Stream: Internet Engineering Task Force (IETF)
RFC: 9435
Category: Informational
Published: July 2023
ISSN: 2070-1721
Authors: A. Custura, *University of Aberdeen*; G. Fairhurst, *University of Aberdeen*; R. Secchi, *University of Aberdeen*

RFC 9435
Considerations for Assigning a New Recommended Differentiated Services Code Point (DSCP)

Abstract

This document discusses considerations for assigning a new recommended Differentiated Services Code Point (DSCP) for a standard Per-Hop Behavior (PHB). It considers the common observed re-marking behaviors 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 DSCP carried in the Diffserv field of the IP packet header [RFC2474]. A common set of DSCPs are defined for both IPv4 and IPv6, and both network protocols use a common IANA registry [DSCP-registry].

Diffserv associates traffic with a service class and categorizes it into Behavior Aggregates (BAs) [RFC4594]. Configuration guidelines for service classes are provided in [RFC4594]. BAs are associated with a DSCP, which in turn maps to a Per-Hop Behavior (PHB). Treatment differentiation can be achieved by using a variety of schedulers and queues and also 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 Behavior (PDB) that is based on one or more PHBs. The PDB defines which PHB (or set of PHBs) and, hence, for a specific operator, which DSCP (or set of DSCPs) will be associated with specific BAs as the packets pass through a Diffserv domain. It also defines whether the packets are re-marked as they are forwarded (i.e., changing the DSCP of a packet [RFC2475]).

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

Figure 1: 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 re-marking behaviors 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 [NQB-PHB], proposals to build mechanisms using DSCPs in other standards-setting organizations, and the desire to use a common set of DSCPs across a range of infrastructure (e.g., [RFC8622], [NQB-PHB], [AX.25-over-IP]).

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] [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 are represented in text using prefix "0x". Binary values use prefix "0b".

In this document, the symbol "&" denotes a bitwise AND of two unsigned values.

3. Current Usage of DSCPs

This section describes the 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 specified, 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]. A summary of the pools is provided in a table (where 'x' refers to a bit position with value either '0' or '1').

DSCP Pool 1: A pool of 32 codepoints with a format of 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 (EXP) Use or Local Use (LU) but was originally specified to be "preferentially utilized for standardized assignments if Pool 1 is ever exhausted" [RFC2474]. Pool 3 codepoints are now "utilized for standardized assignments (replacing the previous availability for 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].

Note that [RFC4594] previously recommended a Local Use of DSCP values 0x01, 0x03, 0x05, and 0x07 (codepoints with the format of 0b000xx1), until this was updated by [RFC8436].

The DSCP space is shown in the following table. Each row corresponds to one setting of the first three bits of the DSCP field, and each column to one setting of the last three bits of the DSCP field.

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**

Table 1: Currently Assigned DSCPs and Their Assigned PHBs

CS	Class Selector	[RFC2474]

BE	Best Effort (CS0)	[RFC2474]
AF	Assured Forwarding	[RFC2597]
EF	Expedited Forwarding	[RFC3246]
VA	Voice Admit	[RFC5865]
LE	Lower Effort	[RFC8622]

Table 2: Abbreviations for DSCPs and PHB Groups

Table 2 summarizes the DSCP abbreviations used in currently published RFCs, [RFC2474] [RFC2597] [RFC3246] [RFC5865] [RFC8622], as described in the IANA registry [DSCP-registry]. The Default PHB is defined in Section 4.1 of [RFC2474]. This provides Best Effort (BE) forwarding, and the recommended DSCP of '000000' (Section 4.2.2.1 of [RFC2474]). This is the lowest value in the set of Class Selector (CS) DSCPs, and hence is also known as "CS0" [DSCP-registry].

NOTE: [RFC4594] specified a now deprecated use of Class Selector 1 (CS1) as the codepoint for the Lower Effort PHB. [RFC8622] updated [RFC4594] and [RFC8325] and obsoleted [RFC3662], assigning the LE DSCP codepoint to the Lower Effort PHB.

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 (TAs) that might result in re-marking (e.g., [RFC5160] suggests Meta-QoS-Classes to help enable deployment of standard end-to-end QoS classes).

3.2. DSCPs Used for 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). The traffic consists of the network control service class and the OAM service class. This traffic is marked with a value from a set of CS DSCPs. This traffic is often a special case within a provider network, and ingress traffic with these DSCP markings can be re-marked.

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] 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], Section 3.1.

Section 4.2.2.2 of [RFC2474] recommends PHBs selected by CS6 and CS7 "**MUST** give packets a 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 [IETF115-IEPG]. Similarly, another study found many routers re-mark all traffic, except for packets carrying a DSCP with the format 0b11xxxx (i.e., setting the higher order bits to 0b11, see Section 4.2.1 of this document).

4. Re-marking the DSCP

It is a feature of the Diffserv architecture that the Diffserv field of packets can be re-marked at the Diffserv domain boundaries (see Section 2.3.4.2 of [RFC2475]). A DSCP can be re-marked at the ingress of a domain. This re-marking 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 re-marked based on common semantics and agreements between providers at a Diffserv domain boundary. Furthermore, [RFC2474] states that re-marking must occur when there is a possibility of theft or denial-of-service attack.

A packet that arrives with a DSCP that is not associated with a PHB, results in an "unknown DSCP." A node could receive a packet with an "unexpected DSCP" due to misconfiguration, or because there is no consistent policy in place. The treatment of packets that are marked with an unknown or an unexpected DSCP at Diffserv domain boundaries is determined by the policy for a Diffserv domain. If packets are received that are marked with an unknown or an unexpected DSCP by a Diffserv domain interior node, [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 re-marking specified by [RFC2474] is intended for interior nodes within a Diffserv domain. For Diffserv ingress nodes, the traffic conditioning required by [RFC2475] applies first.

Reports measuring existing deployments have defined a set of categories for DSCP re-marking [Cus17] [Bar18] in the following seven observed re-marking behaviors:

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

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

Re-mark-ToS: set the first three bits of the DSCP field to any value different from 0b000 (replace the three 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

Re-mark-DSCP: re-marks all traffic to one or more particular (non-zero) DSCP values

These behaviors are explained in the following subsections and cross-referenced in the remainder of the document.

The network nodes forming a particular path might or might not have supported Diffserv. It is not generally possible for an external observer to determine which mechanism results in a specific re-marking solely from the change in an observed DSCP value.

NOTE: More than one mechanism could result in the same DSCP re-marking (see below). These behaviors were measured on both IPv4 and IPv6 Internet paths between 2017 and 2021 [Cus17]. IPv6 routers were found to perform all the types of re-marking described above to a lesser extent than IPv4 ones.

4.1. Bleaching the DSCP Field

A specific form of re-marking 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 [Cus17].

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]. [RFC2474] defines DSCPs 0bxxx000 as the Class Selector codepoints, where PHBs selected by these codepoints **MUST** meet the "Class Selector PHB Requirements" described in Section 4.2.2.2 of [RFC2474].

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 [Cus17].

4.2.1. Impact of ToS Precedence Bleaching

Bleaching of the ToS Precedence field (see Section 4) 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 re-marking. This re-marking 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 re-marking.

```
+-+-+-+-+-+-+
|5|4|3|2|1|0|
+-+-+-+-+-+-+
|0 0 0|x x x|
+-+-+-+-+-+-+
```

Figure 2: Bits in the Diffserv Field Modified by Bleaching of the ToS Precedence

Figure 2 shows bleaching of the ToS Precedence (see Section 4), 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**

Table 3: DSCP Values

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

Data on the observed re-marking at the time of writing was presented in [IETF115-IEPG].

0/CS0	1/LE	2	3	4	5	6	7
Assigned		Re-marked from AF11..41	EXP/ LU	*		Re-marked from AF13..EF	EXP/ LU

Table 4: 0b000xxx DSCPs

* DSCP 4 has been historically used by the SSH application [Kol10]

Table 4 shows 0b000xxx DSCPs. This highlights any current assignments and whether they are affected by any known re-marking behaviors, such as ToS Precedence Bleaching.

DSCPs of the form 0b000xxx can be impacted by known re-marking behaviors of other assigned DSCPs. For example, ToS Precedence Bleaching of popular DSCPs AF11, AF21, AF31, and AF41 would result in traffic being re-marked with DSCP 2 in the Internet core. The Lower Effort (LE) Per-Hop Behavior PHB uses a DSCP of 1. The DSCP value of 4 has been historically used by the SSH application, following semantics that precede Diffserv [Kol10].

Bleach-ToS-Precedence (see Section 4) of packets with a DSCP 'x' results in the DSCP being re-marked to 'x' & 0x07 and then forwarded using the PHB specified for the resulting DSCP in that Diffserv domain. In subsequent networks, the packet will receive treatment as specified by the domain's operator corresponding to the re-marked codepoint.

A variation of this observed re-marking behavior 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 Re-marking

[RFC2474] states:

> Implementors should note that the DSCP field is six bits wide. DS-compliant nodes **MUST** select PHBs by matching against the entire 6-bit DSCP field, e.g., by treating the value of the field as a table index which is used to select a particular packet handling mechanism which has been implemented in that device.

This replaced re-marking according to ToS precedence (see Section 4) [RFC1349]. These practices based on ToS semantics have not yet been eliminated from deployed networks.

```
+-+-+-+-+-+-+
|5|4|3|2|1|0|
+-+-+-+-+-+-+
|0 0 1|x x x|
+-+-+-+-+-+-+
```

Figure 3: Bits in the Diffserv Field Modified by ToS Precedence Re-marking

Figure 3 shows the ToS Precedence Re-marking (see Section 4) observed behavior, based on Section 3 of [RFC1349]. The bit positions marked 'x' remain unchanged.

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

If ToS Precedence Re-marking occurs, packets are forwarded using the PHB specified for the resulting DSCP in that domain. For example, the AF31 DSCP (0x1a) could be re-marked to either AF11 or AF21. If such a re-marked packet further traverses other Diffserv domains, it would receive treatment as specified by each domain's operator corresponding to the re-marked codepoint.

4.3. Re-marking to a Particular DSCP

A network device might re-mark the Diffserv field of an IP packet based on a local policy with a specific set of DSCPs (see Section 4).

Section 3 of [RFC2474] recommends: "Packets received with an unrecognized codepoint **SHOULD** be forwarded as if they were marked for the Default behavior, and their codepoints should not be changed." Some networks might not follow this recommendation and instead re-mark packets with these codepoints to the default class: CS0 (0x00). This re-marking is sometimes performed using a Multi-Field (MF) classifier [RFC2475] [RFC3290] [RFC4594].

If re-marking occurs, packets are forwarded using the PHB specified for the resulting DSCP in that domain. As an example, re-marking traffic AF31, AF32, and AF33 all to a single DSCP, e.g., AF11, stops any drop probability differentiation, which may have been expressed by these three DSCPs. If such a re-marked packet further traverses other domains, it would receive treatment as specified by the domain's operator corresponding to the re-marked codepoint. Bleaching (see Section 4) is a specific example of this observed re-marking behavior that re-marks to CS0 (0x00) (see Section 4.1).

5. Interpretation of the IP DSCP at Lower Layers

Transmission systems and subnetworks can, and do, utilize 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 re-mark the DSCP used at the IP layer. This function is constrained by designs that utilize fewer than 6 bits to express the service class and, therefore, infer a mapping to a smaller L2 QoS field, for example, the 3-bit Priority Code Point (PCP) field in an IEEE Ethernet 802.1Q header, the 3-bit User Priority (UP) field, or the 3-bit Traffic Class field of Multi-Protocol Label Switching (MPLS). A Treatment Aggregate (TA) [RFC5127] is an optional intermediary mapping group of BAs to PHBs.

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

IEEE 802.1Q specified a 3-bit PCP field, which includes a tag that allows Ethernet frames to be marked as one of eight priority values [IEEE-802.1Q]. Use of this field is described by various documents, including IEEE P802.1p and IEEE 802.1D.

The mapping specified in [IEEE-802.1Q] revises a previous standard, [IEEE-802.1D], in an effort to align with Diffserv practice [RFC4594]. In 802.1Q, the traffic types are specified to match the first three bits of a suitable DSCP (e.g., the first three bits of the Expedited Forwarding (EF) DSCP are mapped to a PCP of 5).

In this mapping, PCP0 is used to indicate the default Best Effort treatment, and PCP1 indicates a background traffic class. The remaining PCP values indicate increasing priority. Internet control traffic can be marked as CS6, and network control is marked as CS7.

Other re-marking behaviors have also been implemented in Ethernet equipment. Historically, a previous standard, [IEEE-802.1D], used both PCP1 (Background) and PCP2 (Spare) to indicate lower priority than PCP0, and some other devices do not assign a lower priority to PCP1.

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 Media Access Control (MAC) functions to support QoS in WLANs using Access Categories (ACs). The upstream UP in the 802.11 header has a 3-bit QoS value. A DSCP can be mapped to the UP. [RFC8622] added a mapping for the LE DSCP to AC_BK (Background).

Most current Wi-Fi implementations use a default mapping that maps the first three bits of the DSCP to the 802.11 UP value. This is an example of equipment still classifying on ToS Precedence (which could be seen as a simple method to map IP layer Diffserv to layers offering only 3-bit QoS codepoints). Then, in turn, this is mapped to the four Wi-Fi Multimedia (WMM) Access Categories. The Wi-Fi Alliance has also specified a more flexible mapping that follows [RFC8325] and provides functions at an Access Point (AP) to re-mark packets as well as a QoS Map that maps each DSCP to an AC [WIFI-ALLIANCE].

```
+-+-+-+-+-+-+
|5|4|3|2|1|0|
+-+-+-+-+-+-+
|x x x|. . .|
+-+-+-+-+-+-+
```

Figure 4: 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] notes inconsistencies that can result from such re-marking and recommends a different mapping to perform this re-marking, dependent on the direction in which a packet is forwarded. It provides recommendations for mapping a DSCP to an IEEE 802.11 UP for interconnection between wired and wireless networks. The recommendation in Section 5.1.2 maps network control traffic, CS6 and CS7, as well as unassigned DSCPs, to UP 0 when forwarding in the upstream direction (wireless-to-wired). It also recommends mapping CS6 and CS7 traffic to UP 7 when forwarding in the downstream direction (Section 4.1 of [RFC8325]).

For other UPs, [RFC8325] recommends a mapping in the upstream direction (wireless-to-wired interconnections) that derives the DSCP from the value of the UP multiplied by 8. This mapping, currently used by some Access Points (APs), can result in a specific DSCP re-marking behavior:

> wherein DSCP values are derived from UP values by multiplying the UP values by 8 (i.e., shifting the three UP bits to the left and adding three additional zeros to generate a DSCP value). This derived DSCP value is then 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 re-marking by Bleach-low (see Section 4).

```
+-+-+-+-+-+-+
|5|4|3|2|1|0|
+-+-+-+-+-+-+
|x  x  x|0  0  0|
+-+-+-+-+-+-+
```

Figure 5: Bits in the Diffserv Field Modified by Re-marking Observed as a Result of 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 Wi-Fi 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]. [RFC5127] describes the aggregation of Diffserv service classes and introduces four Diffserv TAs. Traffic marked with multiple DSCPs can be forwarded in a single MPLS TC.

There are three Label Switching Router (LSR) models: the Pipe, the Short Pipe, and the Uniform Model [RFC3270]. In the Uniform and Pipe models, the egress MPLS router forwards traffic based on the received MPLS TC. The Uniform Model includes an egress DSCP rewrite. With the Short Pipe Model, the egress MPLS router forwards traffic based on the Diffserv DSCP as present at the egress router. If the domain supports IP and MPLS QoS differentiation, controlled behavior requires the DSCP of an (outer) IP header to be assigned or re-written by all domain ingress routers to conform with the domain's internal Diffserv deployment. Note that the Short Pipe Model is prevalent in MPLS domains.

5.2.1. Mapping Specified for MPLS

[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] describes four TAs 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 codepoint scheme [RFC8100] are treated as unexpected and will possibly be re-marked (a Re-mark-DSCP, see Section 4 behavior) or dealt with via additional agreements 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 favors re-marking unexpected DSCP values to zero in the absence of additional agreements, as explained in [RFC8100]. This can result in bleaching (see Section 4).

Treatment Aggregate [RFC8100]	DSCP
Telephony Service Treatment Aggregate	EF VA

Treatment Aggregate [RFC8100]	DSCP
Bulk Real-Time Treatment Aggregate	AF41 (AF42)* (AF43)*
Assured Elastic Treatment Aggregate	AF31 AF32 (AF33)**
Default / Elastic Treatment Aggregate	BE/CS0
Network Control: Local Use (LU)	CS6

Table 5: The Short Pipe MPLS Mapping from [RFC8100]

* May be added

** Reserved for the extension of PHBs

5.3. Mapping Specified for Mobile Networks

Mobile LTE and 5G standards have evolved from older Universal Mobile Telecommunications System (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 Data Network Gateway (P-GW) boundary is determined by network operators. However, TS 23.107 (version 16.0.0, applies to LTE and below) mandates that Differentiated Services, defined by the 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 IP Packet eXchange (IPX) Provider networks [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 provided 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 this situation, the IPX Provider can re-mark the DSCP value to a static default value.

QoS Class in [GSMA-IR.34]	PHB
Conversational	EF
Streaming	AF41

QoS Class in [GSMA-IR.34]	PHB
Interactive (ordered by priority, AF3 highest)	AF31
	AF32
	AF21
	AF11
Background	CS0

Table 6: The PHB Mapping Recommended in the Guidelines Recommended in [GSMA-IR.34]

5.4. Mapping Specified for Carrier Ethernet

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

5.5. Re-marking as a Side Effect of Another Policy

This includes any other re-marking that does not happen as a result of traffic conditioning, such as policies and L2 procedures that result in re-marking 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 re-marking behaviors can arise from interactions with lower layers.

A provider service path may consist of sections where multiple and changing layers use their own code points to determine differentiated forwarding (e.g., IP to MPLS to IP to Ethernet to Wi-Fi).

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. This 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 and Re-marking to a Single DSCP

Section 4 describes re-marking of the DSCP. New DSCP assignments should consider the impact of bleaching or re-marking (see Section 4) to a single DSCP, 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]. Forwarding LE using the default PHB is in line with [RFC8622], but it is recommended to maintain the distinct LE DSCP codepoint end-to-end to allow for differentiated treatment by domains supporting LE. Rewriting the LE DSCP to the default class (CS0) results in an undesired promotion of the priority for LE traffic in such a domain. Bleaching the lower three bits of the DSCP (both Bleach-low and Bleach-some-low in Section 4), as well as re-marking to a particular DSCP, can result in similar changes of priority relative to traffic that is marked with other DSCPs.

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

This considers a proposed DSCP with a codepoint larger than 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 the ToS Precedence is bleached (Bleach-ToS-Precedence and Bleach-some-ToS, Section 4) or ToS Precedence Re-marking (Re-mark-ToS, Section 4) is experienced. For example, it can be desirable to avoid choosing a DSCP that could be re-marked to LE, Lower Effort [RFC8622], which could otherwise potentially result in a priority inversion in the treatment.

6.2.1. Where the Proposed DSCP&0x07=0x01

This considers a proposed DSCP where the least significant 3 bits together represent a value of 1 (i.e., 0b001).

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 Bleach-ToS-Precedence (Section 4) 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 PHB. Care should be taken to consider the implications of re-marking when choosing to assign a DSCP with this format.

6.3. Where the Proposed DSCP <= 0x07 (7)

This considers a proposed DSCP where the DSCP is less than or equal to 7.

ToS Precedence Bleaching or ToS Precedence Re-marking 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 that receives the treatment associated with DSCP 2. New DSCP assignments should consider unexpected consequences arising from this observed re-marking behavior.

6.4. Impact on Deployed Infrastructure

Behavior 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 re-marking behaviors described in the previous section can result from different configurations and approaches:

Networks not re-marking Diffserv:
: A network that either does not implement PHBs or implements one or more PHBs while 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 Service Level Agreements (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]). Operators need to classify the received traffic, assign it to a corresponding PHB, and could re-mark the DSCP to a value that is appropriate for the domain's deployed Diffserv infrastructure.

Networks that re-mark 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 re-mark the DSCP as a result of lower layer interactions.

The DSCP re-marking corresponding to the Bleach-ToS-Precedence (Section 4) observed behavior can arise for various reasons, one of which is old equipment that precedes Diffserv. The same re-marking can also arise in some cases when traffic conditioning is provided by Diffserv routers at operator boundaries or as a result of misconfiguration.

6.5. Considerations to Guide the 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.
- What are the characteristics of the proposed service class? What are the characteristics of the traffic to be carried? What are the expectations for treatment?
- Service classes [RFC4594] that can utilize 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 Internet Draft contain an appropriate request to IANA?

- The value selected for a new DSCP can impact the ability of an operator to apply logical functions (e.g., a bitwise mask) to related codepoints when configuring Diffserv. A suitable value can simplify configurations by aggregating classification on a range of DSCPs. This classification based on DSCP ranges can increase the comprehensibility of documenting forwarding differentiation.
- 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 re-marking behavior, and Section 6.4 identifies root causes for why this re-marking is observed. What is the expected effect of currently-deployed re-marking on the service, end-to-end or otherwise?

7. IANA Considerations

IANA has added the following text as a note at the top of the "Differentiated Services Field Codepoints (DSCP)" registry [DSCP-registry]: "See RFC 9435 for considerations when assigning a new codepoint from the DSCP registry."

8. Security Considerations

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

9. References

9.1. Normative References

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

[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>.

9.2. Informative References

[AX.25-over-IP] Learmonth, I. R., "Internet Protocol Encapsulation of AX.25 Frames", Work in Progress, Internet-Draft, draft-learmonth-intarea-rfc1226-bis-03, 23 May 2021, <https://datatracker.ietf.org/doc/html/draft-learmonth-intarea-rfc1226-bis-03>.

[Bar18] Barik, R., Welzl, M., Elmokashfi, A., Dreibholz, T., and S. Gjessing, "Can WebRTC QoS Work? A DSCP Measurement Study", 2018 30th International Teletraffic Congress (ITC 30), DOI 10.1109/ITC30.2018.00034, September 2018, <https://doi.org/10.1109/ITC30.2018.00034>.

[Cus17] Custura, A., Venne, A., and G. Fairhurst, "Exploring DSCP modification pathologies in mobile edge networks", 2017 Network Traffic Measurement and Analysis Conference (TMA), DOI 10.23919/TMA.2017.8002923, June 2017, <https://doi.org/10.23919/TMA.2017.8002923>.

[GSMA-IR.34] GSM Association, "Guidelines for IPX Provider networks (Previously Inter-Service Provider IP Backbone Guidelines)", Version 17.0, IR.34, May 2021, <https://www.gsma.com/newsroom/wp-content/uploads/IR.34-v17.0.pdf>.

[IEEE-802.11] IEEE, "IEEE Standard for Information Technology - Telecommunications and Information Exchange Between Systems - Local and Metropolitan Area Networks - Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications", DOI 10.1109/IEEESTD.2021.9363693, IEEE Standard 802.11, February 2021, <https://standards.ieee.org/ieee/802.11/7028/>.

[IEEE-802.1D] IEEE, "IEEE Standard for Local and metropolitan area network--Media Access Control (MAC) Bridges", IEEE Standard 802.1D-2004, DOI 10.1109/IEEESTD.2004.94569, June 2004, <https://doi.org/10.1109/IEEESTD.2004.94569>.

[IEEE-802.1Q] IEEE, "IEEE Standard for Local and Metropolitan Area Network--Bridges and Bridged Networks", IEEE Standard 802.1Q-2018, DOI 10.1109/IEEESTD.2018.8403927, July 2018, <https://doi.org/10.1109/IEEESTD.2018.8403927>.

[IETF115-IEPG] Custura, A., "Real-world DSCP Traversal Measurements", November 2022, <https://datatracker.ietf.org/meeting/115/materials/slides-115-iepg-sessa-considerations-for-assigning-dscps-01>.

[ITU-T-Y.1566] ITU-T Recommendation, "Quality of service mapping and interconnection between Ethernet, Internet Protocol and multiprotocol label switching networks", ITU-T Y.1566, July 2012, <https://www.itu.int/rec/T-REC-Y.1566/en>.

[Kol10] Kolu, A., "Subject: bogus DSCP value for ssh", message to the freebsd-stable mailing list, 12 July 2010, <https://lists.freebsd.org/pipermail/freebsd-stable/2010-July/057710.html>.

[MEF-23.1] MEF, "Implementation Agreement MEF 23.1 Carrier Ethernet Class of Service - Phase 2", MEF 23.1, January 2012, <https://mef.net/Assets/Technical_Specifications/PDF/MEF_23.1.pdf>.

[NQB-PHB] 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-18, 10 July 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-tsvwg-nqb-18>.

[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., Ed., 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>.

[RFC3662] Bless, R., Nichols, K., and K. Wehrle, "A Lower Effort Per-Domain Behavior (PDB) for Differentiated Services", RFC 3662, DOI 10.17487/RFC3662, December 2003, <https://www.rfc-editor.org/info/rfc3662>.

[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>.

[RFC5415] Calhoun, P., Ed., Montemurro, M., Ed., and D. Stanley, Ed., "Control And Provisioning of Wireless Access Points (CAPWAP) Protocol Specification", RFC 5415, DOI 10.17487/RFC5415, March 2009, <https://www.rfc-editor.org/info/rfc5415>.

[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 the 5G System (5GS)", TS 23.501, 2019.

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

Acknowledgements

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.

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