Internet Engineering Task Force   R. Bless
Internet-Draft   Karlsruhe Institute of Technology (KIT)
Obsoletes: 3662 (if approved)   February 6, 2017
Updates: 4594 (if approved)
Intended status: Standards Track
Expires: August 10, 2017

                   A Lower Effort Per-Hop Behavior (LE PHB)
                         draft-ietf-tsvwg-le-phb-01

Abstract

   This document specifies properties and characteristics of a Lower
   Effort (LE) per-hop behavior (PHB). The primary objective of this LE
   PHB is to protect best-effort (BE) traffic (packets forwarded with
   the default PHB) from LE traffic in congestion situations, i.e., when
   resources become scarce, best-effort traffic has precedence over LE
   traffic and may preempt it. There are numerous uses for this PHB,
   e.g., for background traffic of low precedence, such as bulk data
   transfers with low priority in time, non time-critical backups,
   larger software updates, web search engines while gathering
   information from web servers and so on. This document recommends a
   standard DSCP value for the LE PHB.

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 http://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 August 10, 2017.

Copyright Notice

Table of Contents

1. Introduction

   This document defines a Differentiated Services per-hop behavior
   [RFC2474] called "Lower Effort" (LE) which is intended for traffic of
   sufficiently low urgency, in which all other traffic takes precedence
   over LE traffic in consumption of network link bandwidth. Low
   urgency traffic has got a low priority in time, which does not
   necessarily imply that it is generally of minor importance. From
   this viewpoint, it can be considered as a network equivalent to a
   background priority for processes in an operating system. There may
   or may not be memory (buffer) resources allocated for this type of
   traffic.

   Some networks carry traffic for which delivery is considered
   optional; that is, packets of this type of traffic ought to consume
   network resources only when no other traffic is present.

   Alternatively, the effect of this type of traffic on all other
   network traffic is strictly limited. This is distinct from "best-
   effort" (BE) traffic since the network makes no commitment to deliver
   LE packets. In contrast, BE traffic receives an implied "good faith"
   commitment of at least some available network resources. This
   document proposes a Lower Effort Differentiated Services per-hop
   behavior (LE PHB) for handling this "optional" traffic in a
   differentiated services node.

1.1. Applicability

   A Lower Effort PHB is applicable for most elastic applications that
   otherwise use best-effort delivery. More specifically, it is
   suitable for traffic and services accepting strongly varying
   throughput for their data flows, especially with respect to periods
   of very low throughput or even starvation (i.e., long interruptions
   due to excessive packet loss). Therefore, an application sending an
   LE marked flow must be able to tolerate short or (even very) long
   interruptions due to the presence of severe congestion conditions
   during the transmission of the flow. Thus, there should be an
   expectation that packets of the LE PHB may be excessively delayed or
   dropped when any other traffic is present. The LE PHB is suitable
   for sending traffic of low urgency across a Differentiated Services
   (DS) domain or DS region.

   LE traffic SHOULD be congestion controlled. Since LE traffic may be
   starved completely for a longer period of time, transport protocols
   or applications (and their related congestion control mechanisms)
   SHOULD be able to detect and react to such a situation and should
   resume the transfer as soon as possible. Congestion control is not
   only useful to let the flows within the LE behavior aggregate adapt
   to the available bandwidth that may be highly fluctuating, but also
   in case that LE traffic is mapped to the default PHB (e.g., due to
   DSCP bleaching).

   Use of the LE PHB might assist a network operator in moving certain
   kinds of traffic or users to off-peak times. Alternatively, or in
   addition, packets can be designated for the LE PHB when the goal is
   to protect all other packet traffic from competition with the LE
   aggregate while not completely banning LE traffic from the network.
   An LE PHB should not be used for a customer's "normal internet"
   traffic nor should packets be "downgraded" to the LE PHB used as a
   substitute for dropping packets that ought simply to be dropped as
   unauthorized. The LE PHB is expected to have applicability in
   networks that have at least some unused capacity at some times of
   day.

The LE PHB allows networks to protect themselves from selected types
of traffic rather than giving a selected traffic aggregate
preferential treatment. Moreover, the LE PHB may also exploit all
unused resources from other PHBs.

There is no intrinsic reason to limit the applicability of the LE PHB
to any particular application or type of traffic. It is intended as
an additional tool for administrators in engineering networks. For
instance, it can be used for filling up protection capacity of
transmission links which is otherwise unused. Some network providers
keep link utilization below 50% in order to being able carrying all
traffic without loss in case of rerouting due to a link failure. LE
marked traffic can utilize the normally unused capacity and will be
preempted automatically in case of link failure when 100% of the link
capacity is required for all other traffic. Ideally, applications
mark their packets as LE traffic, since they know the urgency of
flows.

Example uses for the LE PHB comprise:

o For traffic caused by world-wide web search engines while they
   gather information from web servers.

o For software updates or dissemination of new releases of operating
   systems.

o For backup traffic or non-time critical synchronization or
   mirroring traffic.

o For content distribution transfers between caches.

o For Netnews and other "bulk mail" of the Internet.

o For "downgraded" traffic from some other PHB when this does not
   violate the operational objectives of the other PHB or the overall
   network. LE should not be used for the general case of downgraded
   traffic, but may be used by design, e.g., to protect an internal
   network from untrusted external traffic sources. In this case
   there is no way for attackers to preempt internal (non LE) traffic
   by flooding. Another use case is mentioned in [RFC3754]: non-
   admitted multicast traffic.

1.2. Deployment Considerations

Internet-wide deployment of the LE PHB is eased by the following
properties:

   o No harm to other traffic: since the LE PHB has got the lowest
      forwarding priority it does not consume resources from other PHBs.
      Deployment across different provider domains causes no trust
      issues or attack vectors to existing traffic.

   o No parameters or configuration: the LE PHB requires no parameters
      and no configuration of traffic profiles and so on.

   o No traffic conditioning mechanisms: the LE PHB requires no traffic
      meters, droppers, or shapers.

   DS domains that cannot or do not want to support the LE PHB SHOULD
   NOT drop LE marked packets, but rather map them to the default PHB
   and keep the LE DSCP. See also Section 5 for further discussion of
   forwarding LE traffic with the default PHB instead.

1.3. 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 [RFC2119].

2. PHB Description

   The LE PHB is defined in relation to the default PHB (best-effort).
   A packet forwarded with the LE PHB SHOULD have lower precedence than
   packets forwarded with the default PHB, i.e., in case of congestion,
   LE marked traffic SHOULD be dropped prior to dropping any default PHB
   traffic. Ideally, LE packets SHOULD be forwarded only if no best-
   effort packet is waiting for its transmission.

   A straightforward implementation could be a simple priority scheduler
   serving the default PHB queue with higher priority than the lower-
   effort PHB queue. Alternative implementations may use scheduling
   algorithms that assign a very small weight to the LE class. This,
   however, may sometimes cause better service for LE packets compared
   to BE packets in cases when the BE share is fully utilized and the LE
   share not.

3. Traffic Conditioning Actions

   As for most other PHBs an initial classification and marking would
   usually be performed at the first DS boundary node. If possible,
   packets SHOULD be pre-marked in DS-aware end systems by applications
   due to their specific knowledge about the particular precedence of
   packets. There is no incentive for DS domains to distrust this
   initial marking, because letting LE traffic enter a DS domain causes
   no harm. In the worst case it evokes the same effect as it would

have been marked with the default PHB, i.e., as best-effort traffic.
Usually, the amount of LE traffic is implicitly limited by queueing
mechanisms and related discard actions of the PHB. Thus, any
policing such as limiting the rate of LE traffic is not necessary at
the DS boundary.

Non-LE traffic (e.g., BE traffic) SHOULD not be remarked to LE on a
regular basis without consent or knowledge of the user.

4. Recommended DS Codepoint

The RECOMMENDED codepoint for the LE PHB is '000010'.

Earlier specifications [RFC4594] recommended to use CS1 as codepoint
(as mentioned in [RFC3662]). This is problematic since it may cause
a priority inversion in DiffServ domains that treat CS1 as originally
proposed in [RFC2474], resulting in forwarding LE packets with higher
precedence than BE packets. Existing implementations SHOULD
therefore use the unambiguous LE codepoint '000010' whenever
possible.

5. Remarking to other DSCPs/PHBs

"DSCP bleaching", i.e., setting the DSCP to '000000' (default PHB) is
NOT RECOMMENDED for this PHB. This may cause effects that are in
contrast to the original intent in protecting BE traffic from LE
traffic. In case DS domains do not support the LE PHB, they SHOULD
treat LE marked packets with the default PHB instead (by mapping the
LE DSCP to the default PHB), but they SHOULD do so without remarking
to DSCP '000000'. The reason for this is that later traversed DS
domains may then have still the possibility to treat such packets
according the LE PHB. However, operators of DS domains that forward
LE traffic within the BE aggregate should be aware of the
implications, i.e., induced congestion situations and quality-of-
service degradation of the original BE traffic. In this case, the LE
property of not harming other traffic is no longer fulfilled. In
order to limit the impact in such cases, traffic policing of the LE
aggregate may be used.

In case LE marked packets are effectively carried within the default
PHB (i.e., forwarded as best-effort traffic) they get a better
forwarding treatment than expected. For some applications and
services, it is favorable if the transmission is finished earlier
than expected. However, in some cases it may be against the original
intention of the LE PHB user to strictly send the traffic only if
otherwise unused resources are available, i.e., LE traffic may
compete with BE traffic for the same resources and thus adversely
affect the original BE aggregate. One possible solution for a clear

distinction in such cases would be to use two different codepoints,
"LE-min = LE, better treatment allowed", "LE-strict = LE, better
treatment NOT allowed". However, since DSCPs are a scarce resource,
applications that want to ensure the lower precedence compared to BE
traffic SHOULD use additionally a corresponding Lower-than-Best-
Effort transport protocol [RFC6297], e.g., LEDBAT [RFC6817].

A DS domain that still uses DSCP CS1 for marking LE traffic
(including Low Priority-Data as defined in [RFC4594] or the old
definition in [RFC3662]) MUST remark traffic to the LE DSCP '000010'
at the egress to the next DS domain. This increases the probability
that the DSCP is preserved end-to-end, whereas a CS1 marked packet
may be remarked by the default DSCP if the next domain is applying
DiffServ-intercon [I-D.ietf-tsvwg-diffserv-intercon].

6. Changes to RFC 4594

[RFC4594] recommended to use CS1 as codepoint in section 4.10,
whereas CS1 was defined in [RFC2474] to have a higher precedence than
CS0, i.e., the default PHB. Consequently, DiffServ domains
implementing CS1 according to [RFC2474] will cause a priority
inversion for LE packets that contradicts with the original purpose
of LE. Therefore, every occurrence of the CS1 DSCP is replaced by
the LE DSCP.

Changes:

o The Low-Priority Data row in Figure 3 is updated as follows:

```
 |---------------+---------+-------------+-------------------------|
 | Low-Priority  |   LE    |   000010    | Any flow that has no BW |
 |     Data      |         |             | assurance               |
  -----------------------------------------------------------------
```

o The Low-Priority Data row in Figure 4 is updated as follows:

```
 |---------------+------+------------------+---------+--------+----|
 | Low-Priority  | LE   | Not applicable   | RFCXXXX |  Rate  | Yes|
 |     Data      |      |                  |         |        |    |
  -----------------------------------------------------------------
```

o Section 4.10: The RECOMMENDED DSCP marking is LE (Lower Effort).

o [RFC4594] recommended to remark Low-Priority Data to DSCP '000001'
   inside a DS domain that uses IP precedence marking. By using the
   herein defined LE DSCP such remarking is not necessary, so even if
   Low-Priority Data is unsupported (i.e., mapped to the default PHB)

the LE DSCP should be kept across the domain as RECOMMENDED in
Section 5.

7. IANA Considerations

This memo includes a request to assign a Differentiated Services
Field Codepoint (DSCP) '000010' from the Differentiated Services
Field Codepoints (DSCP) registry https://www.iana.org/assignments/
dscp-registry/dscp-registry.xml

8. Security Considerations

There are no specific security exposures for this PHB. Since it
defines a new class of low forwarding priority, other traffic may be
downgraded to this LE PHB in case it is remarked as LE traffic. See
the general security considerations in [RFC2474] and [RFC2475].

9. References

9.1. 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,
            <http://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,
            <http://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,
            <http://www.rfc-editor.org/info/rfc2475>.

9.2. Informative References

[draft-bless-diffserv-lbe-phb-00]
            Bless, R. and K. Wehrle, "A Lower Than Best-Effort Per-Hop
            Behavior", draft-bless-diffserv-lbe-phb-00 (work in
            progress), September 1999, <https://tools.ietf.org/html/
            draft-bless-diffserv-lbe-phb-00>.

[I-D.ietf-tsvwg-diffserv-intercon]
          Geib, R. and D. Black, "Diffserv-Interconnection classes
          and practice", draft-ietf-tsvwg-diffserv-intercon-14 (work
          in progress), December 2016.

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

[RFC3754] Bless, R. and K. Wehrle, "IP Multicast in Differentiated
          Services (DS) Networks", RFC 3754, DOI 10.17487/RFC3754,
          April 2004, <http://www.rfc-editor.org/info/rfc3754>.

[RFC4594] Babiarz, J., Chan, K., and F. Baker, "Configuration
          Guidelines for DiffServ Service Classes", RFC 4594,
          DOI 10.17487/RFC4594, August 2006,
          <http://www.rfc-editor.org/info/rfc4594>.

[RFC6297] Welzl, M. and D. Ros, "A Survey of Lower-than-Best-Effort
          Transport Protocols", RFC 6297, DOI 10.17487/RFC6297, June
          2011, <http://www.rfc-editor.org/info/rfc6297>.

[RFC6817] Shalunov, S., Hazel, G., Iyengar, J., and M. Kuehlewind,
          "Low Extra Delay Background Transport (LEDBAT)", RFC 6817,
          DOI 10.17487/RFC6817, December 2012,
          <http://www.rfc-editor.org/info/rfc6817>.

Appendix A. History of the LE PHB

A first version of this PHB was suggested by Roland Bless and Klaus
Wehrle in 1999 [draft-bless-diffserv-lbe-phb-00]. After some
discussion in the DiffServ Working Group Brian Carpenter and Kathie
Nichols proposed a bulk handling per-domain behavior and believed a
PHB was not necessary. Eventually, Lower Effort was specified as
per-domain behavior and finally became [RFC3662]. More detailed
information about its history can be found in Section 10 of
   [RFC3662].

Appendix B. Acknowledgments

Since text is borrowed from earlier Internet-Drafts and RFCs the co-
authors of previous specifications are acknowledged here: Kathie
Nichols and Klaus Wehrle. Ruediger Geib provided helpful comments
and suggestions.

Appendix C. Change History

   This section briefly lists changes between Internet-Draft versions
   for convenience.

   Changes in Version 01:

   o Now obsoletes RFC 3662.

   o Tried to be more precise in section 1.1 (Applicability) according
      to R. Geib's suggestions, so rephrased several paragraphs. Added
      text about congestion control

   o Change section 2 (PHB Description) according to R. Geib's
      suggestions.

   o Added RFC 2119 language to several sentences.

   o Detailed the description of remarking implications and
      recommendations in Section 5.

   o Added Section 6 to explicitly list changes with respect to RFC
      4594, because this document will update it.

Appendix D. Note to RFC Editor

   This section lists actions for the RFC editor during final
   formatting.

   o Please replace the occurrence of RFCXXXX in Section 6 with the
      assigned RFC number for this document.

   o Delete Appendix C.

   o Delete this section.

Author's Address

   Roland Bless
   Karlsruhe Institute of Technology (KIT)
   Kaiserstr. 12
   Karlsruhe 76131
   Germany

   Phone: +49 721 608 46413
   Email: roland.bless@kit.edu