

Network Working Group
Request for Comments: 3260
Updates: 2474, 2475, 2597
Category: Informational

D. Grossman
Motorola, Inc.
April 2002

New Terminology and Clarifications for Diffserv

Status of this Memo

This memo provides information for the Internet community. It does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

Copyright Notice

Copyright (C) The Internet Society (2002). All Rights Reserved.

Abstract

This memo captures Diffserv working group agreements concerning new and improved terminology, and provides minor technical clarifications. It is intended to update RFC 2474, RFC 2475 and RFC 2597. When RFCs 2474 and 2597 advance on the standards track, and RFC 2475 is updated, it is intended that the revisions in this memo will be incorporated, and that this memo will be obsoleted by the new RFCs.

1. Introduction

As the Diffserv work has evolved, there have been several cases where terminology has needed to be created or the definitions in Diffserv standards track RFCs have needed to be refined. Some minor technical clarifications were also found to be needed. This memo was created to capture group agreements, rather than attempting to revise the base RFCs and recycle them at proposed standard. It updates in part RFC 2474, RFC 2475 and RFC 2597. RFC 2598 has been obsoleted by RFC 3246, and clarifications agreed by the group were incorporated in that revision.

2. Terminology Related to Service Level Agreements (SLAs)

The Diffserv Architecture [2] uses the term "Service Level Agreement" (SLA) to describe the "service contract... that specifies the forwarding service a customer should receive". The SLA may include traffic conditioning rules which (at least in part) constitute a Traffic Conditioning Agreement (TCA). A TCA is "an agreement

specifying classifier rules and any corresponding traffic profiles and metering, marking, discarding and/or shaping rules which are to apply...."

As work progressed in Diffserv (as well as in the Policy WG [6]), it came to be believed that the notion of an "agreement" implied considerations that were of a pricing, contractual or other business nature, as well as those that were strictly technical. There also could be other technical considerations in such an agreement (e.g., service availability) which are not addressed by Diffserv. It was therefore agreed that the notions of SLAs and TCAs would be taken to represent the broader context, and that new terminology would be used to describe those elements of service and traffic conditioning that are addressed by Diffserv.

- A Service Level Specification (SLS) is a set of parameters and their values which together define the service offered to a traffic stream by a DS domain.
- A Traffic Conditioning Specification (TCS) is a set of parameters and their values which together specify a set of classifier rules and a traffic profile. A TCS is an integral element of an SLS.

Note that the definition of "Traffic stream" is unchanged from RFC 2475. A traffic stream can be an individual microflow or a group of microflows (i.e., in a source or destination DS domain) or it can be a BA. Thus, an SLS may apply in the source or destination DS domain to a single microflow or group of microflows, as well as to a BA in any DS domain.

Also note that the definition of a "Service Provisioning Policy" is unchanged from RFC 2475. RFC 2475 defines a "Service Provisioning Policy" as "a policy which defines how traffic conditioners are configured on DS boundary nodes and how traffic streams are mapped to DS behavior aggregates to achieve a range of services." According to one definition given in RFC 3198 [6], a policy is "...a set of rules to administer, manage, and control access to network resources". Therefore, the relationship between an SLS and a service provisioning policy is that the latter is, in part, the set of rules that express the parameters and range of values that may be in the former.

Further note that this definition is more restrictive than that in RFC 3198.

3. Usage of PHB Group

RFC 2475 defines a Per-hop behavior (PHB) group to be:

"a set of one or more PHBs that can only be meaningfully specified and implemented simultaneously, due to a common constraint applying to all PHBs in the set such as a queue servicing or queue management policy. A PHB group provides a service building block that allows a set of related forwarding behaviors to be specified together (e.g., four dropping priorities). A single PHB is a special case of a PHB group."

One standards track PHB Group is defined in RFC 2597 [3], "Assured Forwarding PHB Group". Assured Forwarding (AF) is a type of forwarding behavior with some assigned level of queuing resources and three drop precedences. An AF PHB Group consists of three PHBs, and uses three Diffserv Codepoints (DSCPs).

RFC 2597 defines twelve DSCPs, corresponding to four independent AF classes. The AF classes are referred to as AF1x, AF2x, AF3x, and AF4x (where 'x' is 1, 2, or 3 to represent drop precedence). Each AF class is one instance of an AF PHB Group.

There has been confusion expressed that RFC 2597 refers to all four AF classes with their three drop precedences as being part of a single PHB Group. However, since each AF class operates entirely independently of the others, (and thus there is no common constraint among AF classes as there is among drop precedences within an AF class) this usage is inconsistent with RFC 2475. The inconsistency exists for historical reasons and will be removed in future revisions of the AF specification. It should now be understood that AF is a type of PHB group, and each AF class is an instance of the AF type.

Authors of new PHB specifications should be careful to adhere to the RFC 2475 definition of PHB Group. RFC 2475 does not prohibit new PHB specifications from assigning enough DSCPs to represent multiple independent instances of their PHB Group. However, such a set of DSCPs must not be referred to as a single PHB Group.

4. Definition of the DS Field

Diffserv uses six bits of the IPV4 or IPV6 header to convey the Diffserv Codepoint (DSCP), which selects a PHB. RFC 2474 attempts to rename the TOS octet of the IPV4 header, and Traffic Class octet of the IPV6 header, respectively, to the DS field. The DS Field has a six bit Diffserv Codepoint and two "currently unused" bits.

It has been pointed out that this leads to inconsistencies and ambiguities. In particular, the "Currently Unused" (CU) bits of the DS Field have not been assigned to Diffserv, and subsequent to the publication of RFC 2474, they were assigned for explicit congestion notification, as defined in RFC 3168 [4]. In the current text, a DSCP is, depending on context, either an encoding which selects a PHB or a sub-field in the DS field which contains that encoding.

The present text is also inconsistent with BCP 37, IANA Allocation Guidelines for Values in the Internet Protocol and Related Headers [5]. The IPV4 Type-of-Service (TOS) field and the IPV6 traffic class field are superseded by the 6 bit DS field and a 2 bit CU field. The IANA allocates values in the DS field following the IANA considerations section in RFC 2474, as clarified in section 8 of this memo.

The consensus of the DiffServ working group is that BCP 37 correctly restates the structure of the former TOS and traffic class fields.

Therefore, for use in future documents, including the next update to RFC 2474, the following definitions should apply:

- the Differentiated Services Field (DSField) is the six most significant bits of the (former) IPV4 TOS octet or the (former) IPV6 Traffic Class octet.
- the Differentiated Services Codepoint (DSCP) is a value which is encoded in the DS field, and which each DS Node MUST use to select the PHB which is to be experienced by each packet it forwards.

The two least significant bits of the IPV4 TOS octet and the IPV6 Traffic Class octet are not used by Diffserv.

When RFC 2474 is updated, consideration should be given to changing the designation "currently unused (CU)" to "explicit congestion notification (ECN)" and referencing RFC 3168 (or its successor).

The update should also reference BCP 37.

5. Ordered Aggregates and PHB Scheduling Classes

Work on Diffserv support by MPLS Label Switched Routers (LSRs) led to the realization that a concept was needed in Diffserv to capture the notion of a set of BAs with a common ordering constraint. This presently applies to AF behavior aggregates, since a DS node may not reorder packets of the same microflow if they belong to the same AF class. This would, for example, prevent an MPLS LSR, which was also

a DS node, from discriminating between packets of an AF Behavior Aggregate (BA) based on drop precedence and forwarding packets of the same AF class but different drop precedence over different LSPs. The following new terms are defined.

PHB Scheduling Class: A PHB group for which a common constraint is that, ordering of at least those packets belonging to the same microflow must be preserved.

Ordered Aggregate (OA): A set of Behavior Aggregates that share an ordering constraint. The set of PHBs that are applied to this set of Behavior Aggregates constitutes a PHB scheduling class.

6. Unknown/Improperly Mapped DSCPs

Several implementors have pointed out ambiguities or conflicts in the Diffserv RFCs concerning behavior when a DS-node receives a packet with a DSCP which it does not understand.

RFC 2475 states:

"Ingress nodes must condition all other inbound traffic to ensure that the DS codepoints are acceptable; packets found to have unacceptable codepoints must either be discarded or must have their DS codepoints modified to acceptable values before being forwarded. For example, an ingress node receiving traffic from a domain with which no enhanced service agreement exists may reset the DS codepoint to the Default PHB codepoint [DSFIELD]."

On the other hand, RFC 2474 states:

"Packets received with an unrecognized codepoint SHOULD be forwarded as if they were marked for the Default behavior (see Sec. 4), and their codepoints should not be changed."

RFC 2474 is principally concerned with DS-interior nodes. However, this behavior could also be performed in DS-ingress nodes AFTER the traffic conditioning required by RFC 2475 (in which case, an unrecognized DSCP would occur only in the case of misconfiguration). If a packet arrives with a DSCP that hadn't been explicitly mapped to a particular PHB, it should be treated the same way as a packet marked for Default. The alternatives were to assign it another PHB, which could result in misallocation of provisioned resources, or to drop it. Those are the only alternatives within the framework of RFC 2474. Neither alternative was considered desirable. There has been discussion of a PHB which receives worse service than the default; this might be a better alternative. Hence the imperative was "SHOULD" rather than "SHALL".

The intent of RFC 2475 clearly concerns DS-ingress nodes, or more precisely, the ingress traffic conditioning function. This is another context where the "SHOULD" in RFC 2474 provides the flexibility to do what the group intended. Such tortured readings are not desirable.

Therefore, the statement in RFC 2474 will be clarified to indicate that it is not intended to apply at the ingress traffic conditioning function at a DS-ingress node, and cross reference RFC 2475 for that case.

There was a similar issue, which manifested itself with the first incarnation of Expedited Forwarding (EF). RFC 2598 states:

To protect itself against denial of service attacks, the edge of a DS domain MUST strictly police all EF marked packets to a rate negotiated with the adjacent upstream domain. (This rate must be \leq the EF PHB configured rate.) Packets in excess of the negotiated rate MUST be dropped. If two adjacent domains have not negotiated an EF rate, the downstream domain MUST use 0 as the rate (i.e., drop all EF marked packets).

The problem arose in the case of misconfiguration or routing problems. An egress DS-node at the edge of one DS-domain forwards packets to an ingress DS-node at the edge of another DS domain. These packets are marked with a DSCP that the egress node understands to map to EF, but which the ingress node does not recognize. The statement in RFC 2475 would appear to apply to this case. RFC 3246 [7] clarifies this point.

7. No Backward Compatibility With RFC 1349

At least one implementor has expressed confusion about the relationship of the DSField, as defined in RFC 2474, to the use of the TOS bits, as described in RFC 1349. The RFC 1349 usage was intended to interact with OSPF extensions in RFC 1247. These were never widely deployed and thus removed by standards action when STD 54, RFC 2328, was published. The processing of the TOS bits is described as a requirement in RFC 1812 [8], RFC 1122 [9] and RFC 1123 [10]. RFC 2474 states:

"No attempt is made to maintain backwards compatibility with the "DTR" or TOS bits of the IPv4 TOS octet, as defined in [RFC791].",

In addition, RFC 2474 obsoletes RFC 1349 by IESG action. For completeness, when RFC 2474 is updated, the sentence should read:

"No attempt is made to maintain backwards compatibility with the "DTR/MBZ" or TOS bits of the IPv4 TOS octet, as defined in [RFC791] and [RFC1349]. This implies that TOS bit processing as described in [RFC1812], [RFC1122] and [RFC1123] is also obsoleted by this memo. Also see [RFC2780]."

8. IANA Considerations

IANA has requested clarification of a point in RFC 2474, concerning registration of experimental/local use DSCPs. When RFC 2474 is revised, the following should be added to Section 6:

IANA is requested to maintain a registry of RECOMMENDED DSCP values assigned by standards action. EXP/LU values are not to be registered.

9. Summary of Pending Changes

The following standards track and informational RFCs are expected to be updated to reflect the agreements captured in this memo. It is intended that these updates occur when each standards track RFC progresses to Draft Standard (or if some issue arises that forces recycling at Proposed). RFC 2475 is expected to be updated at about the same time as RFC 2474. Those updates will also obsolete this memo.

RFC 2474: revise definition of DS field. Clarify that the suggested default forwarding in the event of an unrecognized DSCP is not intended to apply to ingress conditioning in DS-ingress nodes. Clarify effects on RFC 1349 and RFC 1812. Clarify that only RECOMMENDED DSCPs assigned by standards action are to be registered by IANA.

RFC 2475: revise definition of DS field. Add SLS and TCS definitions. Update body of document to use SLS and TCS appropriately. Add definitions of PHB scheduling class and ordered aggregate.

RFC 2497: revise to reflect understanding that, AF classes are instances of the AF PHB group, and are not collectively a PHB group.

In addition, RFC 3246 [7] has added a reference to RFC 2475 in the security considerations section to cover the case of a DS egress node receiving an unrecognized DSCP which maps to EF in the DS ingress node.

10. Security Considerations

Security considerations are addressed in RFC 2475.

Acknowledgements

This memo captures agreements of the Diffserv working group. Many individuals contributed to the discussions on the Diffserv list and in the meetings. The Diffserv chairs were Brian Carpenter and Kathie Nichols. Among many who participated actively in these discussions were Lloyd Wood, Juha Heinanen, Grenville Armitage, Scott Brim, Sharam Davari, David Black, Gerard Gastaud, Joel Halpern, John Schnizlein, Francois Le Faucheur, and Fred Baker. Mike Ayers, Mike Heard and Andrea Westerinen provided valuable editorial comments.

Normative References

- [1] 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.
- [2] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z. and W. Weiss, "An Architecture for Differentiated Services", RFC 2475, December 1998.
- [3] Heinanen, J., Baker, F., Weiss, W. and J. Wrocklawski, "Assured Forwarding PHB Group", RFC 2597, June 1999.
- [4] Ramakrishnan, K., Floyd, S. and D. Black "The Addition of Explicit Congestion Notification (ECN) to IP", RFC 3168, September 2001.
- [5] Bradner, S. and V. Paxson, "IANA Allocation Guidelines for Values in the Internet Protocol and Related Headers", BCP 37, RFC2780, March 2000.
- [6] Westerinen, A., Schnizlein, J., Strassner, J., Scherling, M., Quinn, B., Herzog, S., Huynh, A., Carlson, M., Perry, J. and S. Waldbusser, "Terminology for Policy-Based Management", RFC 3198, November 2001.

- [7] Davie, B., Charny, A., Baker, F., Bennett, J.C.R., Benson, K., Le Boudec, J., Chiu, A., Courtney, W., Cavari, S., Firoiu, V., Kalmanek, C., Ramakrishnam, K. and D. Stiliadis, "An Expedited Forwarding PHB (Per-Hop Behavior)", RFC 3246, March 2002.
- [8] Baker, F., "Requirements for IP Version 4 Routers", RFC 1812, June 1995.
- [9] Braden, R., "Requirements for Internet Hosts -- Communications Layers", STD 3, RFC 1122, October 1989.
- [10] Braden, R., "Requirements for Internet Hosts -- Application and Support", STD 3, RFC 1123, October 1989.

Author's Address

Dan Grossman
Motorola, Inc.
20 Cabot Blvd.
Mansfield, MA 02048

EMail: dan@dma.isg.mot.com

Full Copyright Statement

Copyright (C) The Internet Society (2002). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.