Workgroup: Transport Area Working Group
Internet-Draft: draft-ietf-tsvwg-nqb-21
Updates: rfc8325 (if approved)
Published: 7 November 2023
Intended Status: Standards Track
Expires: 10 May 2024
Authors: G. White T. Fossati R. Geib
CableLabs Linaro Deutsche Telekom

A Non-Queue-Building Per-Hop Behavior (NQB PHB) for Differentiated Services

Abstract

This document specifies properties and characteristics of a Non-Queue-Building Per-Hop Behavior (NQB PHB). The NQB PHB provides a shallow-buffered, best-effort service as a complement to a Default deep-buffered best-effort service for Internet services. The purpose of this NQB PHB is to provide a separate queue that enables smooth (i.e. non-bursty), low-data-rate, application-limited traffic microflows, which would ordinarily share a queue with bursty and capacity-seeking traffic, to avoid the latency, latency variation and loss caused by such traffic. This PHB is implemented without prioritization and can be implemented without rate policing, making it suitable for environments where the use of these features is 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, Wi-Fi links, and mobile network radio and core segments are discussed. This document recommends a specific Differentiated Services Code Point (DSCP) to identify Non-Queue-Building microflows.

[NOTE (to be removed by RFC-Editor): This document references an ISE submission draft (I-D.briscoe-docsis-q-protection) that is approved for publication as an RFC. This draft should be held for publication until the queue protection RFC can be referenced.]

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

This document defines a Differentiated Services per-hop behavior (PHB) called "Non-Queue-Building Per-Hop Behavior" (NQB PHB), which isolates traffic microflows (see [RFC2475] for the definition of a microflow) that are relatively low data rate and that do not themselves materially contribute to queuing delay and loss, allowing them to avoid the queuing delays and losses caused by other traffic. Such Non-Queue-Building microflows (for example: interactive voice, game sync packets, machine-to-machine applications, DNS lookups, and real-time IoT analytics data) are low-data-rate application-limited microflows that are distinguished from bursty traffic microflows and high-data-rate traffic microflows managed by a classic congestion control algorithm (as defined in [RFC9330]), both of which cause queuing delay and loss.

In accordance with IETF guidance in [RFC2914] and [RFC8085], most packets carried by broadband access networks are managed by an end-to-end congestion control algorithm. Many of the commonly-deployed congestion control algorithms, such as Reno, Cubic or BBR, are designed 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 delay (variable latency) and packet loss that can affect all the applications that are sharing the bottleneck link. Moreover, many bottleneck links implement a relatively deep buffer (100 ms or more) in order to enable these congestion control algorithms to effectively use the link, which exacerbates the latency and latency variation experienced.

In contrast to applications that frequently cause queuing delay, 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 can 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. For example, AQMs generally allow a significant amount of queue depth variation to accommodate the behaviors of congestion control algorithms such as Reno and Cubic. If the AQM attempted to control the queue much more tightly, applications using those algorithms would not perform well. Alternatively, flow queuing systems, such as fq_codel [RFC8290] can be employed to isolate microflows from one another, but these are not appropriate for all bottleneck links, due to complexity or other reasons.

The NQB PHB supports differentiating between these two classes of traffic in bottleneck links and queuing them separately so that both classes can deliver satisfactory quality of experience for their applications. In particular, the NQB PHB provides a shallow-buffered, best-effort service as a complement to a Default deep-buffered best-effort service. This PHB is primarily applicable for high-speed broadband access network links, where there is minimal aggregation of traffic, and deep buffers are common. The applicability of this PHB to lower-speed links is discussed in Section 5.

To be clear, a network implementing the NQB PHB solely provides isolation for traffic classified as behaving in conformance with the NQB DSCP (and optionally enforces that behavior). It is the NQB senders' behavior itself which results in low latency and low loss.

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. Context

3.1. Non-Queue-Building Behavior

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. Some of these applications are transactional in nature, and might only send one packet (or a few packets) per RTT. These applications might themselves only cause very small, transient queues to form in network buffers, but nonetheless they can be subjected to packet delay and delay variation as a result of sharing a network buffer with applications that tend to cause large and/or standing queues to form. These applications typically implement a response to network congestion that consists of discontinuing (or significantly reducing) transmissions. 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 applications that are the cause of queue build-up, latency and loss.

In contrast, Queue-Building (QB) microflows include those that use TCP or QUIC, with Cubic, Reno or other TCP congestion control algorithms that probe for the link capacity and induce latency and loss as a result. Other types of QB microflows include those that send at a high burst rate even if the long-term average data rate is much lower.

3.2. Relationship to the Diffserv Architecture

The IETF has defined the Differentiated Services architecture [RFC2475] with the intention that it allows traffic to be marked in a manner that conveys the performance requirements of that traffic either qualitatively or in a relative sense (i.e. priority). The architecture defines the use of the Diffserv field [RFC2474] for this purpose, and numerous RFCs have been written that describe recommended interpretations of the values (Diffserv Code Points) of the field, and standardized treatments (traffic conditioning and per-hop-behaviors) that can be implemented to satisfy the performance requirements of traffic so marked.

While this architecture is powerful and flexible enough to be configured to meet the performance requirements of a variety of applications and traffic categories, or to achieve differentiated service offerings, it has not been used for these purposes end-to-end across the Internet.

This is in part due to the fact that meeting the performance requirements of an application in an end-to-end context involves all the networks in the path agreeing on what those requirements are and sharing an interest in meeting them. In many cases this is made more difficult since the performance "requirements" are not strict ones (e.g., applications will degrade in some manner as loss/latency/jitter increase), so the importance of meeting them for any particular application in some cases involves a judgment as to the value of avoiding some amount of degradation in quality for that application in exchange for an increase in the degradation of another application.

Further, in many cases the implementation of Diffserv PHBs has historically involved prioritization of service classes with respect to one another, which sets up the zero-sum game alluded to in the previous paragraph, and results in the need to limit access to higher priority classes via mechanisms such as access control, admission control, traffic conditioning and rate policing, and/or to meter and bill for carriage of such traffic. These mechanisms can be difficult or impossible to implement in an end-to-end context.

Finally, some jurisdictions impose regulations that limit the ability of networks to provide differentiation of services, in large part this seems to be based on the belief that doing so necessarily involves prioritization or privileged access to bandwidth, and thus a benefit to one class of traffic always comes at the expense of another.

In contrast, the NQB PHB has been designed with the goal that it avoids many of these issues, and thus could conceivably be deployed end-to-end across the Internet. The intent of the NQB DSCP is that it signals verifiable behavior that permits the sender to request differentiated treatment. Also, the NQB traffic is to be given a separate queue with priority equal to Default traffic and given no reserved bandwidth other than the bandwidth that it shares with Default traffic. As a result, the NQB PHB does not aim to meet specific application performance requirements. Instead, the goal of the NQB PHB is to provide statistically better loss, latency, and jitter performance for traffic that is itself only an insignificant contributor to those degradations. The PHB is also designed to minimize any incentives for a sender to mismark its traffic, since neither higher priority nor reserved bandwidth are being offered. These attributes eliminate many of the trade-offs that underlie the handling of differentiated service classes in the Diffserv architecture as it has traditionally been defined. These attributes also significantly simplify access control and admission control functions, reducing them to simple verification of behavior. This aspect is discussed further in Section 4.1 and Section 5.2.

The NQB PHB is therefore intended for the prevalent situation where the performance requirements of applications cannot be assured end-to-end, and as a result, applications cannot feasibly place requirements on the network. Instead, many applications have evolved to make the best out of the network environment that they find themselves in. In this context, the NQB PHB provides a better network environment for many applications that send data at relatively low data rates.

In regards to comparison between the NQB PHB and other standardized PHBs in the Diffserv series, the closest similarity is to the Expedited Forwarding (EF) PHB [RFC3246], which also intends to enable low loss, low delay, and low jitter services. The main distinctions between NQB and EF are discussed in Appendix B.

In nodes that support multiple DiffServ Service Classes, NQB traffic is to be treated as a part of the Default treatment. Capacity assigned to this class is not prioritized with respect to other classes, AFxx, EF, etc. Of course, traffic marked as NQB could (like other Default traffic) be prioritized with respect to Lower-Effort (LE) [RFC8622] (i.e. the NQB queue would be emptied in a priority sequence before the LE queue).

3.3. Relationship to L4S

The NQB DSCP and PHB described in this draft have been defined to operate independently of the experimental L4S Architecture [RFC9330]. Nonetheless, traffic marked with the NQB DSCP is intended to be compatible with L4S [RFC9330], with the result being that NQB traffic and L4S traffic can share the low-latency queue in an L4S DualQ node [RFC9332]. Compliance with the DualQ Coupled AQM requirements (Section 2.5 of [RFC9332]) is considered sufficient to support the NQB PHB requirement of fair allocation of bandwidth between the QB and NQB queues (Section 5). Note that these requirements in turn require compliance with all the requirements in Section 5 of [RFC9331].

Applications that comply with both the NQB sender requirements in Section 4.1 and the L4S "Prague" requirements in Section 4 of [RFC9331] could mark their packets both with the NQB DSCP and with the ECT(1) value. NQB network functions SHOULD treat packets marked with the NQB DSCP uniformly, regardless of the value of the ECN field. Here, NQB network functions means the traffic protection function (defined in Section 5.2) and any re-marking/traffic policing function designed to protect unmanaged networks (as described in Section 4.4.1). L4S network functions SHOULD treat packets marked with the NQB DSCP and ECT(1) or CE the same as packets marked with the Default DSCP and the same ECN value. Here, L4S network functions means the L4S Network Node functions (Section 5 of [RFC9331]), and any mechanisms designed to protect the L4S queue (such as those discussed in Section 8.2 of [RFC9330]). The processing by an L4S node of an ECT(0) packet that is classified to the L queue (e.g. as a result of being marked with a NQB DSCP) is specified in Section 5.4.1.1 of [RFC9331] and Section 2.5.1.1 of [RFC9332].

4. DSCP Marking of NQB Traffic

4.1. Non-Queue-Building Sender Requirements

Microflows that are eligible to be marked with the NQB DSCP are typically UDP microflows that send traffic at a low data rate relative to typical network path capacities. Here the data rate is limited by the application itself rather than by network capacity - these microflows send at a data rate of no more than about 1 percent of the "typical" network path capacity. In today's network, where access network data rates are typically on the order of 50 Mbps or more (and see Section 6.1 for a discussion of cases where this isn't true), this implies 500 kbps as an upper limit. In addition, these microflows are required to be sent in a smooth (i.e. paced) manner, where the number of bytes sent in any time interval "T" is less than or equal to R * T + 1500 bytes, where "R" is the maximum rate described above.

Microflows marked with the NQB DSCP are expected to comply with existing guidance for safe deployment on the Internet, including the guidance around response to network congestion, for example the requirements in [RFC8085] and Section 2 of [RFC3551] (also see the circuit breaker limits in Section 4.3 of [RFC8083] and the description of inelastic pseudowires in Section 4 of [RFC7893]). The fact that a microflow's data rate is low relative to typical network capacities is no guarantee that sufficient capacity exists in any particular network, and it is the responsibility of the application to detect and react appropriately if the network capacity is insufficient. To be clear, the description of NQB-marked microflows in this document is not to be interpreted as suggesting that applications generating such microflows are in any way exempt from this responsibility. One way that an application marking its traffic as NQB can handle this is to implement a low latency congestion control mechanism as described in [RFC9331].

Microflows that align with the description of behavior in the preceding paragraphs in this section SHOULD be identified to the network using a Diffserv Code Point (DSCP) of 45 (decimal) so that their packets can be queued separately from QB microflows. The choice of the DSCP value 45 (decimal) is motivated in part by the desire to achieve separate queuing in existing Wi-Fi networks (see Section 7.3) and by the desire to make implementation of the PHB simpler in network gear that has the ability to classify traffic based on ranges of DSCP values (see Section 4.3 for further discussion).

The consideration as to whether an application chooses to mark its traffic as NQB involves the risk of being subjected to a traffic protection algorithm (see Section 5.2) if it contributes to the formation of a queue in a node that supports the PHB. This could result in the excess traffic being discarded or queued separately as default traffic (and thus potentially delivered out of order). As a result, if a microflow's traffic exceeds the rate equation provided in the first paragraph of this section, the application SHOULD NOT mark this traffic with the NQB DSCP. In such a case, the application could instead consider implementing a low latency congestion control mechanism as described in [RFC9331]. At the time of writing, it is believed that 500 kbps is a reasonable upper bound on instantaneous traffic rate for a microflow marked with the NQB DSCP on the Internet. This value is of course subject to the context in which the application is expected to be deployed.

The sender requirements outlined in this section are all related to observable attributes of the packet stream, which makes it possible for network elements (including nodes implementing the PHB) to monitor for inappropriate usage of the DSCP, and take action (such as discarding or re-marking) on traffic that does not comply. This functionality, when implemented as part of the PHB is described in Section 5.2.

4.2. Aggregation of the NQB DSCP into another Diffserv PHB

It is RECOMMENDED that networks and nodes that do not support the NQB PHB be configured to treat traffic marked with the NQB DSCP the same as traffic with the "Default" DSCP. This includes networks and nodes that aggregate service classes as discussed in [RFC5127] and [RFC8100], in which case this recommendation would result in traffic marked with the NQB DSCP being aggregated into the Elastic Treatment Aggregate (for [RFC5127] networks) or the Default / Elastic Treatment Aggregate (for [RFC8100] networks).

In nodes that do not typically experience congestion (for example, many backbone and core network switches), forwarding packets with the NQB DSCP using the Default treatment might be sufficient to preserve loss/latency/jitter performance for NQB traffic.

In nodes that do experience congestion, forwarding packets with the NQB DSCP using the Default treatment could result in degradation of loss/latency/jitter performance but nonetheless preserves the incentives described in Section 5.

Aggregating traffic marked with the NQB DSCP into a PHB designed for real-time, latency sensitive traffic (e.g. the (Bulk) Real-Time Treatment Aggregate), might better preserve loss/latency/jitter performance in the presence of congestion, but would need to be done with consideration of the risk of creating an incentive for non-compliant traffic to be mis-marked as NQB.

Networks and nodes that do not support the NQB PHB should only classify packets with the NQB DSCP value into the appropriate treatment aggregate, or encapsulate such packets for purposes of aggregation, and SHOULD NOT re-mark them with a different DSCP. This preservation of the NQB DSCP value enables hops further along the path to provide the NQB PHB successfully. This aligns with recommendations in [RFC5127].

4.3. Aggregation of other DSCPs in the NQB PHB

Operators of nodes that support the NQB PHB could choose to aggregate other service classes into the NQB queue. This is particularly useful in cases where specialized PHBs for these other service classes are not provided. Candidate service classes for this aggregation would include those that carry low-data-rate inelastic traffic that has low to very-low tolerance for loss, latency and/or jitter. Operators would need to use their own judgment based on the actual traffic characteristics in their networks in deciding whether or not to aggregate other service classes / DSCPs with NQB. For networks that use the [RFC4594] service class definitions, this could include Telephony (EF/VA), Signaling (CS5), and possibly Real-Time Interactive (CS4) (depending on data rate). In some networks, equipment limitations may necessitate aggregating a range of DSCPs (e.g. traffic marked with DSCPs 40-47 (decimal), i.e., those whose three MSBs are 0b101). As noted in Section 4.1, the choice of the DSCP value 45 (decimal) is motivated in part by the desire to make this aggregation simpler in network equipment that can classify packets via comparing the DSCP value to a range of configured values.

A node providing only a NQB queue and a Default queue may obtain an NQB performance similar to that of EF, as described by [RFC2598]. Some caveats and differences are discussed in Appendix B.

4.4. End-to-end usage and DSCP Re-marking

In contrast to some existing standard PHBs, many of which are typically only used within a Diffserv Domain (e.g., an AS or an enterprise network), this PHB is expected to be used end-to-end across the Internet, wherever suitable operator agreements apply. Under the [RFC2474] model, this requires that the corresponding DSCP is recognized and mapped across network boundaries accordingly.

If NQB support is extended across a DiffServ domain boundary, the interconnected networks agreeing to support NQB SHOULD use the DSCP value 45 (decimal) for NQB at network interconnection, unless a different DSCP is explicitly documented in the TCA (Traffic Conditioning Agreement, see [RFC2475]) for that interconnection. Similar to the handling of DSCPs for other PHBs (and as discussed in

[RFC2475]), networks can re-mark NQB traffic to a DSCP other than 45 (decimal) for internal usage. To ensure reliable end-to-end NQB PHB treatment, the appropriate NQB DSCP would need to be restored when forwarding to another network.

4.4.1. Interoperability with Non-DS-Capable Domains

As discussed in Section 4 of [RFC2475], there may be cases where a network operator that supports Diffserv is delivering traffic to another network domain (e.g. a network outside of their administrative control), where there is an understanding that the downstream domain does not support Diffserv or there is no knowledge of the traffic management capabilities of the downstream domain, and no agreement in place. In such cases, Section 4 of [RFC2475] suggests that the upstream domain opportunistically re-mark traffic with a Class Selector codepoint or DSCP 0 (Default) under the assumption that traffic so marked would be handled in a predictable way by the downstream domain.

In the case of a network that supports the NQB PHB (and carries traffic marked with the recommended NQB DSCP value) the same concerns apply. In particular, since the recommended NQB DSCP value could be given high priority in some non-DS-compliant network gear (e.g., legacy Wi-Fi APs as described in Section 7.3.1), it is RECOMMENDED that the operator of the upstream domain re-mark NQB traffic to DSCP 0 (Default) before delivering traffic into a non-DS-capable domain. Network equipment designed for such environments, SHOULD by default re-mark NQB traffic to DSCP 0, and SHOULD support the ability to disable this re-marking.

As an alternative to re-marking, such an operator could deploy a traffic protection (see Section 5.2) or a shaping/policing function on traffic marked with the NQB DSCP that minimizes the potential for negative impacts on Default traffic, should the downstream domain treat traffic with the NQB DSCP as high priority. It should be noted that a traffic protection function as defined in this document might only provide protection from issues occuring in subsequent network hops if the device implementing the traffic protection function is the bottleneck link on the path, so it might not be a solution for all situations. In the case that a traffic policing function or a rate shaping function is applied to the aggregate of NQB traffic destined to such a downstream domain, the policer/shaper rate SHOULD be set to either 5% of the interconnection data rate, or 5% of the typical rate for such interconnections, whichever is greater, with excess traffic being either re-marked and classified for Default forwarding or dropped. A traffic policing function SHOULD allow approximately 100 ms of burst tolerance (e.g. a token bucket depth equal to 100 ms multiplied by the policer rate). A traffic shaping function SHOULD allow approximately 10 ms of burst tolerance, and no more than 50 ms of buffering. The burst tolerance values recommended here are intended to reduce the degradation that could be introduced to latency and loss sensitive traffic marked NQB without significantly degrading Default traffic.

The recommendation to limit NQB traffic to 5% is based on an assumption that internal links in the downstream domain could have data rates as low as one tenth of the interconnect rate, in which case if the entire aggregate of NQB traffic traversed a single instance of such a link, the aggregate would consume no more than 50% of that link's capacity. This SHOULD be adjusted based on any knowledge of the local network environment that is available.

4.5. The NQB DSCP and Tunnels

[RFC2983] discusses tunnel models that support Diffserv. It describes a "uniform model" in which the inner DSCP is copied to the outer header at encapsulation, and the outer DSCP is copied to the inner header at decapsulation. It also describes a "pipe model" in which the outer DSCP is not copied to the inner header at decapsulation. Both models can be used in conjunction with the NQB PHB. In the case of the pipe model, any DSCP manipulation (re-marking) of the outer header by intermediate nodes would be discarded at tunnel egress, potentially improving the possibility of achieving NQB treatment in subsequent nodes.

As is discussed in [RFC2983], tunnel protocols that are sensitive to reordering can result in undesirable interactions if multiple DSCP PHBs are signaled for traffic within a tunnel instance. This is true for traffic marked with the NQB DSCP as well. If a tunnel contains a mix of QB and NQB traffic, and this is reflected in the outer DSCP in a network that supports the NQB PHB, it would be necessary to avoid a reordering-sensitive tunnel protocol. Additionally, since networks supporting the NQB PHB could implement a traffic protection mechanism (see Section 5.2) that results in out-of-order delivery to microflows that are marked with the NQB DSCP, it is RECOMMENDED that reordering-sensitive tunnel protocols not be used with NQB-marked traffic.

5. Non-Queue-Building PHB Requirements

For the NQB PHB to succeed, it is important that incentives are aligned correctly, i.e., that there is a benefit to the application in marking its packets correctly, and a disadvantage (or at least no benefit) to an application in intentionally mismarking its traffic. Thus, a useful property of nodes (i.e. network switches and routers) that support separate queues for NQB and QB microflows is that for microflows consistent with the NQB sender requirements in Section 4.1, the NQB queue would likely be a better choice than the

QB queue; and for microflows inconsistent with those requirements, the QB queue would likely be a better choice than the NQB queue (this is discussed further in this section and Section 11). By adhering to these principles, there is no incentive for senders to mismark their traffic as NQB. As mentioned previously, the NQB designation and marking is intended to convey verifiable traffic behavior, as opposed to simply a desire for differentiated treatment. As a result, any mismarking can be identified by the network.

5.1. Primary Requirements

A node supporting the NQB PHB makes no guarantees on latency or data rate for NQB-marked microflows, but instead aims to provide an upper-bound to queuing delay for as many such marked microflows 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 Default traffic.

A node supporting the NQB PHB SHOULD NOT rate limit or rate police the aggregate of NQB traffic separately from Default traffic. An exception to this recommendation is discussed in Section 4.4.1. Note also that Section 5.2 discusses potential uses of per-microflow (rather than aggregate) rate policing.

The NQB queue SHOULD be given equivalent forwarding preference compared to Default. The node SHOULD provide a scheduler that allows NQB and Default traffic to share the link in a manner that treats the two classes equally, e.g., a deficit round-robin scheduler with equal weights. A node that provides rate limits or rate guarantees for Default traffic SHOULD ensure that such limits and/or guarantees are shared with NQB traffic in a manner that treats the two classes equally. This could be supported using a hierarchical scheduler where the rate limits and guarantees are configured on a parent class, and the two queues (Default and NQB) are arranged as the children of the parent class and given equal access to the capacity configured for the parent class (e.g. with equal DRR scheduling). Compliance with these recommendations helps to ensure that there are no incentives for QB traffic to be mismarked as NQB.

A node supporting the NQB PHB SHOULD by default classify packets marked with the NQB DSCP 45 (decimal) into the queue for Non-Queue-Building traffic. A node supporting the NQB PHB MUST support the ability to configure the DSCP that is used to classify packets into the queue for Non-Queue-Building traffic. A node supporting the NQB PHB SHOULD support the ability to configure multiple DSCPs that are used to classify packets into the queue for Non-Queue-Building traffic.

Support for the NQB PHB is advantageous at bottleneck nodes. Many bottleneck nodes have a relatively deep buffer for Default traffic (e.g., roughly equal to the base RTT of the expected connections, which could be tens or hundreds of ms). Providing a similarly deep buffer for the NQB queue would be at cross purposes to providing very low queueing delay and would erode the incentives for QB traffic to be marked correctly at such a bottleneck node. The NQB queue SHOULD have a buffer size that is significantly smaller than the buffer provided for Default traffic. It is RECOMMENDED to configure an NQB buffer size less than or equal to 10 ms at the shared NQB/Default egress rate.

While not fully described in this document, it may be possible for network equipment to implement a separate QB/NQB pair of queues for additional service classes beyond the Default PHB / NQB PHB pair.

In some cases, existing network gear has been deployed that cannot readily be upgraded or configured to support the PHB requirements. This equipment might however be capable of loosely supporting an NQB service – see Section 7.3.1 for details and an example where this is particularly important. A similar approach might prove to be useful in other network environments.

5.2. Traffic Protection

It is possible that due to an implementation error or misconfiguration, a QB microflow could end up being mismarked as NQB, or vice versa. It is also possible that a malicious actor could introduce a QB microflow marked as NQB with the intention of causing disruptions. In the case of a low data rate microflow that isn't marked as NQB and therefore 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 microflow that is mismarked as NQB would cause queuing delays and/or loss for all the other microflows that are sharing the NQB queue.

To prevent this situation from harming the performance of the microflows that are in compliance with the requirements in Section 4.1, network elements that support the NQB PHB SHOULD support a "traffic protection" function that can identify microflows that are inconsistent with the sender requirements in Section 4.1, and either re-mark those microflows/packets as Default and reclassify them to the QB queue or discard the offending traffic. Such a function SHOULD base its decisions upon the behavior of each microflow rather than on application-layer constructs (such as the port number used by the application or the source/destination IP address).

This specification does not mandate a particular algorithm for traffic protection. This is intentional, since the specifics of traffic protection could need to be different in different network equipment and in different network contexts. Instead this specification provides guidelines and some examples of traffic protection algorithms which could be employed.

One potential implementation of such a traffic protection algorithm is a per-microflow rate policer, designed to identify microflows that exceed the bound provided in Section 4.1, where the value R is set to 1 percent of the egress link capacity available for NQB traffic. An alternative is to use a traffic protection algorithm that bases its decisions on the detection of actual queuing (i.e. by monitoring the queuing delay experienced by packets in the NQB queue) in correlation with the arrival of packets for each microflow. One example traffic protection algorithm along these lines can be found in [I-D.briscoe-docsis-q-protection]. This algorithm maintains per-microflow state for up to 32 simultaneous "queue-building" microflows, and shared state for any additional microflows in excess of that number.

In the case of a traffic protection algorithm that re-marks and reclassifies offending traffic, different levels of hysteresis could be considered. For example, the re-mark/reclassify decision could be made on a packet-by-packet basis, which could result in significant out-of-order delivery for offending microflows as some portion of the microflow's packets remain in the NQB queue and some are re-marked and reclassified to the Default queue. Alternatively, a traffic protection function could employ a certain level of hysteresis to prevent borderline microflows from being reclassified capriciously, thus causing less potential for out-of-order delivery. As a third option, the decision could be made to take action on all the future packets of the microflow, though sufficient logic would be needed to ensure that a future microflow (e.g. with the same 5-tuple) isn't misidentified as the current offending microflow.

In the case of a traffic protection algorithm that discards offending traffic, similar levels of hysteresis could be considered. In this case, it is RECOMMENDED that the decision thresholds be set higher than in the case of designs that use re-mark/reclassify, since the degradation of communications caused by packet discard are likely to be greater than the degradation caused by out-of-order delivery.

The traffic protection function described here requires that the network element maintain microflow state. The traffic protection function MUST be designed such that the node implementing the NQB PHB does not fail (e.g. crash) in the case that the microflow state is exhausted.

There are some situations where traffic protection is potentially not necessary. For example, a network element designed for use in controlled environments (e.g., enterprise LAN) where a network administrator is expected to manage the usage of DSCPs. Additionally, some networks might prefer to police the application of the NQB DSCP at the ingress edge, so that per-hop traffic protection is not needed.

5.3. Impact on Higher Layer Protocols

Network elements that support the NQB PHB and that support traffic protection as discussed in the previous section introduce the possibility that microflows classified into the NQB queue could experience out-of-order delivery or packet loss if their behavior is not consistent with the NQB sender requirements. Out-of-order delivery could be particularly likely if the traffic protection algorithm makes decisions on a packet-by-packet basis. In this scenario, a microflow that is (mis)marked as NQB and that causes a queue to form in this bottleneck link could see some of its packets forwarded by the NQB queue, and some of them either discarded or redirected to the QB queue. In the case of redirection, depending on the queuing latency and scheduling within the network element, this could result in packets being delivered out of order. As a result, the use of the NQB DSCP by a higher layer protocol carries some risk that an increased amount of out-of-order delivery or packet loss will be experienced. This characteristic provides one disincentive for incorrectly setting the NQB DSCP on traffic that doesn't comply with the NQB sender requirements.

6. Configuration and Management

As required in Section 5, nodes supporting the NQB PHB provide for the configuration of classifiers that can be used to differentiate between QB and NQB traffic of equivalent importance. The default for such classifiers is recommended to be the assigned NQB DSCP (to identify NQB traffic) and the Default (0) DSCP (to identify QB traffic).

Additionally, Section 4.2 contains configuration recommendations for nodes that do not support the NQB PHB, and Section 4.4.1 contains configuration recommentations for networks that interconnect with non-DS-capable domains.

6.1. Guidance for Lower-Rate Links

The NQB sender requirements in Section 4.1 place responsibility in the hands of the application developer to determine the likelihood that the application's sending behavior could result in a queue forming along the path. These requirements rely on application developers having a reasonable sense for the network context in which their application is to be deployed. Even so, there will undoubtedly be networks that contain links having a data rate that is below the lower end of what is considered "typical", and some of these links could even be below the instantaneous sending rate of some NQB-marked applications.

To limit the consequences of this scenario, operators of networks with lower rate links SHOULD consider utilizing a traffic protection function on those links that is more tolerant of burstiness (i.e., a temporary queue). This will have the effect of allowing a larger set of NQB-marked microflows to remain in the NQB queue, but will come at the expense of a greater potential for latency variation. In implementations that support [I-D.briscoe-docsis-q-protection], the burst tolerance can be configured via the CRITICALqLSCORE_us input parameter.

Alternatively, operators of networks with lower rate links MAY choose to disable NQB support (and thus aggregate traffic marked with the NQB DSCP with Default traffic) on these lower rate links. For links that have a data rate that is less than ten percent of "typical" path rates, it is RECOMMENDED that the NQB PHB be disabled and for traffic marked with the NQB DSCP to thus be carried using the Default PHB.

7. Mapping NQB to standards of other SDOs

This section provide recommendations for the support of the NQB PHB in certain use cases. This section is not exhaustive.

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 traffic marked with the NQB DSCP. The NQB queue MUST be configured to share the service's rate shaped bandwidth with the queue for QB traffic.

7.2. Mobile Networks

Historically, 3GPP mobile networks have utilized "bearers" to encapsulate each user's user plane traffic through the radio and core networks. A "dedicated bearer" can be allocated a Quality of Service (QoS) to apply any prioritisation to its microflows at queues and radio schedulers. Typically, an LTE operator provides a dedicated bearer for IMS VoLTE (Voice over LTE) traffic, which is prioritized in order to meet regulatory obligations for call completion rates; and a "best effort" default bearer, for Internet traffic. The "best effort" bearer provides no guarantees, and hence its buffering characteristics are not compatible with low-latency traffic. The 5G radio and core systems offer more flexibility over bearer allocation, meaning bearers can be allocated per traffic type (e.g., loss-tolerant, low-latency etc.) and hence support more suitable treatment of Internet real-time microflows.

To support the NQB PHB, the mobile network SHOULD be configured to give User Equipment 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 NOT be routed through the dedicated low-latency EPS bearer.

7.3. Wi-Fi Networks

Wi-Fi networking equipment compliant with 802.11e/n/ac/ax [IEEE802-11] generally supports either four or eight transmit queues and four sets of associated Enhanced Multimedia Distributed Control Access (EDCA) parameters (corresponding to the four Wi-Fi Multimedia (WMM) Access Categories) that are used to enable differentiated media access characteristics. As discussed in [RFC8325], it has been a common practice for Wi-Fi implementations to 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. [RFC8325] also provides an alternative mapping that more closely aligns with the DSCP recommendations provided by the IETF. In the case of some managed Wi-Fi gear, this mapping can be controlled by the network operator, e.g., via TR-369 [TR-369].

In addition to the requirements provided in other sections of this document, to support the NQB PHB, Wi-Fi equipment (including equipment compliant with [RFC8325]) SHOULD map the NQB DSCP 45 (decimal) into a separate queue in the same Access Category as the queue that carries Default traffic (i.e. the Best Effort Access Category). It is RECOMMENDED that Wi-Fi equipment provide a separate queue in UP 0, and map the NQB DSCP 45 (decimal) to that queue. If a separate queue in UP 0 cannot be provided (due to hardware limitations, etc.) a Wi-Fi device MAY map the NQB DSCP 45 (decimal) to UP 3.

7.3.1. Interoperability with Existing Wi-Fi Networks

While some existing Wi-Fi equipment might be capable (in some cases via firmware update) of supporting the NQB PHB requirements, many currently deployed devices cannot be configured in this way. As a result, the remainder of this section discusses interoperability with these existing Wi-Fi networks, as opposed to PHB compliance.

Since this equipment is widely deployed, and the Wi-Fi link is commonly a bottleneck link, the performance of traffic marked with the NQB DSCP across such links could have a significant impact on the viability and adoption of the NQB DSCP and PHB. Depending on the DSCP used to mark NQB traffic, existing Wi-Fi equipment that uses the default mapping of DSCPs to Access Categories and the default EDCA parameters will support either the NQB PHB requirement for separate queuing of NQB traffic, or the recommendation to treat NQB traffic with priority equal to Default traffic, but not both.

The DSCP value 45 (decimal) is recommended for NQB. This maps NQB to UP_5 using the default mapping, which is in the "Video" Access Category. While this choice of DSCP enables these Wi-Fi systems to support the NQB PHB requirement for separate queuing, existing Wi-Fi devices generally utilize EDCA parameters that result in statistical prioritization of the "Video" Access Category above the "Best Effort" Access Category. In addition this equipment does not support the remaining NQB PHB recommendations in Section 5. The rationale for the choice of DSCP 45 (decimal) as well as its ramifications, and remedies for its limitations are discussed further below.

The choice of separated queuing rather than equal priority in existing Wi-Fi networks was motivated by the following:

- Separate queuing is necessary in order to provide a benefit for traffic marked with the NQB DSCP.

- Wi-Fi gear typically has hardware support (albeit generally not exposed for user control) for adjusting the EDCA parameters in order to meet the equal priority recommendation. This is discussed further below.

- Traffic that is compliant with the NQB sender requirements Section 4.1 is unlikely to cause more degradation to lower priority Access Categories than the existing recommended Video

Access Category traffic types: Broadcast Video, Multimedia Streaming, Multimedia Conferencing from [RFC8325], and AudioVideo, ExcellentEffort from [QOS_TRAFFIC_TYPE].

- Application instances on Wi-Fi client devices are already free to choose any Access Category that they wish, regardless of their sending behavior, without any policing of usage. So, the choice of using DSCP 45 (decimal) for NQB creates no new avenues for non-NQB-compliant client applications to exploit the prioritization function in Wi-Fi.

- Several existing client applications that are compatible with the NQB sender requirements already select the Video Access Category, and thus would not see a degradation in performance by transitioning to the NQB DSCP, regardless of whether the network supported the PHB.

- For application traffic that originates outside of the Wi-Fi network, and thus is transmitted by the Access Point, opportunities exist in the network components upstream of the Wi-Fi Access Point to police the usage of the NQB DSCP and potentially re-mark traffic that is considered non-compliant, as is recommended in Section 4.4.1. A residential ISP that re-marks the Diffserv field to zero, bleaches all DSCPs and hence would not be impacted by the introduction of traffic marked as NQB. Furthermore, any change to this practice ought to be done alongside the implementation of those recommendations in the current document.

The choice of Video Access Category rather than the Voice Access Category was motivated by the desire to minimize the potential for degradation of Best Effort Access Category traffic. The choice of Video Access Category rather than the Background Access Category was motivated by the much greater potential of degradation to NQB traffic that would be caused by the vast majority of traffic in most Wi-Fi networks, which utilizes the Best Effort Access Category.

If left unchanged, the prioritization of traffic marked with the NQB DSCP via the Video Access Category (particularly in the case of traffic originating outside of the Wi-Fi network as mentioned above) could erode the principle of alignment of incentives discussed in Section 5. In order to preserve the incentives principle for NQB, Wi-Fi systems SHOULD be configured such that the EDCA parameters for the Video Access Category match those of the Best Effort Access Category. These changes can be deployed in managed Wi-Fi systems or those deployed by an ISP and are intended for situations when the vast majority of traffic that would use AC_VI is NQB. In other situations (e.g., consumer-grade Wi-Fi gear deployed by an ISP's customer) this configuration might not be possible, and the requirements and recommendations in Section 4.4.1 would apply.

Similarly, systems that utilize [RFC8325] but that are unable to fully support the PHB requirements, SHOULD map the recommended NQB DSCP 45 (decimal) (or the locally determined alternative) to UP_5 in the "Video" Access Category.

8. Acknowledgements

Thanks to Gorry Fairhurst, Diego Lopez, Stuart Cheshire, Brian Carpenter, Bob Briscoe, Greg Skinner, Toke Hoeiland-Joergensen, Luca Muscariello, David Black, Sebastian Moeller, Jerome Henry, Steven Blake, Jonathan Morton, Roland Bless, Kevin Smith, Martin Dolly and Kyle Rose for their review comments. Thanks also to Gorry Fairhurst and Ana Custura for their input on selection of appropriate DSCPs.

9. IANA Considerations

This document requests that IANA assign the Differentiated Services Field Codepoint (DSCP) 45 ('0b101101', 0x2D) from the "Differentiated Services Field Codepoints (DSCP)" registry (https://www.iana.org/assignments/dscp-registry/) ("DSCP Pool 3 Codepoints", Codepoint Space xxxx01, Standards Action) as the RECOMMENDED codepoint for Non-Queue-Building behavior.

IANA should update this registry as follows:

*Name: NQB

*Value (Binary): 101101

*Value (Decimal): 45

*Reference: this document

10. Implementation Status

Note to RFC Editor: This section should be removed prior to publication

The NQB PHB is implemented in equipment compliant with the current DOCSIS 3.1 specification, published by CableLabs at: CableLabs Specifications Search.

CableLabs maintains a list of production cable modem devices that are Certified as being compliant to the DOCSIS Specifications, this list is available at https://www.cablelabs.com/wp-content/uploads/2013/10/cert_qual.xlsx. DOCSIS 3.1 modems certified in CW 134 or greater implement the NQB PHB. This includes products from Arcadyan

Technology Corporation, Arris, AVM, Castlenet, Commscope, Hitron, Motorola, Netgear, Sagemcom and Vantiva. There are additional production implementations that have not been Certified as compliant to the specification, but which have been tested in non-public Interoperability Events. These implementations are all proprietary, not available as open source.

11. Security Considerations

When the NQB PHB is fully supported in bottleneck links, there is no incentive for a Queue-Building application to mismark its packets as NQB (or vice versa). If a Queue-Building microflow were to mismark its packets as NQB, it would be unlikely to receive a benefit by doing so, and it would usually experience a degradation. The nature of the degradation would depend on the specifics of the PHB implementation (and on the presence or absence of a traffic protection function), but could include excessive packet loss, excessive latency variation and/or excessive out-of-order delivery. If a Non-Queue-Building microflow was to fail to mark its packets as NQB, it could suffer the latency and loss typical of sharing a queue with capacity seeking traffic.

To preserve low latency performance for NQB traffic, networks that support the NQB PHB will need to ensure that mechanisms are in place to prevent malicious traffic marked with the NQB DSCP from causing excessive queue delays. Section 5.2 recommends the implementation of a traffic protection mechanism to achieve this goal but recognizes that other options might be more desirable in certain situations. The recommendations on traffic protection mechanisms in this document presume that some type of "flow" state be maintained in order to differentiate between microflows that are causing queuing delay and those that aren't. Since this flow state is likely finite, this opens up the possibility of flow-state exhaustion attacks. While this document requires that traffic protection mechanisms be designed with this possibility in mind, the outcomes of flow-state exhaustion would depend on the implementation.

Notwithstanding the above, the choice of DSCP for NQB does allow existing Wi-Fi networks to readily (and by default) support some of the PHB requirements, but without a traffic protection function, and (when left in the default state) by giving NQB traffic higher priority than QB traffic. This is not considered to be a compliant implementation of the PHB. These existing Wi-Fi networks currently provide priority to half of the DSCP space, including the NQB DSCP. While the NQB DSCP value could be abused to gain priority on such links, the potential presence of traffic protection functions in other hops along the path (which likely act on the NQB DSCP value alone) would make it less attractive for such abuse than any of the other 31 DSCP values that are provided priority.

This draft discusses the potential use of the NQB DSCP and NQB PHB in network technologies that are standardized in other SDOs. The details of any security considerations that relate to deployment and operation of NQB in these network technologies are not discussed here.

NQB uses the Diffserv field. The design of Diffserv does not include integrity protection for the DSCP, and thus it is possible for the DSCP to be changed by an on-path attacker. The NQB PHB and associated DSCP don't change this. While re-marking DSCPs is permitted for various reasons (some are discussed in this document, others can be found in [RFC2474] and [RFC2475]), if done maliciously, this might negatively affect the QoS of the tampered microflow.

12. References

12.1. Normative References

[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>.

[RFC2983] Black, D., "Differentiated Services and Tunnels", RFC 2983, DOI 10.17487/RFC2983, October 2000, <https://www.rfc-editor.org/info/rfc2983>.

[RFC8085] Eggert, L., Fairhurst, G., and G. Shepherd, "UDP Usage Guidelines", BCP 145, RFC 8085, DOI 10.17487/RFC8085, March 2017, <https://www.rfc-editor.org/info/rfc8085>.

[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>.

12.2. 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.

[I-D.briscoe-docsis-q-protection] Briscoe, B. and G. White, "The DOCSIS(r) Queue Protection Algorithm to Preserve Low Latency", Work in Progress, Internet-Draft, draft-briscoe-docsis-q-protection-06, 13 May 2022, <https://datatracker.ietf.org/doc/html/draft-briscoe-docsis-q-protection-06>.

[IEEE802-11] IEEE-SA, "IEEE 802.11-2020", IEEE 802, December 2020, <https://standards.ieee.org/standard/802_11-2020.html>.

[QOS_TRAFFIC_TYPE] Microsoft, Corporation, "QOS_TRAFFIC_TYPE enumeration", 2022, <https://learn.microsoft.com/en-us/windows/win32/api/qos2/ne-qos2-qos_traffic_type>.

[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>.

[RFC2598] Jacobson, V., Nichols, K., and K. Poduri, "An Expedited Forwarding PHB", RFC 2598, DOI 10.17487/RFC2598, June 1999, <https://www.rfc-editor.org/info/rfc2598>.

[RFC2914] Floyd, S., "Congestion Control Principles", BCP 41, RFC 2914, DOI 10.17487/RFC2914, September 2000, <https://www.rfc-editor.org/info/rfc2914>.

[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>.

[RFC3551] Schulzrinne, H. and S. Casner, "RTP Profile for Audio and Video Conferences with Minimal Control", STD 65, RFC 3551, DOI 10.17487/RFC3551, July 2003, <https://www.rfc-editor.org/info/rfc3551>.

[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>.

[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>.

[RFC7893] Stein, Y., Black, D., and B. Briscoe, "Pseudowire Congestion Considerations", RFC 7893, DOI 10.17487/RFC7893, June 2016, <https://www.rfc-editor.org/info/rfc7893>.

[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>.

[RFC8083] Perkins, C. and V. Singh, "Multimedia Congestion Control: Circuit Breakers for Unicast RTP Sessions", RFC 8083, DOI 10.17487/RFC8083, March 2017, <https://www.rfc-editor.org/info/rfc8083>.

[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>.

[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>.

[RFC8290] Hoeiland-Joergensen, T., McKenney, P., Taht, D., Gettys, J., and E. Dumazet, "The Flow Queue CoDel Packet Scheduler and Active Queue Management Algorithm", RFC

8290, DOI 10.17487/RFC8290, January 2018, <https://www.rfc-editor.org/info/rfc8290>.

[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>.

[RFC9330] Briscoe, B., Ed., De Schepper, K., Bagnulo, M., and G. White, "Low Latency, Low Loss, and Scalable Throughput (L4S) Internet Service: Architecture", RFC 9330, DOI 10.17487/RFC9330, January 2023, <https://www.rfc-editor.org/info/rfc9330>.

[RFC9331] De Schepper, K. and B. Briscoe, Ed., "The Explicit Congestion Notification (ECN) Protocol for Low Latency, Low Loss, and Scalable Throughput (L4S)", RFC 9331, DOI 10.17487/RFC9331, January 2023, <https://www.rfc-editor.org/info/rfc9331>.

[RFC9332] De Schepper, K., Briscoe, B., Ed., and G. White, "Dual-Queue Coupled Active Queue Management (AQM) for Low Latency, Low Loss, and Scalable Throughput (L4S)", RFC 9332, DOI 10.17487/RFC9332, January 2023, <https://www.rfc-editor.org/info/rfc9332>.

[RFC9435] Custura, A., Fairhurst, G., and R. Secchi, "Considerations for Assigning a New Recommended Differentiated Services Code Point (DSCP)", RFC 9435, DOI 10.17487/RFC9435, July 2023, <https://www.rfc-editor.org/info/rfc9435>.

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

[TR-369] Broadband Forum, "The User Services Platform", January 2022, <https://usp.technology/specification/index.html>.

Appendix A. DSCP Re-marking Policies

Some network operators typically bleach (zero out) the Diffserv field on ingress into their network [RFC9435][Custura][Barik], and in some cases apply their own DSCP for internal usage. Bleaching the NQB DSCP is not expected to cause harm to Default traffic, but it will severely limit the ability to provide NQB treatment end-to-end. Reports on existing deployments of DSCP manipulation [Custura] [Barik] categorize the re-marking behaviors into the following six policies: bleach all traffic (set DSCP to zero), set the top three bits (the former Precedence bits) on all traffic to 0b000, 0b001, or 0b010, set the low three bits on all traffic to 0b000, or re-mark all traffic to a particular (non-zero) DSCP value.

Regarding the DSCP value 45 (decimal), there were no observations of DSCP manipulation reported in which traffic was marked 45 (decimal) by any of these policies. Thus it appears that these re-marking policies would be unlikely to result in QB traffic being marked as NQB (45). In terms of the fate of traffic marked with the NQB DSCP that is subjected to one of these policies, the result would be that traffic marked with the NQB DSCP would be indistinguishable from some subset (possibly all) of other traffic. In the policies where all traffic is re-marked using the same (zero or non-zero) DSCP, the ability for a subsequent network hop to differentiate NQB traffic via DSCP would clearly be lost entirely.

In the policies where the top three bits are overwritten (see Section 4.2 of [RFC9435]), the NQB DSCP (45) would receive the same marking as would the currently unassigned Pool 3 DSCPs 5,13,21,29,37,53,61, with all of these DSCPs getting re-marked to DSCP = 5, 13 or 21 (depending on the overwrite value used). Since none of the DSCPs in the preceding lists are currently assigned by IANA, and they all are reserved for Standards Action, it is believed that they are not widely used currently, but this could vary based on local-usage, and could change in the future. If networks in which this sort of re-marking occurs (or networks downstream) classify the resulting DSCP (i.e. 5, 13, or 21) to the NQB PHB, or re-mark such traffic as 45 (decimal), they risk treating as NQB other traffic, which was not originally marked as NQB. In addition, as described in Section 6 of [RFC9435] future assignments of these 0bxxx101 DSCPs would need to be made with consideration of the potential that they all are treated as NQB in some networks.

For the policy in which the low three bits are set to 0b000, the NQB (45) value would be re-marked to CS5 and would be indistinguishable from CS5, VA, EF (and the unassigned DSCPs 41, 42, 43). Traffic marked using the existing standardized DSCPs in this list are likely to share the same general properties as NQB traffic (non-capacity-seeking, very low data rate or relatively low and consistent data rate). Similarly, any future recommended usage for DSCPs 41, 42, 43 would likely be somewhat compatible with NQB treatment, assuming that IP Precedence compatibility (see Section 1.5.4 of [RFC4594]) is maintained in the future. Here there might be an opportunity for a node to provide the NQB PHB or the CS5 PHB to CS5-marked traffic and retain some of the benefits of NQB marking. This could be another motivation to (as discussed in Section 4.3) classify CS5-marked traffic into NQB queue.

Appendix B. Comparison to Expedited Forwarding

The Expedited Forwarding definition [RFC3246] provides the following text to describe the EF PHB forwarding behavior: "This specification defines a PHB in which EF packets are guaranteed to receive service at or above a configured rate" and "the rate at which EF traffic is served at a given output interface should be at least the configured rate R, over a suitably defined interval, independent of the offered load of non-EF traffic to that interface." Notably, this description is true of any class of traffic that is configured with a guaranteed minimum rate, including the Default PHB if configured per the guidelines in Section 1.5.1 of [RFC4594]. [RFC3246] goes on to formalize the definition of EF by requiring that an EF node be characterizable in terms of the fidelity with which it is able to provide a guaranteed rate.

While the NQB PHB is not required to be configured with a guaranteed minimum rate, [RFC2474] and [RFC4594] recommend assigning some minimum resources for the Default PHB, in particular some dedicated bandwidth. If such a guaranteed minimum rate is configured for the Default PHB, it is recommended (Section 5) that NQB traffic share and be given equal access to that rate. In such cases, the NQB PHB effectively receives a rate guarantee of 50% of the rate guaranteed to the combined NQB/Default PHBs, and so technically complies with the PHB forwarding behavior defined for EF.

However, EF is intended to be a managed service, and requires that traffic be policed such that the arriving rate of traffic into the EF PHB doesn't exceed the guaranteed forwarding rate configured for the PHB, thereby ensuring that low latency and low latency variation are provided. NQB is intended as a best effort service, and hence the aggregate of traffic arriving to the NQB PHB queue could exceed the forwarding rate available to the PHB. Section 5.2 discusses the recommended mechanism for handling excess traffic in NQB. While EF relies on rate policing and dropping of excess traffic, this is only one option for NQB. NQB alternatively recommends that the implementation re-mark and forward excess traffic using the Default PHB, rather than dropping it. Further, NQB recommends a microflow-based mechanism to limit the performance impact of excess traffic to those microflows causing potential congestion of the NQB queue, whereas EF ignores microflow properties. Note that under congestion, low loss for NQB conformant flows is only ensured if such a mechanism is operational. Note also that this mechanism for NQB operates at the available forwarding rate for the PHB (which could vary based on other traffic load) as opposed to a configured guaranteed rate, as in EF.

The lack of a requirement of a guaranteed minimum rate, and the lack of a requirement to police incoming traffic to such a rate, makes the NQB PHB suitable for implementation in networks where link capacity is not or cannot be guaranteed.

There are additional distinctions between EF and NQB arising from the intended usage as described in [RFC4594] and the actual usage in practice in the Internet. In Section 1.5.3 of [RFC4594], EF is described as generally being used to carry voice or data that requires "wire like" behavior through the network. The NQB PHB similarly is useful to carry application traffic requiring wire like performance, characterized by low packet delay and delay variation, but places a pre-condition that each microflow be relatively low data rate and sent in a smooth (non-bursty) manner. In actual practice, EF traffic is oftentimes prioritized over Default traffic. This contrasts with NQB traffic which is to be treated with the same forwarding priority as Default (and sometimes aggregated with Default).

Appendix C. Alternate Diffserv Code Points

In networks where another (e.g., a local-use) DSCP is designated for NQB traffic, or where specialized PHBs are available that can meet specific application requirements (e.g., a guaranteed-latency path for voice traffic), it could be preferred to use another DSCP. In end systems where the choice of using DSCP 45 (decimal) is not available to the application, the CS5 DSCP (40 decimal) could be used as a fallback. See Section 4.3 for rationale as to why this choice could be fruitful.

Authors' Addresses

Greg White
CableLabs

Email: g.white@cablelabs.com

Thomas Fossati
Linaro

Email: thomas.fossati@linaro.org

Rüdiger Geib
Deutsche Telekom

Email: Ruediger.Geib@telekom.de