TSVWG A. Custura
Internet-Draft G. Fairhurst
Intended status: Informational R. Secchi
Expires: September 16, 2021 University of Aberdeen
March 15, 2021

Considerations for Assigning a new receommended DiffServ Codepoint (DCSP)
draft-custura-tsvwg-dscp-considerations-01

Abstract

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

This individual draft aims to seek comments and contributions from the TSVWG working group.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

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

This Internet-Draft will expire on September 16, 2021.

Copyright Notice

Copyright (c) 2021 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (https://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction

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

This document discusses considerations for assigning a new DiffServ Code Point (DCSP). It considers some commonly observed DSCP remarking behaviours that might be expected to be experienced along an Internet path. It also notes some packet forwarding treatments that a packet can receive when using a specific DSCP.

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., [I-D.learmonth-rfc1226-bis]).

2. Requirements Language

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].

3. Current usage of DSCPs

This section describes current usage of DSCPs.

3.1. IP-Layer Semantics

The DiffServ architecture specifies 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 Per Hop Behavior (PHB)[RFC2474]. When a new PHB is standardized, a recommended DSCP value among those 64 values is normally reserved for that PHB.

The DSCP space is divided into three pools for the purpose of assignment and management [DSCP-registry]. The pools are defined in the following 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 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 0bxxxx01. This was initially available for experimental or local use, but which was originally specified to be preferentially utilised for standardized assignments if Pool 1 is ever exhausted. Pool 3

codepoints are now utilised for standardized assignments and are no longer available for experimental or local use [RFC8436].

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. Several IP standards map treatment aggregates to DSCPs, that might result in remarking: RFC5160 [RFC5160] suggests Meta-QoS-Classes to help enable deployment of standardized end-to-end QoS classes.

Note: A DSCP is sometimes referred to by name, such as "CS1", and sometimes by a decimal, hex, or binary value [DSCP-registry].

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 using a set of Class Selector (CS) DSCPs. Such network control 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).

The DCSP 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].

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 [RFC2475]). A DSCP can be remarked at the ingress of a DiffServ domain. This can be with or without restoring the initial DSCP marking at the egress of the same domain. The Diffserv field can also be re-marked based upon common semantics and agreements between providers at an exchange point.

If packets are received that are marked with an unknown or an unexpected DSCP, RFC2474 [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 preserving DSCP markings by routers that have not been configured to support DiffServ. (See also Section 4.3).

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

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

Bleach-ToS: sets the first three bits of the DCSP field to 0b000 (resets the 3 bits of the former ToS Precedence field) ;

Remark-ToS: sets the first three bits of the DCSP field to 0b001 (replace the 3 bits of the former ToS Precedence field);

Remark-ToS: sets the first three bits of the DCSP field to 0b010 (replace the 3 bits of the former ToS Precedence field);

Reset-low: sets the last three bits of the DCSP field to 0b000;

Reset-some-low: sets 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.

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.

Resetting all the bits of the DiffServ field to 0 is 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 defined ToS precedence field for IP packets in RFC1349 [RFC1349]. Since 1998, this practice has been deprecated by RFC2474 [RFC2474]. RFC 2474 defines codepoints 0x xxx000 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, practices based on ToS semantics have not yet been eliminated from Internet, and these need to still currently be considered when making new DCSP assignments.

4.2.1. Impact of ToS Precedence Bleaching

ToS precedence bleaching (/Bleach-ToS/) is a practice that resets to zero the upper 3 bits of the DSCP field (the former ToS Precedence field), leaving the lower bits unchanged. This behaviour is distinctive of non-DiffServ aware routers and one of the more common manipulations of the DiffServ field.

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

Figure showing the ToS bleaching pathology, based on Section 3 of RFC1349 [RFC1349].

If ToS precedence bleaching occurs, packets with a DSCP 'x' would be remarked and then would be treated according to the PHB specified for 'x' & 0x07. For example, AF31 (0x1a) would be remarked to DSCP 2 (0x02).

A variation of this behaviour clears the top three bits of a DSCP, unless these have values 0b110 or 0b111 (corresponding to CS6 and CS7). 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 RFC1349 [RFC1349] were deprecated by RFC2474 [RFC2474]. These practices based on ToS semantics have not yet been eliminated from Internet.

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

Figure showing the ToS remarking pathology, based on Section 3 of RFC1349 [RFC1349].

A less common behaviour, ToS precedence remarking /Remark-ToS/ sets the upper three bits of the DSCP field to a non-zero value corresponding to a CS PHB. This behaviour is distinctive of non-DiffServ aware routers.

If remarking occurs, packets are treated using the PHB specified for the resulting codepoint. For example, the AF31 DSCP (0x1a) could be remarked to 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 to avoid marking with a specific (set of) DSCPs, /remark/. This is sometimes performed using a Multi-Field (MF) classifier [RFC2475] [RFC3290] [RFC4594]. For example, a common behaviour is to mark all traffic to a single DSCP, thus removing any traffic differentiation (see Section 4.1). Bleaching (/Bleach/) is a specific example of this.

In another example, some networks do not follow the recommendation in RFC2475 [RFC2475], and instead remark packets with an unknown or unexpected DSCP to the default class, CS0 (0x00) to ensure that appropriate DSCPs are used within a DiffServ domain. (e.g., see [RFC8100]).

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

5.1. Mapping Specified for IEEE 802.11

<< This section is currently incomplete. >>

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. 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 six, and network control is marked as seven.

Section 6 of RFC 8325 RFC 8325 [RFC8325] provides a brief overview of IEEE 802.11 QoS. The IEEE 802.11 standards [IEEE-802-11] provides 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 existing WiFi implementations [RFC8325] use a default mapping that utilises the three most significant bits of the DiffServ Field to the 802.11 UP. This is then in turn mapped to the four WiFi Multimedia (WMM) Access Categories.

RFC 8325 [RFC8325] proposes several recommendations for mapping a DSCP to an IEEE 802.11 UP for wired-to-wireless interconnection. The DSCP value of a downstream IP packet can be used (e.g., the Control And Provisioning of Wireless Access Points (CAPWAP) protocol, RFC5415, maps an IP packet Diffserv field to the Diffserv field of outer IP headier in a CAPWAP tunnel).

Some current constraints of WiFi mapping are discussed in section 2 of RFC 8325 [RFC8325]. A QoS profile can be used to limit the maximum DSCP value used for the upstream and downstream traffic.

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

Figure showing the DSCP bits commonly mapped to the UP in 802.11.

This UP mapping can result in a specific DSCP remarking pathology. In the upstream direction, some Access Points (APs) currently use a default UP-to-DSCP mapping RFC8325 [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 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 behaviour can result in remarking /Reset-low/.

RFC8325 [RFC8325] notes inconsistencies that can result from such remarking, and recommends how to perform this remarking.

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

Figure showing the DSCP bits commonly mapped to the UP in 802.11.

5.2. Mappings Specified for MPLS

In MPLS there are eight MPLS Traffic Classes (TCs), which restricts the number of different treatments (e.g., see [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.

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 [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 that the original DSCP marking is not changed when treatment aggregates are used. The key requirement is that the DSCP at the network ingress is restored at the network egress. This is only feasible in general for a small number of DSCPs. In this design, packets marked with other DSCPs can be re-marked (This can result in the /Remark/ behaviour) or dealt with via an additional agreement(s) among the operators of the interconnected networks. Use of the MPLS Short Pipe Model favours re-marking unexpected DSCP values to zero in the absence of an additional agreement(s), as explained in RFC8100 [RFC8100]. This can result in the remarking: /Bleach/.

5.3. Mapping Specified for Mobile Networks

<< This section on Standardized 5QI to QoS characteristics mapping is currently incomplete. >>

[SA-5G].

The GSM Association (GSMA) defines four traffic classes and seven associated DSCPs 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.

5.4. Mappings Specified for Carrier Ethernet

<< This section on metro Ethernet Forum (MEF) mapping is currently incomplete. >>

{{Ref to MEF23.1}}

5.5. Remarking as a Side-effect of Another Policy

The result of applying a QoS policy (such as matching the IP address, or traffic reaching a rate limit) could also result in a packet being remarked to a different DSCP when it is not admitted to a service. Traffic marked with an EF and VA DSCP are often policed by such policies.

Policies and L2 procedures can also result in remarking traffic as a side-effect of other functions (e.g., in response to DDos).

6. Considerations for DSCP Selection

This section provides advice for the assignment of a new DSCP value.

Diffserv domains can use the codepoints in Pool 2 for local or experimental use.

New IETF assignments are normally made in Pool 1, but can also be made from Pool 3.

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 to the default class. This is an example of priority inversion.

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

Although the IETF requires systems use DSCP marking semantics for the former ToS field, the current recommendation is that any new assignment where the codepoint 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 bleaching can unintentionally result in extra traffic aggregated to the same DSCP. For example, after experiencing ToS 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 DiffServ codepoint assignments should consider unexpected consequences arising from ToS bleaching.

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

As a consequence of assigning the LE PHB [RFC8622], the IETF allocated the DSCP 000001b from Pool 3.

When making assignments where the DSCP has a format: xxx 001b, the case of ToS Precedence Bleaching 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 that a DSCP in this category could be remarked as LE.

6.5. Impact on deployed infrastructure

Behaviour of deployed PHBs and conditioning treatments also needs to be considered when assigning a new DSCP. In some domains, a network operator can provide transparent transport of unknown or unassigned DSCPs across their domain. In other domains, policing can ensure that only traffic that matches a policy is permitted to use a specific DSCP (e.g., as in MPLS TC).

7. Acknowledgments

The authors acknowledge the helpful discussions and analysis by Greg White and Thomas Fossati in a draft concerning NQB. We look forward to further comments and review.

8. IANA Considerations

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

This memo includes no request to IANA.

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".

[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>.

[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.

[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", draft-ietf-tsvwg-nqb-03 (work in progress), November 2020.

[I-D.learmonth-rfc1226-bis] Learmonth, I., "Internet Protocol Encapsulation of AX.25 Frames", draft-learmonth-rfc1226-bis-03 (work in progress), May 2020.

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

[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.

[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>.

[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>.

[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>.

[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.

Appendix A. Revision Notes

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

- o Individual draft -00
- o Individual draft -01; Comments from Martin Duke; Brian Carpenter; Ruediger Geib, and revsisions to improve language consistency.

Authors' Addresses

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

Email: ana@erg.abdn.ac.uk

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

Email: gorry@erg.abdn.ac.uk

Raffaello Secchi
University of Aberdeen
School of Engineering
Fraser Noble Building
Aberdeen AB24 3UE
UK

Email: r.secchi@abdn.ac.uk