An IPv4 Flowlabel Option
draft-dreibholz-ipv4-flowlabel-02

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

This Internet-Draft will expire on Oktober 30, 2004.

Copyright Notice

Copyright (C) The Internet Society (2004). All Rights Reserved.
Table of Contents

1. Introduction ........................................... 4
   1.1 Terminology ....................................... 4
   1.2 Abbreviations ..................................... 4
   1.3 Conventions ....................................... 4
2. A Flow Label Option for IPv4 .......................... 5
   2.1 Motivation ......................................... 5
      2.1.1 The Flow Label Field of IPv6 .................. 5
      2.1.2 The Limitations of IntServ via IPv4 .......... 6
   2.2 Definition of the Flow Label Option ................. 7
3. Translation between IPv6 and IPv4 ..................... 8
4. References ............................................... 8
Author’s Address ........................................... 9
Intellectual Property and Copyright Statements .......... 10
Abstract

This draft defines an IPv4 option containing a flowlabel that is compatible to IPv6. It is required for simplified usage of IntServ and interoperability with IPv6.
1. Introduction

1.1 Terminology

This document uses the following terms:
- **IntServ (Integrated Services):** Reservation of network resources (bandwidth) on a per-flow basis. See [3], [6], [7], [8], [9], [10] and [11] for details.
- **Flow:** An IntServ reservation between two endpoints.
- **Flow Label:** The Flow Label field of the IPv6 header and the IPv4 option header defined in this draft. It is used for marking a packet to use a specific IntServ reservation. See [4] for a detailed description.

1.2 Abbreviations

- **RSVP:** ReSource Reservation Protocol
- **TCP:** Transmission Control Protocol
- **QoS:** Quality of Service
- **UDP:** User Datagram Protocol

1.3 Conventions

The keywords **MUST, MUST NOT, REQUIRED, SHALL, SHALL NOT, SHOULD, SHOULD NOT, RECOMMENDED, NOT RECOMMENDED, MAY, and OPTIONAL,** when they appear in this document, are to be interpreted as described in [5].
2. A Flow Label Option for IPv4

2.1 Motivation

This section describes the motivation to add a flow label option to the IPv4 protocol.

2.1.1 The Flow Label Field of IPv6

The Flow Label field of the IPv6 header (see [12] and [4]) is a 20-bit pseudo-random number. All packets from the same source address having the same flow label MUST contain the same destination address. Therefore, the flow label combined with the source address is a network-unique identification for a specific packet flow. The idea behind the flow label is marking specific flows for IntServ. That is, the routers on the path from source to destination keep e.g. reservation states for the flows. The flow label provides easy identification and utilizes efficient lookup, e.g. using a hash function on the 3-tuple (source address, destination address, flow label).

Using the IPv6 flow label, packets can be mapped easily to specific flows, with the following features:

- **Transport Layer Protocol Independence**: Since the mapping is directly specified in the IP header, all possible layer 4 protocols are supported, even protocols to be specified in a far future.
- **Support for Network Layer Encryption**: The mapping is independent of payload encryption (e.g. by IPsec).
- **Support for Fragmentation**: If fragmentation of a large IP packet is necessary, all fragments contain the same flow label. Therefore, fragmentation does not cause any flow-marking problem.
- **Flow Sharing**: By marking packets with a flow label, it is possible to share a single flow (IntServ reservation) with several communication associations from host A to host B. For example, a video stream via UDP and a HTTP download via TCP could share a single reservation. For the user, flow sharing has the advantage that if one of its communication associations temporarily requires lower bandwidth than expected, other associations sharing the same flow may use the remaining bandwidth. That is, his possibly expensive reservation is fully utilized. Flow sharing also helps keeping the total number of reservations a router has to handle small, reducing their CPU and memory requirements and therefore cost.
- **Multi-Flow Connections**: One communication association can divide up its packets to several flows, simply by marking packets with different flow labels. This technique can be used for layered
transmission. That is, a stream (e.g. a video) is divided up into several parts (called layers). For example, the first layer (base layer) of a video contains a low-quality version, the second (1st enhancement layer) the data to generate a higher-quality version, etc.. Now, the first layer can be mapped to a high-quality reservation (guaranteed bandwidth, low loss rate) at higher cost, but the following layers can be mapped to lower-quality reservations (e.g. higher loss rate) or even best effort at lower cost. Research shows that the total transmission cost can be highly reduced using layered transmission (see [1] for details).

2.1.2 The Limitations of IntServ via IPv4

Using IntServ with IPv4, there are several problems that can only be solved with high management effort:

- No Transport Layer Protocol Independence: It is necessary to mark the packets within the layer 4 protocol header. For example, the TCP or UDP port numbers can be used to mark flows (with limitations, see below). But for new protocols (e.g. experimental, new standards, proprietary), software updates for *all* IntServ routers are necessary to recognize the packet flow!
- No Support for Network Layer Encryption: Since it is necessary to read fields of the layer 4 protocol header, it may not be encrypted. Therefore, e.g. the usage of IPsec is impossible.
- Support for Fragmentation: Only the first fragment of a large packet contains the layer 4 header necessary to map the packet to a flow. Mapping other fragments would require the hops to remember packet identities and try to map fragments to packet identities. Due to the management effort and memory requirements, this is not realistic for high-bandwidth backbone routers; especially when packet reordering must be considered. Furthermore, load sharing or traffic distribution would be impossible.
- No Flow Sharing: It is usually impossible for two different communication associations to share the same flow, e.g. if TCP flows are recognized using port numbers. This makes it necessary to reserve an IntServ flow for each communication association. This implies an increased number of flow states for routers to keep and maintain. Furthermore, if one association temporarily uses a lower bandwidth, the free bandwidth of its flow cannot easily be borrowed to another association.
- No Multi-Flow Connections: To use layered transmission, e.g. a video via UDP, the transmission of every layer would require own port numbers. In the case of connection-oriented transmission protocols (e.g. TCP, SCTP), every layer would even require its own connection setup and management. Depending on the transport protocol, the number of communication associations and the number of flows, much more work is necessary compared to IPv6 using flow labels.
All in all, using IntServ flows with IPv4 requires much more work compared to IPv6, where simply the flow label can be used. It is therefore useful to add such a field to IPv4, too. An appropriate place to add such a field is an IPv4 option header.

2.2 Definition of the Flow Label Option

IPv4 (see [2]) already defines an option header for a 16-bit SATNET stream identifier. Since this identifier would be incompatible to the 20-bit IPv6 flow label, reuse of this existing option header is inappropriate. Therefore, a new one is defined in the following.

Flow Label Option

+--------+--------+--------+--------+--------+
|10001111|00000010|0000     Flow Label       |
+--------+--------+--------+--------+--------+
Type=143 Length=5

Flow Label: 20 bits

The 20-bit flow label. All definitions of [4] and [12] for the IPv6 flow label are also valid for this field. A value of zero denotes that no flow label is used. In this case, the flow label option is in fact unnecessary. Note, that the option header contains 3 bytes and therefore 24 bits. The first 4 bits are unused and MUST be set to 0.

The Flow Label option MUST be copied on fragmentation. It MAY NOT appear more than once per IPv4 packet.

Note, that the flow label option’s length is 5 bytes. [2] requires that padding must be used to end the IP header on a 32 bit boundary. Therefore, the usual case with only the flowlabel option requires 3 padding bytes.
3. Translation between IPv6 and IPv4

Since the new IPv4 flow label is fully compatible to the IPv6 flow label, the field MAY be translated in the other protocol’s one during protocol translation. That is, a router can translate an IPv6 packet set from an IPv6-only host to an IPv4-mapped address of an IPv4-only host and the flow label may simply be copied. The same may also be applied in the backwards direction.

Note, that copying the flow label during protocol translation is not mandatory. There may be IntServ reservation reasons for not copying but setting the flow label to zero. But a router MAY NOT set the flow label to another value than the copy or 0, since the source is responsible to ensure that the source address combined with the flow label is network-unique.

4 References


Author’s Address

Thomas Dreibholz
University of Duisburg-Essen, Institute for Experimental Mathematics
Ellernstrasse 29
45326 Essen, Nordrhein-Westfalen
Germany

Phone: +49 201 183-7637
Fax: +49 201 183-7673
EMail: dreibh@exp-math.uni-essen.de
URI: http://www.exp-math.uni-essen.de/~dreibh/
Intellectual Property Statement

The IETF takes no position regarding the validity or scope of any intellectual property or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; neither does it represent that it has made any effort to identify any such rights. Information on the IETF’s procedures with respect to rights in standards-track and standards-related documentation can be found in BCP-11. Copies of claims of rights made available for publication and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementors or users of this specification can be obtained from the IETF Secretariat.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights which may cover technology that may be required to practice this standard. Please address the information to the IETF Executive Director.

Full Copyright Statement

Copyright (C) The Internet Society (2004). 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.

Acknowledgment

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