

Transport Area Working Group
Internet-Draft
Intended status: Standards Track
Expires: May 7, 2020

G. White
CableLabs
T. Fossati
ARM
November 4, 2019

**A Non-Queue-Building Per-Hop Behavior (NQB PHB) for Differentiated
Services
draft-ietf-tsvwg-nqb-00**

Abstract

This document specifies properties and characteristics of a Non-Queue-Building Per-Hop Behavior (NQB PHB). The purpose of this NQB PHB is to provide a separate queue that enables low latency and, when possible, low loss for application-limited traffic flows that would ordinarily share a queue with capacity-seeking traffic. The PHB provides low latency and, when possible, low loss without prioritization and without rate policing, making it suitable for environments where the use of either these features may be restricted. The NQB PHB has been developed primarily for use by access network segments, where queuing delays and queuing loss caused by Queue-Building protocols are manifested, but its use is not limited to such segments. In particular, applications to cable broadband links and mobile network radio and core segments are discussed. This document defines a standard Differentiated Services Code Point (DSCP) to identify Non-Queue-Building flows.

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 May 7, 2020.

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](https://trustee.ietf.org/license-info) 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](#) [2](#)
- [2. Requirements Language](#) [3](#)
- [3. Overview: Non-Queue Building Flows](#) [3](#)
- [4. DSCP Marking of NQB Traffic](#) [4](#)
- [5. Non Queue Building PHB Requirements](#) [5](#)
- [6. Relationship to L4S](#) [6](#)
- [7. Use Cases](#) [6](#)
 - [7.1. DOCSIS Access Networks](#) [6](#)
 - [7.2. Mobile Networks](#) [6](#)
 - [7.3. WiFi Networks](#) [7](#)
- [8. Open Points](#) [7](#)
- [9. Acknowledgements](#) [7](#)
- [10. IANA Considerations](#) [8](#)
- [11. Security Considerations](#) [8](#)
- [12. Informative References](#) [8](#)
- Authors' Addresses [9](#)

1. Introduction

This document defines a Differentiated Services (DS) per-hop behavior (PHB) called "Non-Queue-Building Per-Hop Behavior" (NQB PHB), which is intended to enable networks to provide low latency and low loss for traffic flows that are relatively low data rate and that do not themselves materially contribute to queuing delay and loss. Such Non-Queue-Building flows (for example: interactive voice and video, gaming, machine to machine applications) are application limited flows that are distinguished from traffic flows managed by an end-to-end congestion control algorithm.

The vast majority of packets that are carried by broadband access networks are, in fact, managed by an end-to-end congestion control

algorithm, such as Reno, Cubic or BBR. These congestion control algorithms attempt to seek the available capacity of the end-to-end path (which can frequently be the access network link capacity), and in doing so generally overshoot the available capacity, causing a queue to build-up at the bottleneck link. This queue build up results in queuing delay that the application experiences as variable latency, and may result in packet loss as well.

In contrast to traditional congestion-controlled applications, there are a variety of relatively low data rate applications that do not materially contribute to queuing delay and loss, but are nonetheless subjected to it by sharing the same bottleneck link in the access network. Many of these applications may be sensitive to latency or latency variation, as well as packet loss, and thus produce a poor quality of experience in such conditions.

Active Queue Management (AQM) mechanisms (such as PIE [[RFC8033](#)], DOCSIS-PIE [[RFC8034](#)], or CoDel [[RFC8289](#)]) can improve the quality of experience for latency sensitive applications, but there are practical limits to the amount of improvement that can be achieved without impacting the throughput of capacity-seeking applications when only a few of such flows are present.

The NQB PHB supports differentiating between these two classes of traffic in bottleneck links and queuing them separately in order that both classes can deliver satisfactory quality of experience for their applications.

2. Requirements Language

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.

3. Overview: Non-Queue Building Flows

There are many applications that send traffic at relatively low data rates and/or in a fairly smooth and consistent manner such that they are highly unlikely to exceed the available capacity of the network path between source and sink. These applications do not make extensive use of network buffers, but nonetheless can be subjected to packet delay and delay variation as a result of sharing a network buffer with those that do make use of them. Many of these applications are negatively affected by excessive packet delay and delay variation. Such applications are ideal candidates to be queued

separately from the capacity-seeking applications that are the cause of queue buildup, latency and loss.

These Non-queue-building (NQB) flows are typically UDP flows that don't seek the capacity of the link (examples: online games, voice chat, DNS lookups, real-time IoT analytics data). Here the data rate is essentially limited by the Application itself. In contrast, Queue-building (QB) flows include traffic which uses the Traditional TCP or QUIC, with BBR or other TCP congestion controllers.

There are many application flows that fall very neatly into one or the other of these categories, but there are also application flows that may be in a gray area in between (e.g. they are NQB on higher-speed links, but QB on lower-speed links).

Editor's Note: Do we need to answer the following questions? How can an application determine whether it is queue building or not, given that the sending application is generally not aware of the available capacity of the path to the receiving endpoint? Even if the application were to be aware of the capacity of the path, how could it be sure that the available capacity (considering other flows that may be sharing the path) would be sufficient to result in the application's traffic not causing a queue to form?

4. DSCP Marking of NQB Traffic

This document recommends a DiffServ Code Point (DSCP) of 0x2A to identify packets of NQB flows. (editor's note: this value is subject to change)

It is worthwhile to note that the NQB designation and marking is intended to convey verifiable traffic behavior, not needs or wants. Also, it is important that incentives are aligned correctly, i.e. that there is a benefit to the application in marking its packets correctly, and no benefit to an application in intentionally mismarking its traffic. Thus, a useful property of nodes that support separate queues for NQB and QB flows would be that for NQB flows, the NQB queue provides better performance than the QB queue; and for QB flows, the QB queue provides better performance than the NQB queue. By adhering to these principles, there is no incentive for senders to mismark their traffic as NQB, and further, any mismarking can be identified by the network.

In contrast to the existing standard DSCPs, which are typically only meaningful within a DiffServ Domain (e.g. an AS or an enterprise network), this DSCP is expected to be used end-to-end across the Internet. Some network operators typically bleach (zero out) the Diffserv field on ingress into their network [[Custura](#)], and in some

cases apply their own DSCP for internal usage. Networks that support the NQB PHB SHOULD preserve the NQB DSCP when forwarding via an interconnect from another network.

5. Non Queue Building PHB Requirements

A node supporting the NQB PHB makes no guarantees on latency or data rate for NQB marked flows, but instead aims to provide a bound on queuing delay for as many such marked flows as it can, and shed load when needed.

A node supporting the NQB PHB MUST provide a queue for non-queue-building traffic separate from the queue used for queue-building traffic.

NQB traffic SHOULD NOT be rate limited or rate policed separately from queue-building traffic of equivalent importance.

The NQB queue SHOULD be given equal priority compared to queue-building traffic of equivalent importance. The node SHOULD provide a scheduler that allows QB and NQB traffic of equivalent importance to share the link in a fair manner, e.g. a deficit round-robin scheduler with equal weights.

A node supporting the NQB PHB SHOULD treat traffic marked as Default (DSCP=0x00) as QB traffic having equivalent importance to the NQB marked traffic.

The NQB queue SHOULD have a buffer size that is significantly smaller than the buffer provided for QB traffic.

It is possible that due to an implementation error or misconfiguration, a QB flow would end up getting mismarked as NQB, or vice versa. In the case of an NQB flow that isn't marked as NQB and ends up in the QB queue, it would only impact its own quality of service, and so it seems to be of lesser concern. However, a QB flow that is mismarked as NQB would cause queuing delays for all of the other flows that are sharing the NQB queue.

To prevent this situation from harming the performance of the real NQB flows, network elements that support differentiating NQB traffic SHOULD (editor's note: SHOULD vs MUST is TBD) support a "traffic protection" function that can identify QB flows that are mismarked as NQB, and reclassify those flows/packets to the QB queue. Such a function SHOULD be implemented in an objective and verifiable manner, basing its decisions upon the behavior of the flow rather than on application-layer constructs. One example algorithm can be found in [[I-D.briscoe-docsis-q-protection](#)].

6. Relationship to L4S

The dual-queue mechanism described in this draft is intended to be compatible with [[I-D.ietf-tsvwg-l4s-arch](#)], with the result being that NQB traffic and L4S traffic can share the low-latency queue in an L4S dual-queue node [[I-D.ietf-tsvwg-aqm-dualq-coupled](#)].

7. Use Cases

7.1. DOCSIS Access Networks

Residential cable broadband Internet services are commonly configured with a single bottleneck link (the access network link) upon which the service definition is applied. The service definition, typically an upstream/downstream data rate tuple, is implemented as a configured pair of rate shapers that are applied to the user's traffic. In such networks, the quality of service that each application receives, and as a result, the quality of experience that it generates for the user is influenced by the characteristics of the access network link.

To support the NQB PHB, cable broadband services **MUST** be configured to provide a separate queue for NQB marked traffic. The NQB queue **MUST** be configured to share the service's rate shaping bandwidth with the queue for QB traffic.

7.2. Mobile Networks

Historically, mobile networks have been configured to bundle all flows to and from the Internet into a single "default" EPS bearer whose buffering characteristics are not compatible with low-latency traffic. The established behaviour is rooted partly in the desire to prioritise operators' voice services over competing over-the-top services and partly in the fact that the addition of bearers was prohibitive due to expense. Of late, said consideration seems to have lost momentum (e.g., with the rise in Multi-RAB (Radio Access Bearer) devices) and the incentives might now be aligned towards allowing a more suitable treatment of Internet real-time flows.

To support the NQB PHB, the mobile network **MUST** be configured to give UEs a dedicated, low-latency, non-GBR, EPS bearer, e.g. one with QCI 7, in addition to the default EPS bearer; or a Data Radio Bearer with 5QI 7 in a 5G system (see Table 5.7.4-1: Standardized 5QI to QoS characteristics mapping in [[SA-5G](#)]).

A packet carrying the NQB DSCP **SHOULD** be routed through the dedicated low-latency EPS bearer. A packet that has no associated NQB marking **SHOULD** be routed through the default EPS bearer.

7.3. WiFi Networks

WiFi networking equipment compliant with 802.11e generally supports either four or eight transmit queues and four sets of associated EDCA parameters (corresponding to the four WiFi Multimedia Access Categories) that are used to enable differentiated media access characteristics. Implementations typically utilize the IP DSCP field to select a transmit queue, but should be considered as Non-Differentiated Services-Compliant Nodes as described in [Section 4 of \[RFC2475\]](#). As a result this document discusses interoperability with WiFi networks, as opposed to PHB compliance.

As discussed in [\[RFC8325\]](#), most existing implementations use a default DSCP to User Priority mapping that utilizes the most significant three bits of the DiffServ Field to select "User Priority" which is then mapped to the four WMM Access Categories. In order to increase the likelihood that NQB traffic is provided a separate queue from QB traffic in existing WiFi equipment, the 0x2A codepoint is preferred for NQB. This would map NQB to UP_5 which is in the "Video" Access Category.

Systems that utilize [\[RFC8325\]](#), SHOULD map the NQB codepoint to UP_5 in the "Video" Access Category.

In order to preserve the incentives principle, WiFi systems SHOULD configure the EDCA parameters for the Video Access Category to match those of the Best Effort Access Category.

8. Open Points

- o Traffic Protection: SHOULD or MUST?
- o Traffic Protection: what are the detailed requirements?
- o DSCP value vis a vis WMM: VideoAC or BestEffortAC?
- o Are there hidden requirements in [Section 9](#) of the individual draft?
- o Is more discussion needed around applicability in order to give guidance to application devs?

9. Acknowledgements

Thanks to Bob Briscoe, Greg Skinner, Toke Hoeiland-Joergensen, Luca Muscariello, David Black, Sebastian Moeller, Ruediger Geib, Jerome Henry, Steven Blake, Jonathan Morton, Roland Bless, Kevin Smith, Martin Dolly, and Kyle Rose for their review comments.

10. IANA Considerations

This draft proposes the registration of a standardized DSCP = 0x2A to denote Non-Queue-Building behavior.

11. Security Considerations

There is no incentive for an application to mismark its packets as NQB (or vice versa). If a queue-building flow were to mark its packets as NQB, it could experience excessive packet loss (in the case that traffic-protection is not supported by a node) or it could receive no benefit (in the case that traffic-protection is supported). If a non-queue-building flow were to fail to mark its packets as NQB, it could suffer the latency and loss typical of sharing a queue with capacity seeking traffic.

The NQB signal is not integrity protected and could be flipped by an on-path attacker. This might negatively affect the QoS of the tampered flow.

12. Informative References

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

[I-D.briscoe-docsis-q-protection]
Briscoe, B. and G. White, "Queue Protection to Preserve Low Latency", [draft-briscoe-docsis-q-protection-00](#) (work in progress), July 2019.

[I-D.ietf-tsvwg-aqm-dualq-coupled]
Schepper, K., Briscoe, B., and G. White, "DualQ Coupled AQMs for Low Latency, Low Loss and Scalable Throughput (L4S)", [draft-ietf-tsvwg-aqm-dualq-coupled-10](#) (work in progress), July 2019.

[I-D.ietf-tsvwg-l4s-arch]
Briscoe, B., Schepper, K., Bagnulo, M., and G. White, "Low Latency, Low Loss, Scalable Throughput (L4S) Internet Service: Architecture", [draft-ietf-tsvwg-l4s-arch-04](#) (work in progress), July 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>>.

- [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>>.
- [RFC8033] Pan, R., Natarajan, P., Baker, F., and G. White, "Proportional Integral Controller Enhanced (PIE): A Lightweight Control Scheme to Address the Bufferbloat Problem", [RFC 8033](#), DOI 10.17487/RFC8033, February 2017, <<https://www.rfc-editor.org/info/rfc8033>>.
- [RFC8034] White, G. and R. Pan, "Active Queue Management (AQM) Based on Proportional Integral Controller Enhanced (PIE) for Data-Over-Cable Service Interface Specifications (DOCSIS) Cable Modems", [RFC 8034](#), DOI 10.17487/RFC8034, February 2017, <<https://www.rfc-editor.org/info/rfc8034>>.
- [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>>.
- [RFC8289] Nichols, K., Jacobson, V., McGregor, A., Ed., and J. Iyengar, Ed., "Controlled Delay Active Queue Management", [RFC 8289](#), DOI 10.17487/RFC8289, January 2018, <<https://www.rfc-editor.org/info/rfc8289>>.
- [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>>.
- [SA-5G] 3GPP, "System Architecture for 5G", TS 23.501, 2019.

Authors' Addresses

Greg White
CableLabs

Email: g.white@cablelabs.com

Thomas Fossati
ARM

Email: Thomas.Fossati@arm.com

