

Internet Engineering Task Force
Internet Draft
Expiration Date: December, 1998

Constantinos Dovrolis
University of Wisconsin,
Madison
June, 1998

Class-Based Service Differentiation

[draft-dovrolis-cbsd-00.txt](#)

Status of this Memo

This document is an Internet-Draft. Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

To view the entire list of current Internet-Drafts, please check the "1id-abstracts.txt" listing contained in the Internet-Drafts Shadow Directories on ftp.is.co.za (Africa), ftp.nordu.net (Northern Europe), ftp.nis.garr.it (Southern Europe), munnari.oz.au (Pacific Rim), ftp.ietf.org (US East Coast), or ftp.isi.edu (US West Coast).

Abstract

This document describes the Class-Based Service Differentiation (CBSD) architecture and it proposes a set of Per-Hop-Behaviors that can be used for implementing it. According to CBSD, the network traffic is mapped to a set of classes. The differentiating factor between traffic classes is the relative forwarding quality in a DS capable router. The Class PHBs have quite similar semantics with the Precedence PHBs, recently proposed by Baker et.al. in [[PREC](#)], but an attempt is made here to define them in a more general manner, decoupling them from specific implementation possibilities. There is also a semantical overlap between the Expedited Forwarding PHB of [[DSFIELD](#)] and the highest Class PHB. Consequently, the Diffserv working group may want to combine these three PHB proposals in a single one, minimizing the number of DS Field bits used.

Constantinos Dovrolis

Expires: December 1998

[Page 1]

1. Introduction

The IETF Diffserv Working group is working towards a differentiated services architecture for the Internet. The ultimate goal of this effort is to come up with effective and scalable service discrimination for the Internet users, without departing from the original IP network model of stateless packet forwarding. The Diffserv approach targets on defining Per-Hop-Behaviors (PHBs) and not end-to-end services. This is probably a wise decision, since there is currently not much knowledge and experiences about the types of end-to-end services that are required today, and even more, that will be required in the future. A PHB has some well-defined semantics about the forwarding behavior at a DS-capable router. The set of PHBs can be viewed as a set of "legos"; as the experiences or the requirements about differentiated services evolve, we can build a large set of services given only a small set of PHBs.

In this document we propose a set of PHBs, called Class PHBs. The semantics of the Class PHBs refer to the relative service level that a packet will get at a DS capable router. In other words, a Class PHB is meaningful only compared to another Class PHB. The Class-I PHB provides a "better service" than any Class-J PHB, with $I > J$. The term "better service" is deliberately loosely defined, and it should be interpreted only in a statistical manner, and only locally at a DS capable router. One way to view the "better service" differentiation is in terms of traffic load, i.e., the Class-I PHB provides in most of the time less loaded forwarding than any Class-J PHB ($I > J$). Alternatively, the term "better service" can be interpreted based on performance metrics such as queuing delay or drop probability, i.e., the Class-I PHB treats packets in most of the time faster and more reliably than any Class-J PHB ($I > J$). The "better service" constraint is not quantitatively defined, so that network providers can have the flexibility to attach to it different performance ratings, depending on the implementations they use and their traffic management policies. A goal of this document is to decouple the relative service level that a traffic class gets, from semantics that refer only to timely delivery (i.e., priority), or only to relative importance (i.e., drop preference).

In [Section 2](#), we argue that the CBSD architecture follows naturally from the end-to-end principle. [Section 3](#) proposes the Class PHBs. A comparison with the Precedence PHBs of [[PREC](#)] and with the EF PHB of [[DSFIELD](#)] follows in [Section 4](#). Finally, a possible implementation for the Class PHBs is sketched in [Section 5](#). We attempt to be consistent with the terminology and the general framework described in [[DSFIELD](#)], [[DSARCH](#)], and [[DSFRMW](#)].

Constantinos Dovrolis

Expires: December 1998

[Page 1]

2. The End-To-End Principle in the Diffserv Context

We can design different types of network service differentiation to match the specific needs of current or emerging applications. For example, IP telephony applications require low jitter and low loss transfer. Packet video applications require high throughput transfer, but they allow higher drop rates for the less important layers of the video stream. Some transaction-based applications require very low end-to-end delays. Such application requirements can be met by deploying the appropriate network-level service differentiation mechanisms. For example, the packet video requirements may be met by providing different levels of drop preference. In this way, we attempt basically to make the network follow the application requirements. It is generally accepted however, that this approach is not flexible and extendible, as new applications with new requirements keep emerging. For example, it is hard to predict at this point the requirements of virtual reality applications. Besides, the lesson learned from adaptive multimedia applications is that applications can always be built to best match whatever service the network provides, without requiring hard network performance guarantees.

An alternative approach would be to provide network service differentiation in terms of network performance metrics, instead of application requirements. In other words, in this approach it is the network that "makes the rules", by defining a service level hierarchy that is general, simple, and application independent. The applications and the users, then, match their needs to specific levels of this hierarchy, based on their performance and cost constraints. The CBSD architecture attempts to follow this approach, as illustrated in the next examples.

An IP telephony application can select while running the service class that best meets its jitter and loss requirements, given also some pricing or policy constraints. Suppose, for example, that two friends that have Internet access only up to Class-3 have an IP phone conversation. The application starts the session at the lowest class (Class-1) and gradually switches to higher classes until the service level is acceptable. After a few seconds the application converges to Class-2, where it finds the service level to be adequate. At the same time, at the same network, and using the same application, two CEOs also have a IP phone conversation, but this time with tighter quality requirements, and without any constraints on the cost of the call. The application this time starts immediately from Class-4, to save time, and finally converges to Class-6 where the performance requirements are met.

Constantinos Dovrolis

Expires: December 1998

[Page 3]

Similarly, consider a packet video application. The application generates three layers of video packets, A, B, and C, that differ in their relative importance for the final video stream (say that A is the base layer). After a few seconds, the application converges to use Class-4 for the A packets, Class-3 for the B packets, and Class-1 for the C packets. Again, this selection is a function of the pricing or policy constraints. Also, depending on the network dynamics, the application may find out later that it can switch the A packets to Class-3, or that it has to switch them to Class-5, instead.

Consider finally a corporate user that wants to get adequate service from a network provider for the company's VPN. The two parts can make, for example, a service level agreement for a certain amount of Class-5 traffic, and for unlimited Class-2 traffic.

3. Class PHBs

We are proposing a set of Class PHBs, defined in the following way:

A Class-I PHB provides a better forwarding service than any Class-J PHB, for $I > J$.

As discussed earlier, the exact meaning of the term "better service" can be viewed in terms of traffic load, or in terms of local performance metrics, such as packet delay or loss. Also, in the absence of admission control or signaling, it is likely that the term "better service" can be only interpreted in a statistical manner, i.e., in most of the time, or for most of the serviced packets, Class-I gets a better service than Class-J. Network providers may want to attach to the above definition quantitative semantics, depending on the implementation mechanisms they use and their traffic management or marketing policies. For example, in a hypothetical implementation the Class-I PHB may lead with a 99% probability to at least two times smaller queuing delay and drop rate than the Class-(I-1) PHB.

A first issue is the appropriate number of classes. In this document we propose eight classes. This is the same with the number of IPv4 precedences, IPv6 priorities (now obsolete, [[IPv6](#)]), or with the number of priorities in some link level technologies, such as IEEE 802.1d. Obviously, there is nothing magic in this selection, and it may be shown in practice that this number should be smaller or larger. Further experimentation in large networks is required for a better decision on the number of classes.

A second issue is how much better should the service of a class be compared to the service of the directly lower class. This is also an implementation issue that can vary between network providers. A possible constraint, however, can be that in the worst case, each traffic class should get on the long run

about the same level of service as if it was processed in a FIFO manner by a non-DS capable router.

Constantinos Dovrolis

Expires: December 1998

[Page 4]

For the eight classes architecture, we need eight PHBs. The selection of bits in the DS-field can be arbitrary, but since legacy routers already use the DS field bits, it is more appropriate to map the class PHBs to the IPv4 three precedence bits of the TOS byte. Since [[IPv4](#)] defines precedence seven as the highest one, we map the PHB for the best class (Class-7) to the precedence seven value. Specifically, the following PHBs are proposed:

000 XXX	Class 0 (worst-service class)
001 XXX	Class 1
010 XXX	Class 2
011 XXX	Class 3
100 XXX	Class 4
101 XXX	Class 5
110 XXX	Class 6
111 XXX	Class 7 (best-service class)

Other PHB selectors can be combined with the CBSD PHBs to produce further service differentiation. Alternatively, it may be better to reserve some "don't care" bits for defining additional service classes later.

[4. Comparison with the Precedence and Expedited Forwarding PHBs](#)

Recently, Baker et.al. proposed a set of PHBs for transferring the IPv4 Precedence semantics in the Diffserv architecture [[PREC](#)]. If we map each class to a precedence, the Precedence PHBs seem to be the same with the Class PHBs. In [[PREC](#)], however, or in the original specification of the IPv4 Precedence bits in [[IPv4](#)], the semantical or "forwarding behavior" differences that the Precedence bits specify are not clear. [[PREC](#)] states that "In essence, a higher precedence (queue or class number) should afford a higher probability of timely delivery than a lower precedence packet (...)" . We believe that this statement does not define the same service differentiation as the CBSD architecture. This becomes clear if we consider the four potential implementation mechanisms that [[PREC](#)] suggests. We show next that neither of these mechanisms can actually implement the Class PHB semantics.

The first possible implementation in [[PREC](#)] is a FIFO queue where the precedence is directly mapped to the drop priority, meaning that lower precedence packets are dropped before higher precedence packets in the presence of congestion. This mechanism does not provide lower queuing delays for the higher precedence packets, however, when there is persistent queuing but not significant packet drops. Consequently, the higher precedence packets can consistently get "worse service" than packets from lower precedences.

Constantinos Dovrolis

Expires: December 1998

[Page 5]

The second possible implementation in [PREC] is a set of priority queues, where higher precedence values correspond to higher priority. This configuration, however, allows for the lower priorities to starve and effectively get the same low quality service. In the CBSD architecture this is not allowed, since every class has to provide a better service than the directly lower class.

The third possible implementation in [PREC] is a round robin scheduler, where a certain rate is assigned to the queue servicing the traffic with a certain precedence. However, as the authors of [PREC] also note, if a lower precedence queue is less loaded than some higher precedence queue, the service level of the higher precedence will be actually worse. Obviously, this violates the CBSD architecture class semantics.

The final possible implementation in [PREC] maps each precedence to a virtual circuit or virtual channel of an underlying link technology such as ATM or Frame Relay. As the authors of [PREC] also mention, this mechanism is not fundamentally different than the weighted round-robin one.

The Expedited Forwarding PHB, proposed in [DSFIELD], has also close similarities with some of the Class PHBs. Specifically, [DSFIELD] states: "An EF-marked packet is queued for an outgoing interface in a queue that is expected to be relatively small and which is configured to have the highest level of service priority (in some loose sense) within the router's scheduling and buffer management sub-systems". However, this is basically the same description for the highest Class-PHB, namely Class-8. If we allow different levels of EF behavior, it may be possible that other high classes can also have similar forwarding behaviors with the EF PHB.

5. A Possible Implementation for the Class PHBs

A possible implementation for the Class PHBs could be a CBQ scheduler [CBQ], where the weights that are assigned to each queue are dynamically varied depending on the load of that queue. Note that in such a CBQ implementation it may be impossible to guarantee that a Class-I packet will always get a lower delay compared to being serviced by the Class-J queue ($I > J$), or that if a packet has to be dropped in the Class-I queue then it would also had to be dropped in the Class-J queue ($I > J$). However, by applying appropriate class-weight adjustments in the CBQ scheduler over certain timescales, these services can be met in the statistical sense. Other mechanisms, based on admission control, class-based routing, or other types of buffer management and scheduling policies may also be applicable.

Constantinos Dovrolis

Expires: December 1998

[Page 6]

6. Acknowledgements

There were two sources of motivation for writing this document. The first was the interesting discussions at the diffserv mailing list. The second was the presentation of the IPv6 Class field in Christian Huitema's book "IPv6: The New Internet Protocol" (although this does not necessarily mean that Christian agrees with what is written in this document).

7. References

- [DSFIELD] K.Nichols, S.Blake,
"Definition of the Differentiated Services Field (DS Byte)
in the IPv4 and IPv6 Headers",
<[draft-ietf-diffserv-00.txt](#)>, May 1998
- [DSARCH] D.Black, S.Blake, M.Carlson, E.Davies, Z.Wang, W.Weiss,
"An Architecture for Differentiated Services",
<[draft-ietf-diffserv-arch-00.txt](#)>, May 1998
- [DSFRMW] Y.Bernet, J.Binder, S. Blake, M.Carlson, E.Davies, B.Ohlman,
D.Verma, Z.Wang, W.Weiss,
"A Framework for Differentiated Services",
<[draft-ietf-diffserv-framework-00.txt](#)>, May 1998
- [PREC] F.Baker, S.Brim, T.Li, F.Kastenholz, S.Jagannath, J.K.Renwick,
"IP Precedence in Differentiated Services",
<[draft-ietf-diffserv-precedence-00.txt](#)>, April 1998
- [IPv4] J.Postel,
"Internet Protocol",
[RFC 791](#), September 1981
- [IPv6] S.Deering, R.Hinden,
"Internet Protocol, Version 6 (IPv6) Specification",
<[draft-ietf-ipngwg-ipv6-spec-v2-01.txt](#)>, November 1997
- [CBQ] S.Floyd, V.Jacobson,
"Link-Sharing and Resource Management Models for Packet Networks",
IEEE/ACM Transactions on Networking, pp.365-386, August 1995

8. Author's Address

Constantinos Dovrolis

Graduate Student
Dept. of Electrical & Computer Engineering
University of Wisconsin-Madison
[1415 Engineering Drive](#)
Madison, WI 53706
Phone: (608) 262-5130
Email: dovrolis@ece.wisc.edu

Constantinos Dovrolis

Expires: December 1998

[Page 7]