

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 14, 2013

S. Bryant, Ed.
J. Guichard
C. Pignataro
S. Spraggs
Cisco Systems
June 12, 2013

Carrying Metadata in IP Networks
draft-bryant-ip-metadata-00

Abstract

This document defines the mechanism for carrying and identifying the presence of metadata carried in addition to the payload in IPv4 and IPv6 packets.

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](#)].

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 14, 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

Internet-Draft

IP Metadata

June 2013

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	2
2.	Metadata Component Structure	2
3.	Metadata Channel Encapsulation	2
3.1.	The Metadata Insertion Process	3
3.2.	The Metadata Receive Process	4
3.3.	The Metadata removal Process	4
4.	Load-balancing Considerations	5
5.	IANA Considerations	5
6.	Security Considerations	6
7.	Contributing Authors	6
8.	Normative References	6

[1.](#) Introduction

This document defines the mechanism for carrying and identifying the presence of metadata carried in addition to the payload in IP packets. The metadata header is common across all encapsulations (including IPv4, IPv6, and MPLS) and is defined in [I-D.guichard-metadata-header]. In this document ant reference to IP is to be taken as IPv4 and IPv6 unless otherwise specified.

[2.](#) Metadata Component Structure

The addition of metadata to packets enables the instrumentation of user packets, and service chaining, although it is anticipated that the ability to allow packets to carry data of use to the infrastructure and specific handling instructions will enable other uses.

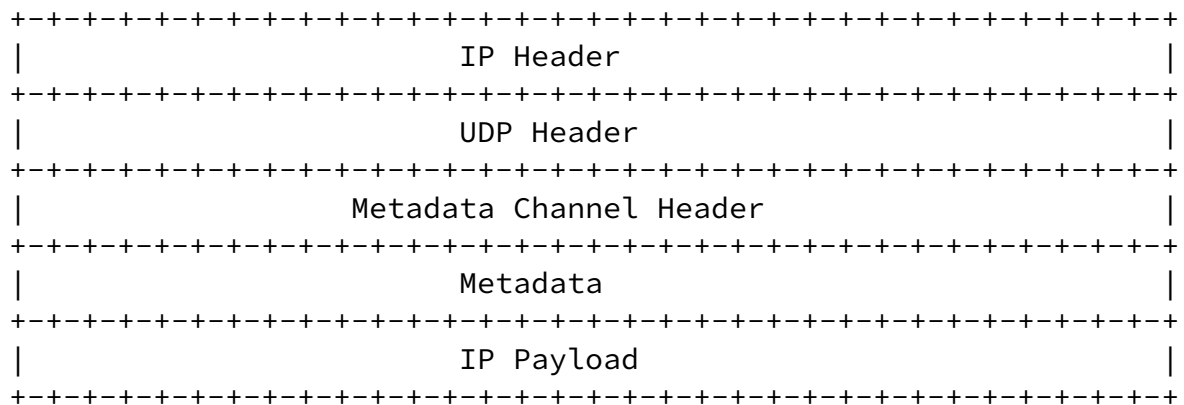
Metadata carried within an IP packet is carried in UDP and is prefaced by a Metadata Channel Header (MCH) as defined in [I-D.guichard-metadata-header].

The presence of metadata is identified by the UDP port number

assigned for this purpose.

[3.](#) Metadata Channel Encapsulation

An IP packet carrying metadata consists of the original IP header (except for the IP protocol type), a UDP header, the MCH, the metadata and the original IP payload. This is shown in Figure 1



Metadata in IP Encapsulation

Figure 1

[3.1.](#) The Metadata Insertion Process

In this section we describe the process of inserting metadata into an existing IP packet. This existing IP packet has been constructed in the normal way including the use of the normal protocol type or final next header type (IPv4 and Ipv6 respectively) to indicate the IP payload type, and the calculation of any transport layer checksums over the normal pseudoheader.

The IP header, and in the case of IPv6, the next headers are removed from the packet and stored. The metadata is prepended to the IP payload. The MCH is prepended to the metadata. The IPv4 protocol type or the final IPv6 next header is copied from the stored IP header into the protocol field of the MCH.

The UDP header is prepended to the MCH. The UDP destination port is set to the value "MCH-UDP" to indicate that an MCH follows. The UDP source port is used for load balancing as described in [Section 4](#). As in this case UDP is providing an encapsulation for the IP payload, and the IP payload can be assumed to have been adequately protected before the inclusion of the metadata and encapsulation in UDP, the UDP checksum [[RFC0768](#)], [[RFC6935](#)] MAY therefore be safely set to zero, [[RFC6936](#)].

The IP header is restored to the packet by prepending it to the UDP header. The IPv4 protocol field or the IPv6 final next header is overwritten with the value 17 to indicate that a UDP header follows.

The length of the IP packet is increased to compensate for the length of the UDP header, and the length of the metadata plus the metadata itself.

The IP checksum is also incrementally corrected to compensate for these modifications.

Any process functionally equivalent to the above is considered a compliant implementation.

It goes without saying that a packet can be constructed ab initio with the metadata included.

Fragmented IP packets are not supported.

[3.2](#). The Metadata Receive Process

The receipt of a packet with a UDP port of type "MCH-UDP" indicates the presence of metadata.

The MCH is examined and checked for the initial nibble of zero [I-D .guichard-metadata-header] and the MCH type is used to dispatch to the metadata process. The processing of the metadata is MCH type specific and outside the scope of this document.

If the MCH does not have an initial nibble of zero the packet is discarded and the configured management recording of the event undertaken (for example the event is counted).

[3.3.](#) The Metadata removal Process

In this section we describe the removal of the metadata and the reconstruction of the original IP packet.

The protocol field in the MCH is stored.

The IP header is prepended to the IP payload overwriting the MCH and the metadata.

The protocol field in the IP header is overwritten by the protocol field extracted from the MCH.

The length of the IP packet is decreased to compensate for the length of the UDP header, and the length of the metadata plus the metadata itself.

The IP checksum is also incrementally corrected to compensate for these modifications.

The reconstructed IP packet is processed as normal.

Any process functionally equivalent to the above is considered a compliant implementation.

It goes without saying that a packet does not need to be deconstructed by an application that is metadata aware.

[4.](#) Load-balancing Considerations

In an IPv4 packet with included metadata, the only field in the packet available to provide Equal Cost Multi-path (ECMP) Entropy is the UDP source port.

The choice of which components of the original packet to extract entropy from and how calculate the entropy value stored in the UDP source port is implementation and protocol specific and is not specified in the document. The following example of a method of calculating the source port is provided by way of example and is not normative. The source port value MAY be calculated by hashing the IP source and destination addresses, the original IP protocol type (or final next header type in the case of IPv6) and, if applicable the

source and destination ports of the original transport header and the result taken modulo 2^{16} . As the MCH is a unidirectional channel a source port value of zero is allowed.

Nothing in the above precludes the use of the IPv6 flow label to signal entropy to the routers on the network path, in which case the choice of UDP source port is outside the scope of this document.

[5.](#) IANA Considerations

This document requests that a UDP port is assigned to indicate the presence of metadata in the packet. This value is referred to as "MCH-UDP" in this document. The IANA specification of this UDP source port is:

Service Name = NLmetadata

Port Number = Next available from System Port Range.

Transport Protocol = UDP (only)

Description = Network Layer Metadata

Assignee = IESG

Contact = IETF Chair

Bryant, et al.

Expires December 14, 2013

[Page 5]

Internet-Draft

IP Metadata

June 2013

[6.](#) Security Considerations

The addition of metadata to a packet increases the amount of processing required by the router receiving the packet, and thus may be used in a denial of service attack vector. Metadata SHOULD be blocked on ingress to the network other than from trusted sources. Metadata SHOULD only be processed when received from a trusted source.

The metadata itself may be an attack vector with either the originating router or a man in the middle inserting malicious content, however the trust required from a router in the network is no different from the normal trust needed in this regard.

Any management action taken to record the occurrence of a defective

MCH or metadata payload SHOULD fall back to a lightweight process in the event of excessive events. This minimises the use of defective metadata packets as an attack vector.

If the ingress router is taking instructions from a third party in the specific metadata to insert, there MUST be a sufficient trust relationship between the ingress router and the third party.

The security considerations associated with each metadata type MUST be specified as part of its definition.

7. Contributing Authors

- o Clarence Filsfils (cfilsfil@cisco.com)
- o Dan Frost (danfrost@cisco.com)

8. Normative References

- [RFC0768] Postel, J., "User Datagram Protocol", STD 6, [RFC 768](#), August 1980.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC6935] Eubanks, M., Chimento, P., and M. Westerlund, "IPv6 and UDP Checksums for Tunneled Packets", [RFC 6935](#), April 2013.
- [RFC6936] Fairhurst, G. and M. Westerlund, "Applicability Statement for the Use of IPv6 UDP Datagrams with Zero Checksums", [RFC 6936](#), April 2013.

Authors' Addresses

Stewart Bryant (editor)
Cisco Systems

EMail: stbryant@cisco.com

Jim Guichard
Cisco Systems

EMail: jguichar@cisco.com

Carlos Pignataro
Cisco Systems

EMail: cpignata@cisco.com

Simon Spraggs
Cisco Systems

EMail: sspraggs@cisco.com