TSVWG
Internet-Draft
Intended status: Informational
Expires: December 26, 2013

R. Geib, Ed.
Deutsche Telekom
June 24, 2013

DiffServ interconnection classes and practice
draft-geib-tsvwg-diffserv-intercon-03

Abstract

This document proposes a limited set of interconnection QoS PHBs and PHB groups. It further introduces some DiffServ deployment aspects. The proposals made here should be integrated into a revised version of RFC5127.

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 December 26, 2013.

Copyright Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (http://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

This draft proposes a DiffServ interconnection class and codepoint scheme. At least one party of an interconnection often is a network provider. Many network providers operate Aggregated DiffServ classes. This draft contains concepts and current practice relevant for a revised version of RFC5127 [RFC5127]. Its main purpose is to be considered as an input for the latter task.

DiffServ sees deployment in many networks for the time being. As described in the introduction of the draft DiffServ problem statement [I-D.polk-tsvwg-diffserv-stds-problem-statement], remarking of packets at domain boundaries is a DiffServ feature. This draft proposes a set of standard QoS classes and codepoints at interconnection points to which and from which locally used classes and codepoints should be mapped. Such a scheme simplifies interconnection negotiations and ensures that end to end class properties remain roughly the same while codepoints may change.

The proposed Interconnection class and codepoint scheme tries to reflect and consolidate related DiffServ and QoS standardisation efforts outside of the IETF, namely MEF, GSMA and ITU.

IP Precedence has been deprecated when DiffServ was standardised. It is common practice today however to copy the DSCPs Bits 0-2 (called Class Selector Codepoints in the following) into MPLS TC or Ethernet P-Bits. This is also reflected by the DiffServ codepoint definitions of AF and EF. The Class Selector Codepoints shouldn't be used for backward compatibility only. Class based PHBs may be applied in core network sections rather than then DSCP based PHBs.

The set of available router and traffic management tools to configure and operate DiffServ classes is limited. This should be reflected by class definitions. These may in the end be more related to transport properties than to application requirements. Please interpret transport properties as "congestion aware" and "not congestion aware" rather then TCP or UDP.

Finally, this draft proposes to leave some lass Selector Codepoint and by that MPLS TC codepoint space to allow for future DiffServ extensions like ECN/PCN and domain internal classes. An example for an internal PHB may be CS6. Some operators protect their network internal routing and / or management traffic by CS6. This PHB is possibly not available to transport customer or interconnection partner signaling and management traffic.

In addition to the standardisation activities which triggered this work, other authors published RFCs or drafts which may benefit from an interconnection class- and codepoint scheme. RFC 5160 suggests Meta-QoS-Classes to enable deployment of standardised end to end QoS classes [RFC5160]. The authors agree that the proposed interconnection class- and codepoint scheme as well as the idea of standardised end to end classes would complement their own work. Work on signaling Class of Service at interconnection interfaces by BGP [I-D.knoll-idr-cos-interconnect], [ID.idr-sla] is beyond the scope of this draft. Should the basic transport and class properties be standardised as proposed here, signaled access to QoS classes may be of interest. The current BGP drafts focus on exchanging SLA and traffic conditioning parameters. They seem to assume that common interpretation of the PHB properties identified by DSCPs has been established prior to exchanging further details by BGP signaling.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. Terminology

This draft re-uses existing terminology.

Class Selector Codepoints The bits 0-2 of the DSCP (marked "x" in this generic DSCP field: xxx000) are called the Class Selector Codepoints [RFC2474]. As their purpose is not just backwards compatibility, they are used to enable IP to MPLS DiffServ interoperability.

Class A class is a set of one or more PHBs utilising the same PHB if classified by a single identical Class Selector Codepoint (e.g. an AF class [RFC2597]). It is a PHB Scheduling Class [RFC3260] or an Ordered Aggregate. A class is a PHB group [RFC2575]. Different classes must not be aggregated.

PHB On IP layer, a single DSCP identifies a single PHB. In addition, this document proposes an MPLS like classification of traffic for a single PHB based on the Class Selector Codepoint (see [RFC3270]).

The above references may be incomplete and mostly refer to the early DiffServ RFCs only.

On MPLS layer, the available DiffServ Coding space is called Traffic Class (TC) [RFC5462]. A Class Selector Codepoint may be set to the same value as the MPLS TC. This allows MPLS DiffServ treatment by

MPLS routers if a DSCP is at packet top after a Pen Ultimate Hop label pop (which seems to be best practice by the time of writing). Note that supporting Class Selector Codepoint based DiffServ means support of MPLS like DiffServ only. This document neither argues for nor supports any scheme based on two 3 bit field based PHB assignment on IP layer.

To gain clarity, "DSCP based PHB selection" is only meant if expressed exactly that way in the remaining document. "PHB" relates to Class Selector Codepoint based PHB selection.

The following current practice issues relate to the concept of the DiffServ interconnection class proposal rather than to terminology. They serve as additional motivation of this activity:

- o Abstract class names like "EF" are preferential over those being close to an application, like "Voice". Unfortunately, non QoS experts can’t handle abstract class names. Hence and usually sooner than later, classes are named for applications or groups of them. One consequence however is, that people tend to combine application group class names and SLA parameters. Based on an application specific name and some worst case performance numbers on a paper, they often decide that their application needs a separate new QoS class.

- o Worse than that, but very present in practice, is the class abstraction level which is preferred by those dealing with QoS (as experts or non experts): the DSCPs or the Class Selector Codepoints values. These are the commodity abstractions applied for QoS classes. Most of these persons have fixed class to codepoint mappings in their minds, which they can’t easily adapt on per customer or per interconnection partner basis.

While these issues aren’t to be solved by IETF (QoS experts could and should of course teach staff to use proper Diffserv terminology and concepts), a simple and comprehensible QoS interconnection class scheme also is helpful in this area.

3. An Interconnection class and codepoint scheme

DiffServ deployments mostly follow loose class specification schemes (often one or two AF classes, EF and Best Effort). Especially DSCP assignment for the AF classes varies between deployments. Basic AF class property definitions are often similar however. Applying provider specific DSCPs is in line with the DiffServ architecture. This document doesn’t propose to change that.

Interconnecting parties face the problem of matching classes to be interconnected and then to agree on codepoint mapping. As stated by draft DiffServ problem statement [I-D.polk-tsvwg-diffserv-stds-problem-statement], remarking is a standard behaviour at interconnection interfaces. This draft proposes a standard interconnection set of 4 QoS classes with well defined DSCP and Class Selector Codepoints values A sending party remarks DSCPs from internal schemes to the Interconnection codepoints. The receiving party remarks Class Selector Codepoints and / or DSCPs to her internal scheme. Thus the interconnection codepoint scheme fully complies with the DiffServ architecture. An interconnection class and codepoint scheme was introduced by ITU-T [Y.1566] (there also including Ethernet). It is specified to a higher level of detail in this document.

At first glance, this looks like an additional effort. But there are obvious benefits: each party sending or receiving traffic has to specify the mapping from or to the interconnection class and codepoint scheme only once. Without it, this is to be negotiated per interconnection party individually. Further, end-to-end QoS in terms of traffic being classified for the same class in all passed domains is likely to result if an interconnection codepoint scheme is used. It is not necessarily resulting from individual per network mapping negotiations.

The standards and deployments known to the author of this draft are limited to 4 DiffServ classes at interconnection points (or less).Draft RFC 4597 update [I-D.polk-tsvwg-rfc4594-update]doesn’t seem to generally contradict to this, as it proposes to standardise "many services classes, not all will be used in each network at any period of time." Some reasons favour working with 4 DiffServ interconnection classes:

o There should be a coding reserve for interconnection classes. This leaves space for future standards, for private bilateral agreements and for provider internal classes.

o MPLS and Ethernet support only 8 PHBs, classes or ECN indications. Assignment of 3 bit codepoints for whatever purpose must be well thought through. Limiting interconnection QoS to four classes is MPLS and Ethernet friendly in that sense.

o Migrations from one codepoint scheme to another may require spare QoS codepoints.

The proposed class and codepoint scheme is designed for point to point IP layer interconnections. Other types of interconnections are out of scope of this document. The basic class and codepoint scheme is applicable on Ethernet layer too.

4. Consolidation of QoS standards by the interconnection codepoint scheme

The interconnection class and codepoint scheme proposed by Y.1566 also tries to consolidate related DiffServ and QoS standardisation efforts outside of the IETF [Y.1566]. The interconnection class and codepoint scheme may be a suitable approach to consolidate these standards. MEF 23.1 specifies 3 aggregated classes, consuming up to 5 codepoints on Ethernet layer (EF, AF3, AF1 and Best Effort) and 5 PHBs [MEF23.1]. MEF aggregates AF1 and Default PHB in a single class. This is not recommended for interconnection, as it is not in line with RFC 2597 (which requires separate forwarding resources for each AF class and doesn’t foresee aggregation of Default PHB and an AF class).

GSMA IR.34 proposes four classes, EF, AF4, another AF class and Best Effort with 7 PHBs in sum [IR.34]. IR.34 specifies an "Interactive" class consisting of 3 PHBs with different priorities. IR.34 assigns the PHBS AF31, AF21 and AF11 to this Interactive class. This breaks RFC 2597. The proposed interconnection class and codepoint scheme supports an GSMA Interactive like class but assigns AF3 with PHBs AF31, AF32 and AF33.

If IETF picks up this draft, it may be a good idea to inform MEF and GSMA about conflicts of their standards with DiffServ and suggest joint activities to improve the situation. Information on interworking with MEF 23 and GSMA IR.34 with the interconnection QoS scheme could be given by a later version of this draft.

The classes to be supported at interconnection interfaces are specified by Y.1566 as:

Class Priority: EF, expecting the figures of merit describing the PHB to be in the range of low single digit milliseconds. See [RFC3246].

Bulk inelastic: Optimised for low loss, low delay, low jitter at high bandwidth. Traffic load in this class must be controlled, e.g. by application servers. One example could be flow admission control. There may be infrequent retransmissions requested by the application layer to mitigate low levels of packet losses. Discard of packets through active queue management should be avoided in this class. Congestion in this class may result in bursty packet loss. If used to carry multimedia traffic, it is recommended to carry audio and video traffic in a single PHB. All of these properties influence the buffer design.

Assured: This class may be optimised to transport traffic without bandwidth requirements. It aims on Very low loss at high bandwidths. Retransmissions after losses characterise the class and influence the buffer design. Active queue management with probabilistic dropping may be deployed.

Default: Default. This class may be optimised to transport traffic without bandwidth requirements. Retransmissions after losses characterise the class and influence the buffer design. Active queue management with probabilistic dropping may be deployed.

Note that other DiffServ related standards trim down class requirements to SLA parameters. To quote e.g. RFC 4594-update, "A "service class" represents a similar set of traffic characteristics for delay, loss, and jitter as packets traverse routers in a network." This draft adds traffic PHB properties corresponding to expected transport layer characteristics as a key factor to a class definition: the desired class performance like delay, jitter and worst case loss are met only if PHB and transport properties meet the ones described by the class definition. This is not to say, the other standards ignore PHB properties. They are e.g. a core part of RFC 4594-update. They do not directly refer to transport protocol properties, as most existing QoS standards prefer the approach of assigning QoS classes to applications or application sets. This may result in undesirable class mappings, if an e.g. IP TV application demanding low loss is matched to a class whose low loss guarantees depend on AQM mechanisms.

Y.1566 does not define a complete set of DSCP based PHBs to be supported at an interconnection interface. This information is added by this draft. At interconnection points, the following DSCP based PHBs should be accepted between interconnected parties:

Class: PHB (one or more)

Class Priority: EF

Bulk inelastic: AF41 (AF42 and AF43 are reserved for extension)

Assured: AF31, AF32 and AF33

Default: Default (i.e. Best Effort)

Class names (and property specification) have been picked from Y.1566 above.

A provider may prefer to operate an internal PHB for the routing and management traffic of own systems. The PHB may not be available for traffic of peers or customers classified for the same HB within their networks. By default, many routers mark this traffic by CS6. Several scenarios are possible:

o CS6 marked traffic originating within a domain should be mapped to a suitable PHB at interconnection interfaces, if the receiving provider isn't offering transport with CS6. AF31 is recommended to that purpose.

o BGP traffic terminating in the adjacent AS border router could carry any codepoint whose traffic is not dropped by the receiving AS border router.

o An AS border router may not be able to mark BGP traffic by any different DSCP than CS6 and this traffic might be destined to a distant BGP peer, like a routing arbiter. In that case, the interconnecting parties should negotiate the treatment of this traffic. Standard DiffServ remarking, picking e.g. AF31 or Best Effort are possible options.

Operating a provider internal network management and routing class is an option only. Providers may of course bilaterally agree to exchange CS6 marked traffic without changing the DSCP.

Maintaining a separate PHB for network management, routing or signaling traffic also for traffic transiting through or terminating in a remote AS may be desirable. AF31 is recommended to that purpose. This is simple in the case of VPNs or point to point services. If this traffic is multiplexed with arbitrary traffic using this DSCP based PHB, distinction by the codepoint only isn't possible any more. Hence a standard agreement would best solve the issue. This document recommends picking an Assured class DSCP based PHB, AF31.

5. MPLS, Ethernet and Class Selector Codepoints for aggregated classes

Ethernet and MPLS support 3 bit codepoint fields to differentiate service quality. Mapping of the Class Selector Codepoints to these 3 Bit fields has been a configuration restriction in the early days of DiffServ. The concept of classifying DiffServ traffic classes by the bits 0-2 of a DSCP has however been part of Diffserv from start on. EF's Class Selector Codepoints is 5, that of AF4 is 4 and so on. The interconnection class and codepoint scheme respects properties and limits of a 3 bit PHB coding space in different ways:

- o it allows to classify four interconnection classes based on Class Selector Codepoints.

- o it supports a single PHB group (AF3), whose DSCP based PHBs may be mapped to up to three different MPLS TC's or Ethernet P-Bits. Note that this draft doesn's favour or recommend doing that, but it is possible. The author isn't aware of deployed service offers with 3 different drop levels in a single class.

The above statement is no requirement to depricate any DSCP to MPLS TC or Ethernet P-Bit mapping functionality. In the opposite, by limiting the interconnection scheme to 7 DSCP based PHBs, each PHB may be mapped to a 3 Bit based PHB scheme.

6. QoS class name selection

This is more of an informational discussion, proposed best practice, and mainly relates to human behaviour (including QoS experts) rather than technical issues. Above the human preference for conceivable class names has been mentioned. Network engineers (including the former Diffserv WG authors) recommend avoiding application related QoS class names. Focus should be put on class properties. These can be irritating again. Just looking at SLA parameters like Delay, Jitter and packet loss doesn't tell the reader, which transport properties guided the related scheduler engineering of a PHB. A router produces QoS with a scheduling mechanism, a settable queue depth and optional active queue management (including ECN), and may be a policer. Some kind of resource management may be present (also in Diffserv domains). It's beyond the imagination of the author how one would engineer more than half a dozen classes with distinguishable properties using this set of tools.

There's no perfect solution to the problem, as PHB configurations are not comprehensible to most readers, even if they were communicated (they are operational secrets of course). There are (or should be) engineering assumptions, when designing QoS PHBs. They closer relate to layer 3 or layer 4 level properties than to specific applications. In most cases, an application responds to congestion by reducing traffic, or it ignores congestion. Active queue management doesn't help to avoid congestion in the latter case, only resource management does. EF may be a special case. If the EF traffic is not responsive to congestion, and packets are assumed to be short, rather small jitter values can be reached if engineering ensures that the packet arrival rate never exceeds the transmission rate of that queue (see RFC 3246 [RFC3246]). There's other non congestion-responsive traffic, for which the EF engineering assumptions may not fit. So support of a PHB like bulk inelastic is reasonable.

Active queue management may be deployed for QoS classes designed to transport traffic responding to congestion by traffic reduction.

The class names of this document follow Y.1566. TCP_optimised and especially UDP_optimised are inappropriate class names, as some UDP based applications are or may be expected to become TCP friendly.

7. Allow for DiffServ extendibility on MPLS and Ethernet level

Any aggregated Diffserv deployment faces codepoint depletion issues rather soon, if deployed on MPLS or Ethernet. Coding space should be left for new features, like ECN, PCN or Conex. In addition to carrying customer traffic, internal routing and network management traffic may be protected by using a separate class. Offering interconnection with up to four classes and 4 - 6 MPLS TC's (or Ethernet P-bits) to that respect is probably at least a fair compromise.

8. Acknowledgements

David Black gave many helpful comments to this work. Al Morton and Sebastien Jobert provided feedback on many aspects during private discussions. Brian Carpenter, Mohamed Boucadair and Thomas Knoll helped adding awareness of further potentially related work.

9. IANA Considerations

This memo includes no request to IANA.

10. Security Considerations

This document does not introduce new features, it describes how to use existing ones. The security section of RFC 4597 [RFC4597] applies.

11. References

11.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[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, December 1998.

[RFC2575] Wijnen, B., Presuhn, R., and K. McCloghrie, "View-based Access Control Model (VACM) for the Simple Network Management Protocol (SNMP)", RFC 2575, April 1999.

[RFC2597] Heinanen, J., Baker, F., Weiss, W., and J. Wroclawski, "Assured Forwarding PHB Group", RFC 2597, June 1999.

[RFC3246] Davie, B., Charny, A., Bennet, J., Benson, K., Le Boudec, J., Courtney, W., Davari, S., Firoiu, V., and D. Stiliadis, "An Expedited Forwarding PHB (Per-Hop Behavior)", RFC 3246, March 2002.

[RFC3260] Grossman, D., "New Terminology and Clarifications for Diffserv", RFC 3260, April 2002.

[RFC3270] Le Faucheur, F., Wu, L., Davie, B., Davari, S., Vaananen, P., Krishnan, R., Cheval, P., and J. Heinanen, "Multi-Protocol Label Switching (MPLS) Support of Differentiated Services", RFC 3270, May 2002.

[RFC5462] Andersson, L. and R. Asati, "Multiprotocol Label Switching (MPLS) Label Stack Entry: "EXP" Field Renamed to "Traffic Class" Field", RFC 5462, February 2009.

[min_ref] authSurName, authInitials., "Minimal Reference", 2006.

11.2. Informative References

[I-D.knoll-idr-cos-interconnect]
Knoll, T., "BGP Class of Service Interconnection", draft-knoll-idr-cos-interconnect-10 (work in progress), May 2013.

[I-D.polk-tsvwg-diffserv-stds-problem-statement]
Polk, J., "The Problem Statement for the Standard Configuration of DiffServ Service Classes", draft-polk-tsvwg-diffserv-stds-problem-statement-00 (work in progress), July 2012.

[I-D.polk-tsvwg-rfc4594-update]
Polk, J., "Standard Configuration of DiffServ Service Classes", draft-polk-tsvwg-rfc4594-update-03 (work in progress), March 2013.

[ID.idr-sla]
IETF, "Inter-domain SLA Exchange", IETF, http://datatracker.ietf.org/doc/draft-ietf-idr-sla-exchange/, 2013.

[IR.34] GSMA Association, "IR.34 Inter-Service Provider IP Backbone Guidelines Version 7.0", GSMA, GSMA IR.34 http://www.gsma.com/newsroom/wp-content/uploads/2012/03/ir.34.pdf, 2012.

[MEF23.1] MEF, "Implementation Agreement MEF 23.1 Carrier Ethernet Class of Service Phase 2", MEF, MEF23.1 http://metroethernetforum.org/PDF_Documents/technical-specifications/MEF_23.1.pdf, 2012.

[RFC4597] Even, R. and N. Ismail, "Conferencing Scenarios", RFC 4597, August 2006.

[RFC5127] Chan, K., Babiarz, J., and F. Baker, "Aggregation of Diffserv Service Classes", RFC 5127, February 2008.

[RFC5160] Levis, P. and M. Boucadair, "Considerations of Provider-to-Provider Agreements for Internet-Scale Quality of Service (QoS)", RFC 5160, March 2008.

[Y.1566] ITU-T, "Quality of service mapping and interconnection between Ethernet, IP and multiprotocol label switching networks", ITU, http://www.itu.int/rec/T-REC-Y.1566-201207-I/en, 2012.

Appendix A. Change log

00 to 01 Added terminology and references. Added details and information to interconnection class and codepoint scheme. Editorial changes.

01 to 02 Added some references regarding related work. Clarified class definitions. Further editorial improvements.

02 to 03 Consistent terminology. Discussion of Network Management PHB at interconnection interfaces. Editorial review.

Author's Address

Ruediger Geib (editor)
Deutsche Telekom
Heinrich Hertz Str. 3-7
Darmstadt, 64295
Germany

Phone: +49 6151 5812747
Email: Ruediger.Geib@telekom.de