should take into account separately the load level of every individual PHB (not only the average load level).
Resource reservation	The fundamental idea of the DRT-PHB group is to avoid explicit resource reservation whenever possible.
Network dimensioning	This is the most critical part of the DRT-PHB model, because the customer service relies on the assumption that the share bought by the customer provides a predictable and stable bit rate even for arbitrary destinations.
Contracts between domains	Both pricing and re-marking of packets are possible approaches. In particular, pricing should take into account the importance levels of each packet. Re-marking should be as stable as possible—for instance, the importance level of all packets coming from a network domain are decreased by one step for a relatively long duration.

7.5.4 Evaluation of Attributes

A simple and consistent system such as the DRT-PHB group is likely to provide highly efficient implementation on a traffic-handling level. The main concerns are traffic control and network management. The control mechanisms must be able to effectively prevent the misuse of high importance levels. The management system should support a workable network dimensioning and operation system that appropriately takes into account the requirements of both PHB classes and all importance levels.

The fairness of the service provision depends decisively on the NBR management and pricing structure. The logical structure of the DRT-PHB system itself gives tools to realize a fair service based on the resource-sharing model. Versatility is achieved by both delay classes and by wide dynamics that make it possible to apply exquisite pricing schemes. Robustness is appropriate if the network operation and management gives appropriate support for customer care and network management.

Summary

The introduction to this chapter defined the essence of PHB—that is, the rule that determines the treatment of packets inside the network. The rule should be useful and comprehensible to make it possible to build practical services. But the rules should not necessitate a specific implementation nor should they exactly specify the service model. This discussion reviewed four PHB models: Class Selector, Expedited Forwarding, Assured Forwarding and Dynamic RT/NRT. Did they meet the fundamental targets?

The Class Selector PHB group lacks clear rules because it is mainly intended to provide backward compatibility for current, various uses of the DS octet. Expedited Forwarding appears to be a service specification rather than a PHB. This is in a sense a significant advantage; on the other hand, however, the EF specification does not considerably advance the development of the general-purpose framework for Differentiated Services.

Assured Service defines requirements for mechanisms in interior nodes rather than a clear PHB structure from a packet or flow perspective. Although the AF-PHB model may allow the operator to introduce numerous new services, it does not give much guidance for interoperability. Different service providers may understand the (largely hidden) AF rules in different ways.

The Dynamic RT/NRT-PHB model satisfies the goal better; the essence of the DRT specification is about the relationships among packets belonging to different PHBs. Logical behavior of the PHB system requires that the boundary mechanisms comply with certain rules, however. In particular, if the importance level of each packet does not depend systematically on the bit rate of the flow, the whole system model is somewhat questionable.

It is apparent that the current PHB specifications as such do not provide one clear way to the utopia of Differentiated Services. It seems that the reality is too complicated and too many matters resist any significant, quick changes in the Internet technology and service models. Because there is such strong pressure on service evolution, however, some significant changes will certainly happen within the next few years. This chapter's evaluation of PHBs offers some tools to direct the evolution toward reasonable goals.