Workgroup: TSVWG
Internet-Draft:
draft-ietf-tsvwg-dscp-considerations-05
Published: 10 August 2022
Intended Status: Informational
Expires: 11 February 2023
Authors: A. Custura, University of Aberdeen
G. Fairhurst, University of Aberdeen
R. Secchi, University of Aberdeen

Considerations for Assigning a new Recommended DiffServ Codepoint (DSCP)

Abstract

This document discusses considerations for assigning a new recommended DiffServ Code Point (DSCP) for a new standard Per Hop Behaviour (PHB). It considers the common observed remarking behaviours that the DiffServ field might be subjected to along an Internet path. It also notes some implications of using a specific DSCP.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

The Differentiated Services (DiffServ) architecture has been deployed in many networks. It provides differentiated traffic forwarding based on the DiffServ Code Point (DSCP) [RFC2474] carried in the DiffServ field [RFC2474] of the IP packet header.

A DiffServ node associates traffic with a service class [RFC4594], and categorises it into Behavior Aggregates [RFC4594]. Configuration guidelines for service classes are provided in RFC4594 [RFC4594]. In IP networks, behaviour aggregates are associated with a DiffServ Code Point (DSCP), which in turn maps to a Per Hop Behaviour (PHB). A Treatment Aggregate (TA) is concerned only with the forwarding treatment of the traffic forming a behaviour aggregate, which could be mapped from a set of DSCP values [RFC5127]. Treatment differentiation can be realised using a variety of schedulers and queues, and also by algorithms that implement access to the physical media.

Within a DiffServ domain, operators can set service level specifications [RFC3086], each of which maps to a particular Per Domain Behaviour (PDB). The PDB defines which DSCP (or set of DSCPs) will be associated with specific TAs as the packets pass through a DiffServ domain, and whether the packets are remarked as they are forwarded.

```
Application -> Service
Traffic        Class
                 |
               Behavior  -> DiffServ -> Per Hop
               Aggregate    Codepoint   Behavior
                                          |
                                        Treatment -> Schedule
                                        Aggregate    Queue, Drop
```

Figure showing the role of DSCPs in classifying IP traffic for differential network treatment by a DiffServ Node.

This document discusses considerations for assigning a new DSCP for a standard PHB. It considers some commonly observed DSCP remarking behaviours that might be experienced along an Internet path. It also describes some packet forwarding treatments that a packet with a specific DSCP can expect to receive when forwarded across a link or subnetwork.

The document is motivated by new opportunities to use DiffServ end-to-end across multiple domains, such as [I-D.ietf-tsvwg-nqb], proposals to build mechanisms using DSCPs in other standards-setting organisations, and the desire to use a common set of DSCPs across a range of infrastructure (e.g., [RFC8622], [I-D.ietf-tsvwg-nqb], [I-D.learmonth-rfc1226-bis]).

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC2119] [RFC8174] [RFC8174] when, and only when, they appear in all capitals, as shown here.

DSCPs are specified in the IANA registry [DSCP-registry] where a variety of different formats are described. A DSCP can sometimes be referred to by name, such as "CS1", and sometimes by a decimal, hex, or binary value. Hex values will be represented in text using prefix 0x. Binary values will use prefix 0b.

3. Current usage of DSCPs

This section describes current usage of DSCPs.

3.1. IP-Layer Semantics

The DiffServ architecture specifies the use of the DiffServ field in the IPv4 and IPv6 packet headers to carry one of 64 distinct DSCP values. Within a given administrative boundary, each DSCP value can be mapped to a distinct PHB [RFC2474]. When a new PHB is standardized, a recommended DSCP value among those 64 values is normally reserved for that PHB, and is assigned by IANA. An operator is not formally required to use the recommended value; indeed [RFC2474] states that "the mapping of codepoints to PHBs MUST be configurable." However, use of the recommended value is usually convenient and avoids confusion.

The DSCP space is divided into three pools for the purpose of assignment and management [DSCP-registry]. The pools can be summarised in a table (where 'x' refers to either a bit position with value '0' or '1').

DSCP Pool 1: A pool of 32 codepoints with a format 0bxxxxx0, to be assigned by IANA Standards Action [RFC8126].

DSCP Pool 2: A pool of 16 codepoints with a format of 0bxxxx11, reserved for experimental or local (private) use by network operators (see sections 4.1 and 4.2 of [RFC8126].

DSCP Pool 3: A pool of 16 codepoints with a format of 0bxxxx01. This was initially available for experimental or local use, but was originally specified to be preferentially utilised for standardized assignments if Pool 1 is ever exhausted. [RFC4594] had recommended a local use of DSCP values 0x01, 0x03, 0x05 and 0x07 (codepoints with the format of 0b000xx1). Pool 3 codepoints are now utilised for standardized assignments and are no longer available for assignment to experimental or local use [RFC8436]. [RFC8622] assigned 0x01 from this pool and consequentially updated [RFC4594]. Any future request to assign 0x05 would be expected to similarly update [RFC4594].

The DSCP space is shown in the following Figure.

56/CS7	57	58	59	60	61	62	63
48/CS6	49	50	51	52	53	54	55
40/CS5	41	42	43	44/VA	45	46/EF	47
32/CS4	33	34/AF41	35	36/AF42	37	38/AF43	39
24/CS3	25	26/AF31	27	28/AF32	29	30/AF33	31
16/CS2	17	18/AF21	19	20/AF22	21	22/AF23	23
8/CS1	9	10/AF11	11	12/AF12	13	14/AF13	15
0/CS0	1/LE	2	3	4	5	6	7

Figure showing the current list of assigned DSCPs and their assigned PHBs.

CS	Class Selector	RFC 2474
BE	Best Effort (CS0)	RFC 2474
AF	Assured Forwarding	RFC 2597
EF	Expedited Forwarding	RFC 3246
VA	Voice Admit	RFC 5865
LE	Lower Effort	RFC 8622

Figure showing the summary of the DSCP abbreviations used in previous RFCs [RFC2474] [RFC2597] [RFC3246] [RFC5865] [RFC8622], as described in the IANA registry [DSCP-registry]. BE, also known as CS0, describes the default forwarding treatment.

The DiffServ architecture allows forwarding treatments to be associated with each DSCP, and the RFC series describes some of these as PHBs. Although DSCPs are intended to identify specific treatment requirements, multiple DSCPs might also be mapped (aggregated) to the same forwarding treatment. DSCPs can be mapped to treatment aggregates that might result in remarking (e.g., RFC5160 [RFC5160] suggests Meta-QoS-Classes to help enable deployment of standardized end-to-end QoS classes)

3.2. Network Control Traffic

Network Control Traffic is defined as packet flows that are essential for stable operation of the administered network (see [RFC4594], Section 3). This traffic is marked with a value from a set of Class Selector (CS) DSCPs. This traffic is often a special case within a provider network, and ingress traffic with these DSCP markings can be remarked.

DSCP CS2 is recommended for the OAM (Operations, Administration, and Maintenance) service class (see [RFC4594], Section 3.3).

DSCP CS6 is recommended for local network control traffic. This includes routing protocols and OAM traffic that are essential to network operation administration, control and management. Section 3.2 of RFC4594 [RFC4594] recommends that "CS6 marked packet flows from untrusted sources (for example, end-user devices) SHOULD be dropped or remarked at ingress to the DiffServ network".

DSCP CS7 is reserved for future use by network control traffic. "CS7 marked packets SHOULD NOT be sent across peering points" [RFC4594].

RFC2474 [RFC2474] recommends PHBs selected by CS6 and CS7 "MUST give packets preferential forwarding treatment by comparison to the PHB selected by codepoint '000000'".

At the time of writing, there is evidence to suggest CS6 is actively used by network operators for control traffic. A study of traffic at a large Internet Exchange showed around 40% of ICMP traffic carried this mark [IETF113DSCP]. Similarly, another study found many routers remark all traffic except those packets with a DSCP that sets the higher order bits to 0b11 (see Section 4 of this document).

4. Remarking the DSCP

It is a feature of the DiffServ architecture that the DiffServ field of packets can be remarked at domain boundaries (see section 2.3.4.2 of [RFC2475]). A DSCP can be remarked at the ingress of a DiffServ domain. This remarking can change the DSCP value used on the remainder of an IP path, or the network can restore the initial DSCP marking at the egress of the domain. The DiffServ field can also be remarked based on common semantics and agreements between providers at an exchange point. Furthermore, [RFC2474] states that remarking must occur when there is a possibility of theft/denial-of-service attack.

If packets are received that are marked with an unknown or an unexpected DSCP, [RFC2474] recommends forwarding the packet using a default (best effort) treatment, but without changing the DSCP. This seeks to support incremental DiffServ deployment in existing networks as well as preserve DSCP markings by routers that have not been configured to support DiffServ. (See also Section 4.3). [RFC3260] clarifies that this remarking specified by RFC2474 is intended for interior nodes within a DiffServ domain. For DiffServ ingress nodes the traffic conditioning required by RFC 2475 applies first.

Reports measuring existing deployments have categorised DSCP remarking [Custura] [Barik] into the following seven observed remarking behaviours:

Bleach: bleaches all traffic (sets the DSCP to zero);

Bleach-ToS-Precedence: set the first three bits of the DSCP field to 0b000 (reset the 3 bits of the former ToS Precedence field, defined in [RFC0791], and clarified in [RFC1122]);

Bleach-some-ToS: set the first three bits of the DSCP field to 0b000 (reset the 3 bits of the former ToS Precedence field), unless the first two bits of the DSCP field are 0b11;

Remark-ToS: set the first three bits of the DSCP field to any value different than 0b000 (replace the 3 bits of the former ToS Precedence field);

Bleach-low: set the last three bits of the DSCP field to 0b000;

Bleach-some-low: set the last three bits of the DSCP field to 0b000, unless the first two bits of the DSCP field are 0b11;

Remark: remarks all traffic to one or more particular (non-zero) DSCP values.

NOTE: More than one mechanism could result in the same DSCP remarking (see below). It is not generally possible for an external observer to determine which mechanism results in a specific remarking solely from the change in an observed DSCP value.

4.1. Bleaching the DSCP Field

A specific form of remarking occurs when the DiffServ field is re-assigned to the default treatment, CS0 (0x00). This results in traffic being forwarded using the BE PHB. For example, AF31 (0x1a) would be bleached to CS0.

A survey reported that resetting all the bits of the DiffServ field to 0 was seen to be more prevalent at the edge of the network, and rather less common in core networks [Custura].

4.2. IP Type of Service manipulations

The IETF first defined ToS precedence for IP packets in [RFC0791], and updated it to be part of the ToS Field in [RFC1349]. Since 1998, this practice has been deprecated by [RFC2474]. RFC 2474 defines DSCPs 0bxxx000 as the Class Selector codepoints, where PHBs selected by these codepoints MUST meet the Class Selector PHB Requirements" described in Sec. 4.2.2.2 of that RFC.

However, a recent survey reports practices based on ToS semantics have not yet been eliminated from the Internet, and need to still be considered when making new DSCP assignments [Custura].

4.2.1. Impact of ToS Precedence Bleaching

ToS Precedence Bleaching (/Bleach-ToS-Precedence/) is a practice that resets the first three bits of the DSCP field to zero (the former ToS Precedence field), leaving the last three bits unchanged (see section 4.2.1 of [RFC2474]). A DiffServ node can be configured in a way that results in this remarking. This remarking can also occur when packets are processed by a router that is not configured with DiffServ (e.g., configured to operate on the former ToS precedence field [RFC0791]). At the time of writing, this is a common manipulation of the DiffServ field. The following Figure depicts this remarking.

```
+-+-+-+-+-+-+
|0 0 0|x x x|
+-+-+-+-+-+-+
```

Figure showing the ToS Precedence Bleaching (/Bleach-ToS-Precedence/) observed remarking behaviour, based on Section 3 of [RFC1349]. The bit positions marked "x" are not changed.

56/CS7	57	58	59	60	61	62	63
48/CS6	49	50	51	52	53	54	55
40/CS5	41	42	43	44/VA	45	46/EF	47
32/CS4	33	34/AF41	35	36/AF42	37	38/AF43	39
24/CS3	25	26/AF31	27	28/AF32	29	30/AF33	31
16/CS2	17	18/AF21	19	20/AF22	21	22/AF23	23
8/CS1	9	10/AF11	11	12/AF12	13	14/AF13	15
0/CS0	1/LE	2	3	4	5	6	7

As a result of ToS Precedence Bleaching, all the DSCPs in each column are remarked to the smallest DSCP in that column. The DSCPs in the bottom row therefore have higher survivability across an end-to-end Internet path.

0/CS0	1/LE	2	3	4	5	6	7
Assigned		ToS Prec Bl. of AF11..41	EXP/LU	Used by SSH	Future NQB	ToS Prec Bl. of AF13..EF	Exp/LU

Figure showing 0b000xxx DSCPs, highlighting any current assignments and whether they are affected by any known remarking behaviours. For example, ToS Precedence Bleaching of popular DSCPs AF11,21,31,41 would result in traffic being remarked with DSCP 2 in the Internet core. DSCP 4 has been historically used by the SSH application, following semantics which precede DiffServ[DSCP4].

If ToS Precedence Bleaching occurs, packets with a DSCP 'x' would be remarked to 'x' & 0x07and then would be treated according to the corresponding PHB.

A variation of this observed remarking behaviour clears the top three bits of a DSCP, unless these have values 0b110 or 0b111 (corresponding to the CS6 and CS7 DSCPs). As a result, a DSCP value greater than 48 decimal (0x30) is less likely to be impacted by ToS Precedence Bleaching.

4.2.2. Impact of ToS Precedence Remarking

Practices based on ToS Precedence Remarking (/Remark-ToS/) [RFC1349] were deprecated by [RFC2474]. These practices based on ToS semantics have not yet been eliminated from deployed networks.

```
+-+-+-+-+-+-+
|0 0 1|x x x|
+-+-+-+-+-+-+
```

Figure showing ToS Precedence Remarking (/Remark-ToS/) observed behaviour, based on Section 3 of [RFC1349]. The bit positions marked "x" remain unchanged.

A less common remarking, ToS Precedence Remarking sets the first three bits of the DSCP to a non-zero value corresponding to a CS PHB. This remarking occurs when routers are not configured to perform DiffServ remarking.

If remarking occurs, packets are forwarded using the PHB specified for the resulting DSCP. For example, the AF31 DSCP (0x1a) could be remarked to either AF11 or AF21.

4.3. Remarking to a Particular DSCP

A network device might remark the DiffServ field of an IP packet based on a local policy with a specific (set of) DSCPs, /Remark/.

Both [RFC2474] and [RFC8100] recommend that DiffServ boundary nodes use remarking, if necessary, to avoid theft/denial of service or ensure that appropriate DSCPs are used within a DiffServ domain. Some networks therefore may not follow the earlier recommendation in [RFC2474] to carry unknown or unexpected DSCPs without modification, and instead remark packets with these codepoints to the default class, CS0 (0x00).

Remarking is sometimes performed using a Multi-Field (MF) classifier [RFC2475] [RFC3290] [RFC4594]. For example, a common remarking is to remark all traffic to a single DSCP, thus removing any traffic differentiation (see Section 4.1). Bleaching (/Bleach/) is a specific example of this observed remarking behaviour that remarks to CS0 (0x00).

5. Interpretation of the IP DSCP at Lower Layers

Transmission systems and subnetworks can, and do, utilise the DiffServ field in an IP packet to set a QoS-related field or function at the lower layer. A lower layer could also implement a traffic conditioning function that could remark the DSCP used at the

IP layer. In many cases, this use is constrained by designs that utilise fewer than 6 bits to express the service class, and therefore infer mapping to a smaller L2 QoS field, for example, WiFi or Multi-Protocol Label Switching (MPLS).

5.1. Mapping Specified for IEEE 802

The IEEE specifies standards that include mappings for DSCPs to lower layer elements.

5.1.1. Mapping Specified for IEEE 802.1

A 3-bit Priority Code Point (PCP) is specified in the IEEE 802.1Q tag to mark Ethernet frames to one of eight priority values [IEEE-802-1Q]. The value zero indicates the default best effort treatment, and the value one indicates a background traffic class. The remaining values indicate increasing priority. Internet control traffic can be marked as CS6, and network control is marked as CS7.

The mapping specified in [IEEE-802-1Q] revises a previous standard [IEEE-802-1D], in an effort to align with DiffServ practice: the traffic types are specified to match the first three bits of a suitable DSCP (e.g., the first three bits of the EF DSCP are mapped to a PCP of 5). However, [IEEE-802-1D] maps both PCP1 (Background) and PCP2 (Spare) to indicate lower priority than PCP0, RFC8622. Therefore, different remarking behaviours are expected depending on the age of deployed devices.

5.1.2. Mapping Specified for IEEE 802.11

Section 6 of [RFC8325] provides a brief overview of IEEE 802.11 QoS. The IEEE 802.11 standards [IEEE-802-11] provide MAC functions to support QoS in WLANs using Access Classes (AC). The upstream User Priority (UP) in the 802.11 header has a 3-bit QoS value. A DSCP can be mapped to the UP.

Most current WiFi implementations [RFC8325] use a default mapping that maps the first three bits of the DSCP to the 802.11 UP value. This is then in turn mapped to the four WiFi Multimedia (WMM) Access Categories. The Wi-Fi Alliance has also specified a more flexible mapping that follows RFC8325 and provides functions at an AP to remark packets as well as a QoS Map that maps each DSCP to an AC [WIFI-ALLIANCE].

```
+-+-+-+-+-+-+
|x x x|. . .|
+-+-+-+-+-+-+
```

Figure showing the DSCP bits commonly mapped to the UP in 802.11. The bit positions marked "x" are mapped to the 3-bit UP value, while the ones marked "." are ignored.

RFC8325 [RFC8325] notes inconsistencies that can result from such remarking, and recommends how to perform this remarking. It proposes several recommendations for mapping a DSCP to an IEEE 802.11 UP for wired-to-wireless interconnection. The recommendation includes mapping network control traffic CS6 and CS7, as well unassigned DSCPs, to UP 0.

In the upstream direction (wireless-to-wired interconnections, this mapping can result in a specific DSCP remarking behaviour. Some Access Points (APs) currently use a default UP-to-DSCP mapping [RFC8325], wherein "DSCP values are derived from the layer 2 UP values by multiplying the UP values by eight (i.e., shifting the three UP bits to the left and adding three additional zeros to generate a 6-bit DSCP value). This derived DSCP value is used for QoS treatment between the wireless AP and the nearest classification and marking policy enforcement point (which may be the centralized wireless LAN controller, relatively deep within the network). Alternatively, in the case where there is no other classification and marking policy enforcement point, then this derived DSCP value will be used on the remainder of the Internet path." This can result in remarking /Bleach-low/.

```
+-+-+-+-+-+-+
|x x x|0 0 0|
+-+-+-+-+-+-+
```

Figure showing the observed remarking behaviour resulting from deriving from UP-to-DSCP mapping in some 802.11 networks.

An alternative to UP-to-DSCP remapping uses the DSCP value of a downstream IP packet (e.g., the Control And Provisioning of Wireless Access Points (CAPWAP) protocol, RFC5415, maps an IP packet DiffServ field to the DiffServ field of the outer IP header in a CAPWAP tunnel).

Some current constraints of WiFi mapping are discussed in section 2 of [RFC8325]. A QoS profile can be used to limit the maximum DSCP value used for the upstream and downstream traffic.

5.2. DiffServ and MPLS

Multi-Protocol Label Switching (MPLS) specified eight MPLS Traffic Classes (TCs), which restrict the number of different treatments [RFC5129]. RFC 5127 describes aggregation of DiffServ TCs [RFC5127], and introduces four DiffServ Treatment Aggregates. Traffic marked with multiple DSCPs can be forwarded in a single MPLS TC.

There are three Label-Switched Router (LSR) behaviours: the Pipe, the Short Pipe and the Uniform Model [RFC3270]. These only differ when a LSP performs a push or a pop.

5.2.1. Mapping Specified for MPLS

RFC3270 [RFC3270] defines a flexible solution for support of DiffServ over MPLS networks. This allows an MPLS network administrator to select how BAs (marked by DSCPs) are mapped onto Label Switched Paths (LSPs) to best match the DiffServ, Traffic Engineering and protection objectives within their particular network.

Mappings from the IP DSCP to the MPLS header are defined in Section 4.2 of [RFC3270].

The Pipe Model conveys the "LSP Diff-Serv Information" to the LSP Egress so that its forwarding treatment can be based on the IP DSCP.

When Penultimate Hop Popping (PHP) is used, the Penultimate LSR needs to be aware of the encapsulation mapping for a PHB to the label corresponding to the exposed header to perform DiffServ Information Encoding (Section 2.5.2 of [RFC3270]).

5.2.2. Mapping Specified for MPLS Short Pipe

The Short Pipe Model is an optional variation of the Pipe Model [RFC3270].

ITU-T Y.1566 [ITU-T-Y-1566] further defined a set of four common QoS classes and four auxiliary classes to which a DSCP can be mapped when interconnecting Ethernet, IP and MPLS networks. [RFC8100] proposes four treatment aggregates for interconnection with four defined DSCPs. This was motivated by the requirements of MPLS network operators that use Short-Pipe tunnels, but can be applicable to other networks, both MPLS and non-MPLS.

RFC8100 recommends preserving the notion of end-to-end service classes, and recommends a set of standard DSCPs mapped to a small set of standard PHBs at interconnection. The key requirement is that the DSCP at the network ingress is restored at the network egress. The current version of RFC8100 limits the number of DSCPs to 6 and 3 more are suggested for extension. RFC8100 respects the deployment of PHB groups for DS domain internal use, which limits the number of acceptable external DSCPs (and possibilities for their transparent transport or restoration at network boundaries). In this design, packets marked with DSCPs not part of the RFC8100 codepoint scheme are treated as unexpected and will possibly be remarked (a /Remark/ behaviour) or dealt with via an additional agreement(s) among the operators of the interconnected networks. RFC8100 can be extended to support up to 32 DSCPs by future standards. RFC8100 is operated by at least one Tier 1 backbone provider. Use of the MPLS Short Pipe Model favours remarking unexpected DSCP values to zero in the absence of an additional agreement(s), as explained in [RFC8100]. This can result in bleaching (/Bleach/).

RFC8100 Agg. Class	DSCP
Telephony Service Treatment Aggregate	EF
	VA
Bulk Real-Time Treatment Aggregate	AF41
May be added	(AF42)
May be added	(AF43)
Assured Elastic Treatment Aggregate	AF31
	AF32
Reserved for the extension of PHBs	(AF33)
Default / Elastic Treatment Aggregate	BE/CS0
Network Control: Local Use	CS6

The short-pipe MPLS mapping from RFC 8100.

5.3. Mapping Specified for Mobile Networks

Mobile LTE and 5G standards have evolved from older UMTS standards, and support DiffServ. LTE (4G) and 5G standards [SA-5G] identify traffic classes at the interface between User Equipment (UE) and the mobile Packet Core network by QCI (QoS Class Identifiers) and 5QI (5G QoS Identifier). The 3GPP standards do not define or recommend any specific mapping between each QCI or 5QI and DiffServ (and mobile QCIs are based on several criteria service class definitions). The way packets are mapped at the Packet Gateway (P-GW) boundary is determined by operators. However, TS 23.107 (version 16.0.0, applies to LTE and below) mandates that Differentiated Services, defined by IETF, shall be used to interoperate with IP backbone networks.

The GSM Association (GSMA) has defined four aggregated classes and seven associated PHBs in their guidelines for IPX Provider networks

GSMA IR.34 [GSMA-IR-34]. This was previously specified as the Inter-Service Provider IP Backbone Guidelines, and provides a mobile ISP to ISP QoS mapping mechanism, and interconnection with other IP networks in the general Internet. If realised by an IP VPN, the operator is free to apply its DS Domain internal codepoint scheme at outer headers and inner IPX DSCPs may be transported transparently. The guidelines also describe a case where the DSCP marking from a Service Provider cannot be trusted (depending on the agreement between the Service Provider and its IPX Provider), in which situation the IPX Provider can remark the DSCP value to a static default value.

```
+---------------+-------+
|  GSMA IR.34   |  PHB  |
|  Agg. Class   |       |
+---------------+-------+
|Conversational |  EF   |
+---------------+-------+
| Streaming     | AF41  |
+---------------+-------+
| Interactive   | AF31  |
+               +-------+
| (ordered by   | AF32  |
+   priority,   +-------+
| AF3 highest)  | AF21  |
+               +-------+
|               | AF11  |
+---------------+-------+
| Background    | CS0   |
+---------------+-------+
```

Figure showing the PHB mapping recommended in the guidelines proposed in GSMA IR.34 [GSMA-IR-34].

5.4. Mapping Specified for Carrier Ethernet

Metro Ethernet Forum (MEF) provides a mapping of DSCPs at the IP layer to quality of service markings in the Ethernet frame headers MEF 23.1 [MEF23.1].

5.5. Remarking as a Side-effect of Another Policy

This includes any other remarking that does not happen as a result of traffic conditioning, such as policies and L2 procedures that result in remarking traffic as a side-effect of other functions (e.g., in response to Distributed Denial of Service, DDoS).

5.6. Summary

This section has discussed the various ways in which DSCP remarking behaviours can arise from interactions with lower layers.

6. Considerations for DSCP Selection

This section provides advice for the assignment of a new DSCP value. It is intended to aid the IETF and IESG in considering a request for a new DSCP. The section identifies known issues that might influence the finally assigned DSCP, and provides a summary of considerations for assignment of a new DSCP.

6.1. Effect of Bleaching

New DSCP assignments should consider the impact of bleaching, which can limit the ability to provide the expected treatment end-to-end. This is particularly important for cases where the codepoint is intended to result in lower than best effort treatment, as was the case when defining the LE PHB [RFC8622]. In this case, bleaching, or remarking to "CS0" would result in elevating the lower effort traffic (LE) to the default class (BE/CS0). This is an example of priority inversion.

6.2. Where the proposed DSCP > 0x07 (7)

Although the IETF specifications require systems to use DSCP marking semantics in place of methods based on the former ToS field, the current recommendation is that any new assignment where the DSCP is greater than 0x07 should consider the semantics associated with the resulting DSCP when ToS Precedence Bleaching is experienced. For example, it can be desirable to avoid choosing a DSCP that could be remarked to LE, Lower Effort [RFC8622], which could otherwise potentially result in a priority inversion in the treatment.

6.3. Where the proposed DSCP < 0x07 (7)

ToS Precedence Bleaching can unintentionally result in extra traffic aggregated to the same DSCP. For example, after experiencing ToS Precedence Bleaching, all traffic marked AF11, AF21, AF31 and AF41 would be aggregated with traffic marked with DSCP 2 (0x02), increasing the volume of traffic which receives the treatment associated with DSCP 2. New DSCP assignments should consider unexpected consequences arising from this observed remarking behaviour.

6.3.1. Where the proposed DSCP&0x07=0x01

As a consequence of assigning the LE PHB [RFC8622], the IETF allocated the DSCP 0b000001 from Pool 3.

When making assignments where the DSCP has a format: 0bxxx001, the case of ToS Precedence Bleaching (/Bleach-ToS-Precedence/) of a DSCP to a value of 0x01 needs to be considered. ToS Precedence Bleaching will result in demoting the traffic to the lower effort traffic class. Care should be taken to consider the implications of remarking when shooing to assign a DSCP with this format.

6.4. Impact on deployed infrastructure

Behaviour of deployed PHBs and conditioning treatments also needs to be considered when assigning a new DSCP. Network operators have choices when it comes to configuring DiffServ support within their domains, and the observed remarking behaviours described in the previous section can result from different configurations and approaches:

Networks not remarking DiffServ: A network that either does not implement PHBs, or implements one or more PHBs whilst restoring the DSCP field at network egress with the value at network ingress. Operators in this category pass all DSCPs transparently.

Networks that condition the DSCP: A network that implements more than one PHB and enforces SLAs with its peers. Operators in this category use conditioning to ensure that only traffic that matches a policy is permitted to use a specific DSCP (see [RFC8100]). This requires operators to choose to support or remark a new DSCP assignment.

Networks that remark in some other way , which includes:

1. Networks containing misconfigured devices that do not comply with the relevant RFCs.

2. Networks containing devices that implement an obsolete specification or contain software bugs.

3. Networks containing devices that remark the DSCP as a result of lower layer interactions.

For example, the ToS Precedence Bleaching (/Bleach-ToS-Precedence/) remarking behaviour cannot be observed in the case of networks not remarking DiffServ, but can arise as a result of traffic conditioning at operator boundaries. It can also arise in the case of misconfiguration or in networks using old equipment which precedes DiffServ.

6.5. Questions to guide discussion of a proposed new DSCP

A series of questions emerge that need to be answered when considering a proposal to the IETF that requests a new assignment. These questions include:

* Is the request for local use within a DiffServ domain that does not require interconnection with other DiffServ domains? This request can use DSCPs in Pool 2 for local or experimental use, without any IETF specification for the DSCP or associated PHB.

* How is the proposed service class characterised: What are the characteristics of the traffic to be carried? What are the expectations for treatment?

* Service classes [RFC4594] that can utilise existing PHBs should use assigned DSCPs to mark their traffic: Could the request be met by using an existing IETF DSCP?

* Specification of a new recommended DSCP requires Standards Action. RFC2474 states: "Each standardized PHB MUST have an associated RECOMMENDED codepoint". If approved, new IETF assignments are normally made by IANA in Pool 1, but the IETF can request assignments to be made from Pool 3 [RFC8436]. Does the ID contain an appropriate request to IANA?

* Section 5.2 describes examples of treatment aggregation. What are the effects of treatment aggregation on the proposed DSCP?

* Section 5 describes some observed treatments by layers below IP. What are the implications of the treatments and mapping described in Section 5 on the proposed DSCP?

* DSCPs are assigned to PHBs and can be used to enable nodes along an end-to-end path to classify the packet for a suitable PHB. Section 4 describes some observed remarking behaviour, and Section 6.4 identifies root causes for why this remarking is observed. What is the expected effect of currently-deployed remarking on the end-to-end service?

7. Acknowledgments

The authors acknowledge the helpful discussions and analysis by Greg White and Thomas Fossati in a draft concerning NQB. Ruediger Geib and Brian Carpenter contributed comments, along with other members of the TSVWG.

8. IANA Considerations

This memo provides information to assist in considering new assignments to the IANA DSCP registry (https://www.iana.org/assignments/dscp-registry/dscp-registry.xhtml).

This memo includes no request to IANA, or update to the IANA procedures.

9. Security Considerations

The security considerations are discussed in the security considerations of each cited RFC.

10. References

10.1. Normative References

[DSCP-registry] IANA, "Differentiated Services Field Codepoints (DSCP) Registry", https://www.iana.org/assignments/dscp-registry/dscp-registry.xhtml, 2019.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[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, DOI 10.17487/RFC2474, December 1998, <https://www.rfc-editor.org/info/rfc2474>.

[RFC2475] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., and W. Weiss, "An Architecture for Differentiated Services", RFC 2475, DOI 10.17487/RFC2475, December 1998, <https://www.rfc-editor.org/info/rfc2475>.

[RFC3260] Grossman, D., "New Terminology and Clarifications for Diffserv", RFC 3260, DOI 10.17487/RFC3260, April 2002, <https://www.rfc-editor.org/info/rfc3260>.

[RFC3290] Bernet, Y., Blake, S., Grossman, D., and A. Smith, "An Informal Management Model for Diffserv Routers", RFC 3290, DOI 10.17487/RFC3290, May 2002, <https://www.rfc-editor.org/info/rfc3290>.

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

10.17487/RFC4594, August 2006, <https://www.rfc-editor.org/info/rfc4594>.

[RFC5129] Davie, B., Briscoe, B., and J. Tay, "Explicit Congestion Marking in MPLS", RFC 5129, DOI 10.17487/RFC5129, January 2008, <https://www.rfc-editor.org/info/rfc5129>.

[RFC8100] Geib, R., Ed. and D. Black, "Diffserv-Interconnection Classes and Practice", RFC 8100, DOI 10.17487/RFC8100, March 2017, <https://www.rfc-editor.org/info/rfc8100>.

[RFC8436] Fairhurst, G., "Update to IANA Registration Procedures for Pool 3 Values in the Differentiated Services Field Codepoints (DSCP) Registry", RFC 8436, DOI 10.17487/RFC8436, August 2018, <https://www.rfc-editor.org/info/rfc8436>.

10.2. Informative References

[Barik] Barik, R., Welzl, M., Elmokashfi, A., Dreibholz, T., and S. Gjessing, "Can WebRTC QoS Work? A DSCP Measurement Study", ITC 30, September 2018.

[Custura] Custura, A., Venne, A., and G. Fairhurst, "Exploring DSCP modification pathologies in mobile edge networks", TMA , 2017.

[DSCP4] Kolu, A., "Bogus DSCP value for SSH", online https://lists.freebsd.org/pipermail/freebsd-stable/2010-July/057710.html, 2010.

[GSMA-IR-34] GSM Association, "IR.34 Guidelines for IPX Provider networks (Previously Inter-Service Provider IP Backbone Guidelines)", IR 34, 2017.

[I-D.ietf-tsvwg-nqb] White, G. and T. Fossati, "A Non-Queue-Building Per-Hop Behavior (NQB PHB) for Differentiated Services", Work in Progress, Internet-Draft, draft-ietf-tsvwg-nqb-10, March 2022, <https://www.ietf.org/archive/id/draft-ietf-tsvwg-nqb-10.txt>.

[I-D.learmonth-rfc1226-bis] Learmonth, I. R., "Internet Protocol Encapsulation of AX.25 Frames", Work in Progress, Internet-Draft, draft-learmonth-rfc1226-bis-03, 19 May 2020, <https://www.ietf.org/archive/id/draft-learmonth-rfc1226-bis-03.txt>.

[IEEE-802-11] IEEE, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications", IEEE 802.11, 2007.

[IEEE-802-1D]
IEEE, "IEEE Standard for Local and Metropolitan Area Network-- Media Access Control (MAC) Bridges", IEEE 802.1D, 2004.

[IEEE-802-1Q] IEEE, "IEEE Standard for Local and Metropolitan Area Network-- Bridges and Bridged Networks", IEEE 802.1Q, 2018.

[IETF113DSCP] Custura, A. and G. Fairhurst, "Considerations for assigning DSCPs", IETF 113 Proceedings https://datatracker.ietf.org/meeting/113/materials/slides-113-tsvwg-61-dscp-considerations-07, 2022.

[ITU-T-Y-1566] ITU-T, "Quality of Service Mapping and Interconnection Between Ethernet, Internet Protocol and Multiprotocol Label Switching Networks", ITU-T Y.1566, 2012.

[MEF23.1] MEF, "MEF Technical Specification MEF 23.1-- Carrier Ethernet Class of Service ? Phase 2", MEF 23.1, 2012.

[RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, DOI 10.17487/RFC0791, September 1981, <https://www.rfc-editor.org/info/rfc791>.

[RFC1122] Braden, R., Ed., "Requirements for Internet Hosts - Communication Layers", STD 3, RFC 1122, DOI 10.17487/RFC1122, October 1989, <https://www.rfc-editor.org/info/rfc1122>.

[RFC1349] Almquist, P., "Type of Service in the Internet Protocol Suite", RFC 1349, DOI 10.17487/RFC1349, July 1992, <https://www.rfc-editor.org/info/rfc1349>.

[RFC2597] Heinanen, J., Baker, F., Weiss, W., and J. Wroclawski, "Assured Forwarding PHB Group", RFC 2597, DOI 10.17487/RFC2597, June 1999, <https://www.rfc-editor.org/info/rfc2597>.

[RFC3086] Nichols, K. and B. Carpenter, "Definition of Differentiated Services Per Domain Behaviors and Rules for their Specification", RFC 3086, DOI 10.17487/RFC3086, April 2001, <https://www.rfc-editor.org/info/rfc3086>.

[RFC3246]
Davie, B., Charny, A., Bennet, J C R., Benson, K., Le Boudec, J Y., Courtney, W., Davari, S., Firoiu, V., and D. Stiliadis, "An Expedited Forwarding PHB (Per-Hop

Behavior)", RFC 3246, DOI 10.17487/RFC3246, March 2002, <https://www.rfc-editor.org/info/rfc3246>.

[RFC3270] Le Faucheur, F., Wu, L., Davie, B., Davari, S., Vaananen, P., Krishnan, R., Cheval, P., and J. Heinanen, "Multi-Protocol Label Switching (MPLS) Support of Differentiated Services", RFC 3270, DOI 10.17487/RFC3270, May 2002, <https://www.rfc-editor.org/info/rfc3270>.

[RFC5127] Chan, K., Babiarz, J., and F. Baker, "Aggregation of Diffserv Service Classes", RFC 5127, DOI 10.17487/RFC5127, February 2008, <https://www.rfc-editor.org/info/rfc5127>.

[RFC5160] Levis, P. and M. Boucadair, "Considerations of Provider-to-Provider Agreements for Internet-Scale Quality of Service (QoS)", RFC 5160, DOI 10.17487/RFC5160, March 2008, <https://www.rfc-editor.org/info/rfc5160>.

[RFC5865] Baker, F., Polk, J., and M. Dolly, "A Differentiated Services Code Point (DSCP) for Capacity-Admitted Traffic", RFC 5865, DOI 10.17487/RFC5865, May 2010, <https://www.rfc-editor.org/info/rfc5865>.

[RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <https://www.rfc-editor.org/info/rfc8126>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC8325] Szigeti, T., Henry, J., and F. Baker, "Mapping Diffserv to IEEE 802.11", RFC 8325, DOI 10.17487/RFC8325, February 2018, <https://www.rfc-editor.org/info/rfc8325>.

[RFC8622] Bless, R., "A Lower-Effort Per-Hop Behavior (LE PHB) for Differentiated Services", RFC 8622, DOI 10.17487/RFC8622, June 2019, <https://www.rfc-editor.org/info/rfc8622>.

[SA-5G] 3GPP, "System Architecture for 5G", TS 23.501, 2019.

[WIFI-ALLIANCE] Wi-Fi Alliance, "Wi-Fi QoS Management Specification Version 2.0", Wi-Fi QoS Management Specification Version 2.0, 2021.

Appendix A. Revision Notes

Note to RFC-Editor: please remove this entire section prior to publication.

- *Individual draft -00, initial document.

- *Individual draft -01, address comments from Martin Duke; Brian Carpenter; Ruediger Geib, and revisions to improve language consistency.

- *Individual draft -02, revise to improve language consistency.

- *Working Group -00, replace individual draft.

- *Working Group -01, address feedback in preparation for IETF 113 Vienna.

- *Working Group -02:

  Consolidate terminology after IETF 113 Vienna.

  Add clarification to RFC2474 and RFC2475 addressed in RFC3260 (Comments from Ruediger Geib).

  Include figures to show the full list of codepoints, their assigned PHBs & impact of ToS Precedence Bleaching.

  Add network categories that differentiate between network operator approaches to DiffServ.

  Add Terminology section to clarify representations of DSCPs.

- *Working Group -03

  Add table to explain DSCP abbreviations (comment from Brian Carpenter).

  Add some refs, improve language consistency (comments from Brian Carpenter).

  Clarify figure captions.

- *Working Group -04

Reference RFC3086 (comment from Brian Carpenter).

Improve references (comments from Ruediger Geib).

Clarify intended document audience and scope (comments from Ruediger Geib).

Clarify terms and language around re-marking, DiffServ domains and nodes, RFC8100 (comments from Ruediger Geib).

More in-depth on mappings specified for mobile networks/MPLS short-pipe (comments from Ruediger Geib).

*Working Group -05

Clarify meaning of RFC2474 with respect to IP precedence (Comments from RG).

Add note on understanding the process of remarking. (comments from Ruediger Geib).

Improve readibility.

Authors' Addresses

Ana Custura
University of Aberdeen
School of Engineering
Fraser Noble Building
Aberdeen
AB24 3UE
United Kingdom

Email: ana@erg.abdn.ac.uk

Godred Fairhurst
University of Aberdeen
School of Engineering
Fraser Noble Building
Aberdeen
AB24 3UE
United Kingdom

Email: gorry@erg.abdn.ac.uk

Raffaello Secchi
University of Aberdeen
School of Engineering
Fraser Noble Building

Aberdeen
AB24 3UE
United Kingdom

Email: r.secchi@abdn.ac.uk