Internet Engineering Task Force (IETF)
Request for Comments: 5865
Updates: 4542, 4594
Category: Standards Track
ISSN: 2070-1721

F. Baker
J. Polk
Cisco Systems
M. Dolly
AT&T Labs

May 2010

A Differentiated Services Code Point (DSCP) for Capacity-Admitted Traffic

Abstract

This document requests one Differentiated Services Code Point (DSCP) from the Internet Assigned Numbers Authority (IANA) for a class of real-time traffic. This traffic class conforms to the Expedited Forwarding Per-Hop Behavior. This traffic is also admitted by the network using a Call Admission Control (CAC) procedure involving authentication, authorization, and capacity admission. This differs from a real-time traffic class that conforms to the Expedited Forwarding Per-Hop Behavior but is not subject to capacity admission or subject to very coarse capacity admission.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc5865.

Table of Contents

1. Introduction

This document requests one Differentiated Services Code Point (DSCP) from the Internet Assigned Numbers Authority (IANA) for a class of real-time traffic. This class conforms to the Expedited Forwarding (EF) [RFC3246] [RFC3247] Per-Hop Behavior. It is also admitted using a CAC procedure involving authentication, authorization, and capacity admission. This differs from a real-time traffic class that conforms to the Expedited Forwarding Per-Hop Behavior but is not subject to capacity admission or subject to very coarse capacity admission.

In addition, this document recommends that certain classes of video described in [RFC4594] be treated as requiring capacity admission.

Real-time traffic flows have one or more potential congestion points between the endpoints. Reserving capacity for these flows is important to application performance. All of these applications have low tolerance to jitter (aka delay variation) and loss, as summarized in Section 2, and most (except for multimedia conferencing) have inelastic flow behavior from Figure 1 of [RFC4594]. Inelastic flow behavior and low jitter/loss tolerance are the service characteristics that define the need for admission control behavior.

One of the reasons behind the requirement for capacity admission is the need for classes of traffic that are handled under special policies. Service providers need to distinguish between special-policy traffic and other classes, particularly the existing Voice over IP (VoIP) services that perform no capacity admission or only very coarse capacity admission and can exceed their allocated resources.

The requested DSCP applies to the Telephony Service Class described in [RFC4594].

Since video classes have not had the history of mixing admitted and non-admitted traffic in the same Per-Hop Behavior (PHB) as has occurred for EF, an additional DSCP code point is not recommended within this document for video. Instead, the recommended "best practice" is to perform admission control for all traffic in three of the video classes from [RFC4594]:

- The Interactive Real-Time Traffic (CS4, used for Video conferencing and Interactive gaming),
- The Broadcast TV (CS3) for use in a video on demand context, and
- The AF4 Multimedia Conferencing (video conferencing).

Other video classes are believed not to have the current problem of confusion with unadmitted traffic and therefore would not benefit from the notion of a separate DSCP for admitted traffic. Within an ISP and on inter-ISP links (i.e., within networks whose internal paths are uniform at hundreds of megabits per second or faster), one would expect all of this traffic to be carried in the Real-Time Traffic (RTP) class described in [RFC5127].

1.1. Definitions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

The following terms and acronyms are used in this document.

PHB: A Per-Hop Behavior (PHB) is the externally observable forwarding behavior applied at a Differentiated Services compliant node to a DS behavior aggregate [RFC2475]. It may be thought of as a program configured on the interface of an Internet host or router, specified in terms of drop probabilities, queuing priorities or rates, and other handling characteristics for the traffic class.

DSCP: The Differentiated Services Code Point (DSCP), as defined in [RFC2474], is a value that is encoded in the DS field, and that each DS Node MUST use to select the PHB that is to be experienced by each packet it forwards [RFC3260]. It is a 6-bit number embedded into the 8-bit TOS (type of service) field of an IPv4 datagram or the Traffic Class field of an IPv6 datagram.

CAC: Call Admission Control includes concepts of authorization and capacity admission. "Authorization" refers to any procedure that identifies a user, verifies the authenticity of the identification, and determines whether the user is authorized to use the service under the relevant policy. "Capacity Admission" refers to any procedure that determines whether capacity exists supporting a session’s requirements under some policy.

In the Internet, these are separate functions; while in the Public Switched Telephone Network (PSTN), they and call routing are carried out together.

UNI: A User/Network Interface (UNI) is the interface (often a physical link or its virtual equivalent) that connects two entities that do not trust each other, and in which one (the user) purchases connectivity services from the other (the network).

Figure 1 shows two user networks connected by what appears to each of them to be a single network ("The Internet", access to which is provided by their service provider) that provides connectivity services to other users.

UNIs tend to be the bottlenecks in the Internet, where users purchase relatively low amounts of bandwidth for cost or service reasons, and as a result are most subject to congestion issues and therefore issues requiring traffic conditioning and service prioritization.

NNI: A Network/Network Interface (NNI) is the interface (often a physical link or its virtual equivalent) that connects two entities that trust each other within limits, and in which the two are seen as trading services for value. Figure 1 shows three service networks that together provide the connectivity services that we call "the Internet". They are different administrations and are very probably in competition, but exchange contracts for connectivity and capacity that enable them to offer specific services to their customers.

NNIs may not be bottlenecks in the Internet if service providers contractually agree to provision excess capacity at them, as they commonly do. However, NNI performance may differ by ISP, and the performance guarantee interval may range from a month to a much shorter period. Furthermore, a peering point NNI may not have contractual performance guarantees or may become overloaded under certain conditions. They are also policy-controlled interfaces, especially in BGP. As a result, they may require a traffic prioritization policy.

Queue: There are multiple ways to build a multi-queue scheduler. Weighted Round Robin (WRR) literally builds multiple lists and visits them in a specified order, while a calendar queue (often used to implement Weighted Fair Queuing, or WFQ) builds a list for each time interval and queues at most a stated amount of data in each such list for transmission during that time interval. While these differ dramatically in implementation, the external difference in behavior is generally negligible when they are properly configured. Consistent with the definitions used in the Differentiated Services Architecture [RFC2475], these are treated as

equivalent in this document, and the lists of WRR and the classes of a calendar queue will be referred to uniformly as "queues".

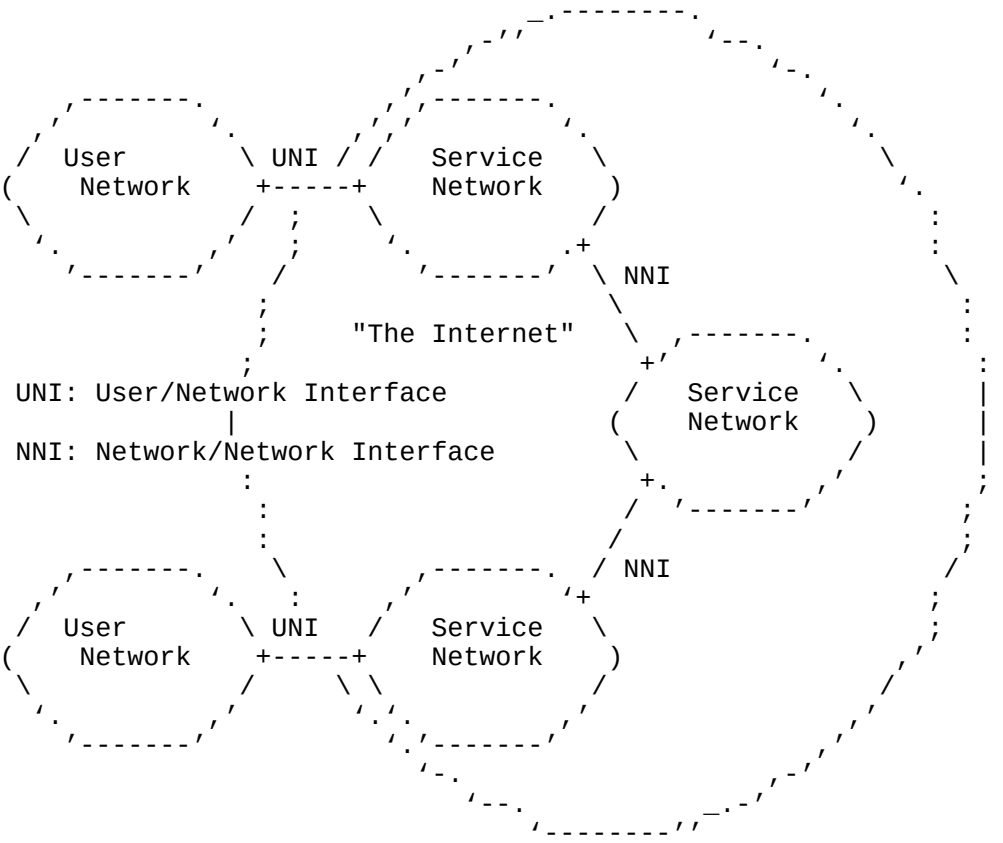

Figure 1: UNI and NNI Interfaces

1.2. Problem

In short, the Telephony Service Class, described in [RFC4594], permits the use of capacity admission in implementing the service, but present implementations either provide no capacity admission services or do so in a manner that depends on specific traffic engineering. In the context of the Internet backbone, the two are essentially equivalent; the edge network depends on specific engineering by the service provider that might not be present, especially in a mobile environment.

However, services are being requested of the network that would specifically make use of capacity admission, and would distinguish among users or the uses of available Voice-over-IP or Video-over-IP capacity in various ways. Various agencies would like to provide services as described in RFC [RFC4190] or in Section 2.6 of [RFC4504].

This requires the use of capacity admission to differentiate among users to provide services to them that are not afforded to non-capacity admitted customer-to-customer IP telephony sessions.

2. Candidate Implementations of the Admitted Telephony Service Class

2.1. Potential Implementations of EF in This Model

There are at least two possible ways to implement isolation between the Capacity Admitted PHB and the Expedited Forwarding PHB in this model. They are to implement separate classes as a set of

- Multiple data plane traffic classes, each consisting of a policer and a queue, with the queues enjoying different priorities, or

- Multiple data plane traffic classes, each consisting of a policer but feeding into a common queue or multiple queues at the same priority.

We will explain the difference and describe in what way they differ in operation. The reason this is necessary is that there is current confusion in the industry.

The multi-priority model is shown in Figure 2. In this model, traffic from each service class is placed into a separate priority queue. If data is present in more than one queue, traffic from one of them will always be selected for transmission. This has the effect of transferring jitter from the higher-priority queue to the lower-priority queues, and reordering traffic in a way that gives the higher-priority traffic a smaller average queuing delay. Each queue must have its own policer, however, to protect the network from errors and attacks; if a traffic class thinks it is carrying a certain data rate but an abuse sends significantly more, the effect of simple prioritization would not preserve the lower priorities of traffic, which could cause routing to fail or otherwise impact a service level agreement (SLA).

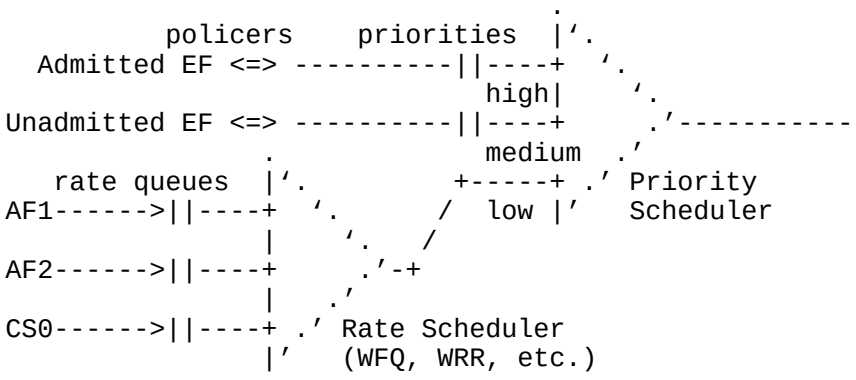

Figure 2: Implementation as a Data Plane Priority

The multi-policer model is shown in Figure 3. In this model, traffic from each service class is policed according to its SLA requirements, and then placed into a common priority queue. Unlike the multi-priority model, the jitter experienced by the traffic classes in this case is the same, as there is only one queue, but the sum of the traffic in this higher-priority queue experiences less average jitter than the elastic traffic in the lower-priority.

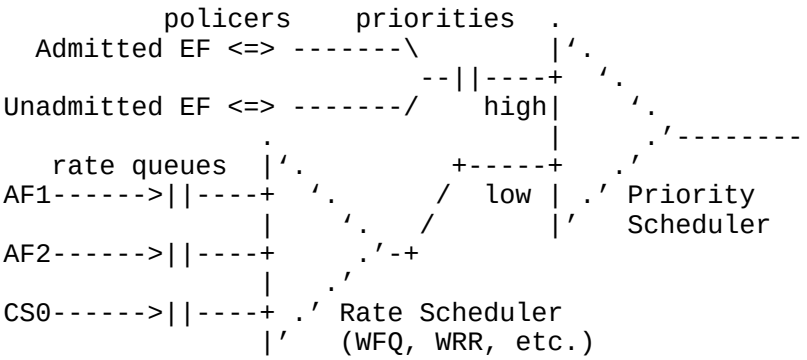

Figure 3: Implementation as a Data Plane Policer

The difference between the two operationally is, as stated, the issues of loss due to policing and distribution of jitter.

If the two traffic classes are, for example, voice and video, datagrams containing video data can be relatively large (often of variable sizes up to the path MTU), while datagrams containing voice are relatively small, on the order of only 40 to 200 bytes, depending on the codec. On lower-speed links (less than 10 MBPS), the jitter introduced by video to voice can be disruptive, while at higher

speeds, the jitter is nominal compared to the jitter requirements of voice. Therefore, at access network speeds, [RFC4594] recommends the separation of video and voice into separate queues, while at optical speeds, [RFC5127] recommends that they use a common queue.

If, on the other hand, the two traffic classes are carrying the same type of application with the same jitter requirements, then giving one preference in this sense does not benefit the higher-priority traffic and may harm the lower-priority traffic. In such a case, using separate policers and a common queue is a superior approach.

2.2. Capacity Admission Control

There are at least six major ways that capacity admission is done or has been proposed to be done for real-time applications. Each will be described below, and Section 3 will judge which ones are likely to meet the requirements of the Admitted Telephony service class. These include:

- Drop Precedence used to force sessions to voluntarily exit,
- Capacity admission control by assumption or engineering,
- Capacity admission control by call counting,
- Endpoint capacity admission performed by probing the network,
- Centralized capacity admission control via bandwidth broker, and
- Distributed capacity admission control using protocols such as Resource Reservation Protocol (RSVP) or Next Steps in Signaling (NSIS).

The problem with dropping traffic to force users to hang up is that it affects a broad class of users -- if there is capacity for N calls and the N+1 calls are active, data is dropped randomly from all sessions to ensure that offered load doesn’t exceed capacity. On very fast links, that is acceptable, but on lower speed links it can seriously affect call quality. There is also a behavioral issue involved here, in which users who experience poor quality calls tend to hang up and call again, making the problem better -- then worse.

The problem with capacity admission by assumption, which is widely deployed in today’s VoIP environment, is that it depends on the assumptions made. One can do careful traffic engineering to ensure needed bandwidth, but this can also be painful, and has to be revisited when the network is changed or network usage changes.

The problem with call-counting-based admission control is that it gets exponentially worse the farther you get from the control point (e.g., it lacks sufficient scalability on the outskirts of the network).

There are two fundamental problems with depending on the endpoint to perform capacity admission: it may not be able to accurately measure the impact of the traffic it generates on the network, and it tends to be greedy (e.g., it doesn't care). If the network operator is providing a service, he must be able to guarantee the service, which means that he cannot trust systems that are not controlled by his network.

The problem with capacity controls via a bandwidth broker is that centralized servers lack far away awareness, and also lack effective real-time reaction to dynamic changes in all parts of the network at all instances of time.

The problem with mechanisms that do not enable the association of a policy with the request is that they do not allow for multi-policy services, which are becoming important.

The operator's choice of admission procedure MUST, for this DSCP, ensure the following:

- The actual links that a session uses have enough bandwidth to support it.
- New sessions are refused admission if there is inadequate bandwidth under the relevant policy.
- A user is identified and the correct policy is applied if multiple policies are in use in a network.
- Under periods of network stress, the process of admission of new sessions does not disrupt existing sessions, unless the service explicitly allows for disruption of calls.

2.3. Recommendations on Implementation of an Admitted Telephony Service Class

When coupled with adequate Authentication, Authorization, and Accounting (AAA) and capacity admission procedures as described in Section 2.2, either of the two PHB implementations described in Section 2.1 is sufficient to provide the services required for an Admitted Telephony service class. If preemption is required, Section 2.3.5.2 of [RFC4542] provides the tools for carrying out the preemption. If preemption is not in view, or if used in addition to

preemptive services, the application of different thresholds depending on call precedence has the effect of improving the probability of call completion by admitting preferred calls at a time when other calls are being refused. Routine and priority traffic can be admitted using the same DSCP value, as the choice of which calls are admitted is handled in the admission procedure executed in the control plane, not the policing of the data plane.

On the point of what protocols and procedures are required for authentication, authorization, and capacity admission, we note that clear standards do not exist at this time for bandwidth brokers, NSIS has not been finalized at this time and in any event is limited to unicast sessions, and that RSVP has been standardized and has the relevant services. We therefore RECOMMEND the use of a protocol, such as RSVP, at the UNI. Procedures at the NNI are business matters to be discussed between the relevant networks, and are RECOMMENDED but NOT REQUIRED.

3. Summary: Changes from RFC 4594

To summarize, there are two changes to [RFC4594] discussed in this document:

Telephony class: The Telephony Service Class in RFC 4594 does not involve capacity admission, but depends on application layer admission that only estimates capacity, and does that through static engineering. In addition to that class, a separate Admitted Telephony Class is added that performs capacity admission dynamically.

Video classes: Capacity admission is added to three video classes. These are the Interactive Real-Time Traffic class, Broadcast TV class when used for video on demand, and the Multimedia Conferencing class.

4. IANA Considerations

IANA assigned a DSCP value to a second EF traffic class consistent with [RFC3246] and [RFC3247] in the "Differentiated Services Field Codepoints" registry. It implements the Telephony Service Class described in [RFC4594] at lower speeds and is included in the Real-Time Treatment Aggregate [RFC5127] at higher speeds. The code point value should be from pool 1 within the dscp-registry. The value is parallel with the existing EF code point (101110), as IANA assigned

the code point 101100 -- keeping the (left-to-right) first 4 binary values the same in both. The code point described in this document is referred to as VOICE-ADMIT and has been registered as follows:

Sub-registry: Pool 1 Codepoints
Reference: [RFC2474]
Registration Procedures: Standards Action

Registry:		
Name	Space	Reference
VOICE-ADMIT	101100	[RFC5865]

This traffic class REQUIRES the use of capacity admission, such as RSVP services together with AAA services, at the User/Network Interface (UNI); the use of such services at the NNI is at the option of the interconnected networks.

5. Security Considerations

A major requirement of this service is effective use of a signaling protocol, such as RSVP, with the capabilities to identify its user as either an individual or a member of some corporate entity, and assert a policy such as "normal", "routine", or some level of "priority".

This capability, one has to believe, will be abused by script kiddies and others if the proof of identity is not adequately strong or if policies are written or implemented improperly by the carriers. This goes without saying, but this section is here for it to be said.

Many of the security considerations from RFC 3246 [RFC3246] apply to this document, as well as the security considerations in RFC 2474 and RFC 4542. RFC 4230 [RFC4230] analyzes RSVP, providing some gap analysis to the NSIS WG as they started their work. Keep in mind that this document is advocating RSVP at the UNI only, while RFC 4230 discusses (mostly) RSVP from a more complete point of view (i.e., e2e and edge2edge). When considering the RSVP aspect of this document, understanding Section 6 of RFC 4230 is a good source of information.

6. Acknowledgements

Kwok Ho Chan, Georgios Karagiannis, Dan Voce, and Bob Briscoe commented and offered text. The impetus for including video in the discussion, which initially only targeted voice, is from Dave McDysan.

7. References

7.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2474] Nichols, K., Blake, S., Baker, F., and D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", RFC 2474, December 1998.

[RFC3246] Davie, B., Charny, A., Bennet, J., Benson, K., Le Boudec, J., Courtney, W., Davari, S., Firoiu, V., and D. Stiliadis, "An Expedited Forwarding PHB (Per-Hop Behavior)", RFC 3246, March 2002.

7.2. Informative References

[RFC2475] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., and W. Weiss, "An Architecture for Differentiated Service", RFC 2475, December 1998.

[RFC3247] Charny, A., Bennet, J., Benson, K., Boudec, J., Chiu, A., Courtney, W., Davari, S., Firoiu, V., Kalmanek, C., and K. Ramakrishnan, "Supplemental Information for the New Definition of the EF PHB (Expedited Forwarding Per-Hop Behavior)", RFC 3247, March 2002.

[RFC3260] Grossman, D., "New Terminology and Clarifications for Diffserv", RFC 3260, April 2002.

[RFC4190] Carlberg, K., Brown, I., and C. Beard, "Framework for Supporting Emergency Telecommunications Service (ETS) in IP Telephony", RFC 4190, November 2005.

[RFC4504] Sinnreich, H., Ed., Lass, S., and C. Stredicke, "SIP Telephony Device Requirements and Configuration", RFC 4504, May 2006.

[RFC4542] Baker, F. and J. Polk, "Implementing an Emergency Telecommunications Service (ETS) for Real-Time Services in the Internet Protocol Suite", RFC 4542, May 2006.

[RFC4594] Babiarz, J., Chan, K., and F. Baker, "Configuration Guidelines for DiffServ Service Classes", RFC 4594, August 2006.

[RFC5127] Chan, K., Babiarz, J., and F. Baker, "Aggregation of DiffServ Service Classes", RFC 5127, February 2008.

[RFC4230] Tschofenig, H. and R. Graveman, "RSVP Security Properties", RFC 4230, December 2005.

Authors' Addresses

Fred Baker
Cisco Systems
Santa Barbara, California 93117
USA

Phone: +1-408-526-4257
EMail: fred@cisco.com

James Polk
Cisco Systems
Richardson, Texas 75082
USA

Phone: +1-817-271-3552
EMail: jmpolk@cisco.com

Martin Dolly
AT&T Labs
Middletown Township, New Jersey 07748
USA

Phone: +1-732-420-4574
EMail: mdolly@att.com