opsawg
Internet-Draft
Updates: 7012 (if approved)
Intended status: Standards Track
Expires: September 6, 2018

Z. Li
R. Gu
China Mobile
J. Dong
Huawei Technologies

March 5, 2018

Export BGP community information in IP Flow Information Export (IPFIX) draft-ietf-opsawg-ipfix-bgp-community-05

Abstract

This draft updates RFC7012 IPFIX information model by introducing several information elements (IEs) to enable IPFIX to export the BGP community information, including BGP standard community defined in RFC1997, BGP extended community defined in RFC4360, and BGP large community defined in RFC8092. Network traffic flow information can then be accumulated and analysed at the granularity specified by the BGP communities, which is suitable for and needed by some traffic optimization applications located in IPFIX collector, SDN controller or PCE (Path Computation Element).

To clarify, no new BGP community attribute is defined in this document and this document has no purpose to replace BGP Monitoring Protocol BMP defined in RFC7854. The IEs introduced in this document are used by IPFIX together with other IEs to facilitate the IPFIX collector analyzing the traffic in BGP community granularity without running the heavy BGP protocol.

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 https://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 September 6, 2018.

Table of Contents

1. Introduction

IP Flow Information Export (IPFIX) [RFC7011] provides network administrators with traffic flow information using the information elements (IEs) defined in [IANA-IPFIX] registries. Based on the traffic flow information, network administrators know the amount and direction of the traffic in their network, then they can optimize their network when needed. For example, they can shift some flows from the congested links to the low utilized links through a SDN controller or PCE [RFC4655].

[IANA-IPFIX] has already defined the following IEs for traffic flow information exporting in different granularities: sourceIPv4Address, sourceIPv4Prefix, destinationIPv4Address, destinationIPv4Prefix, bgpSourceAsNumber, bgpDestinationAsNumber, bgpNextHopIPv4Address, etc. In some circumstances, however, especially when traffic engineering and optimization are executed in the Tier 1 or Tier 2 operators' backbone networks, traffic flow information based on these IEs may not be suitable. Flow information based on IP address or IP prefix may provide much too fine granularity for a large network. On the contrary, flow information based on AS number may be too coarse.

BGP community is a BGP path attribute defined in IDR (Inter Domain Routing) working group. The already defined BGP community attribute includes the standard community defined in [RFC1997], the extended community defined in [RFC4360], and the large community defined in [RFC8092]. BGP community attribute has a variety of use cases, one common practice of which for the operators is to use BGP community with planned specific values in their field networks to represent the groups of customers, peers, geographical and topological regions. Please refer to [RFC4384], [RFC8195] and Section 3 of this document for the detailed examples. To know the traffic generated by differnt kinds of customers, from differnt geographical or topological regions, by differnt kinds of customers in differnt regions, we need the corresponding community information related to the traffic flow exported by IPFIX. Netwok traffic statistic in BGP community granularity is useful not only for the traffic analyzing, but also can then be used by other applications, such as the traffic optimization applications located in IPFIX collector, SDN controller or PCE. [Community-TE] also states analyzing network traffic information at the granularity specified by BGP community is prefered for inbound traffic engineering. However, there is no IE defined for BGP community attribute in [IANA-IPFIX] yet.

Flow information based on BGP community may be collected by a mediator defined in [RFC6183]. Mediator is responsible for the correlation between flow information and BGP community. However no IEs are defined in [RFC6183] for exporting BGP community information

in IPFIX. Furthermore, to correlate the BGP community with the flow information, mediator needs to learn BGP routes and perform lookup in the BGP routing table to get the matching entry for a specific flow. Neither BGP route learning nor routing table lookup is trivial for a mediator. Mediator is mainly introduced to release the performance requirement for the exporter [RFC5982]. In fact, to obtain the information for BGP related IEs that have already been defined, such as bgpSourceAsNumber, bgpDestinationAsNumber, and bgpNextHopIPv4Address, etc, exporter has to hold the up-to-date BGP routing table and perform lookup in the BGP routing table. The exporter can obtain the BGP community information in the same procedure, thus exporting BGP community information adds no more requirement for exporter. It is RECOMMENDED that the BGP community information be exported by the exporter directly using IPFIX.

Through running BGP [RFC4271] or BMP [RFC7854] and performing lookup in the BGP routing table to get the matching entry for a specific flow (we call it correlation), IPFIX collectors and other applications, such as SDN controller or PCE, can figure up the network traffic at BGP community granularity. However,neither running BGP or BMP protocol nor routing table lookup is trivial for the IPFIX collectors and other applications. Moreover correlation between IPFIX flow information and the BGP RIB on the exporter (such as router) is more accurate, compared to the correlation on a collector, since the BGP routing table may be updated when the IPFIX collectors and other applications reveive the IPFIX flow information. And as stated above, the exporter can obtain the BGP community information in the same procedure when it obtains other BGP related informaiton. So exporting the BGP community information directly by the exporter to the collector is the efficient and accurate way. If the IPFIX collectors and other applications only want to figure up the network traffic at BGP community granularity, they do not need to run the heavy BGP or BMP protocol when the BGP community information can be obtained by IPFIX. However, we have to clarify, the BMP protocol has its own application scenario, the mechanisum introduced in this document has no purpose to replace it.

This draft introduces new IEs to extend the IPFIX information model defined in [RFC7012] to export the BGP community information, including BGP standard community defined in [RFC1997], BGP extended community defined in [RFC4360], and BGP large community defined in [RFC8092]. Flow information, including packetDeltaCount, octetDeltaCount [RFC7012] etc, can then be accumulated and analysed by the collector or other applications, such as SDN controller or PCE [RFC4655], at the granularity specified by BGP community , which is useful for knowing the traffic generted by different kinds of customers, from differnt geographical or topological regions according to the operator's BGP community plan, and can then be used

by the traffic engineering or traffic optimization applications, especially in the backbone network. To clarify, no new BGP community attribute is defined in this document, IDR (Inter Domain Routing) working group is the right place to define new community attributes for the BGP protocol.

The IEs introduced in this document are applicable for both IPv4 and IPv6 traffic. Both exporter and mediator can use these IEs to export BGP community information in IPFIX.

Please refer Appendix A for the encoding example and Section 3 for a detailed use case.

2. Terminology

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].

3. BGP Community based Traffic Collection

[RFC4384] introduces the mechanism of using BGP standard communities and extended communities to collect the geographical and topological related information in BGP routing system. [RFC8195] gives some examples about the application of BGP large communities to represent the geographical regions. Since the network traffic at the BGP community granularity represents the traffic generted by different kinds of customers, from differnt geographical regions according to the network operator's BGP community plan, it is useful for the network operators to analyze and optimize the network traffic among different customers and regions. This section gives a use case in which the network operator uses the BGP community based traffic information to adjust the network paths for different traffic flows.

Considering the following scenario, AS C provides transit connection between AS A and B. By tagging with different BGP communities, the routes of AS A and B are categorized into several groups respectively with the operator's plan. For example community A:X and A:Y are used for the routes originated from different geographical regions in AS A, and community B:M and B:N are used for the routes representing the different kinds of customers in AS B, such as B:M is for the mobile customers and B:N is for the fixed line customers. By default, all traffic originated from AS A and destinated to AS B (we call it traffic A-B) goes through path C1-C2-C3 (call it Path-1) in AS C. When the link between C1 and C2 is congested, we cannot simply steer all the traffic A-B from Path-1 to Path C1-C4-C3 (call it Path-2), becuse it will cause the congestion in Path-2.

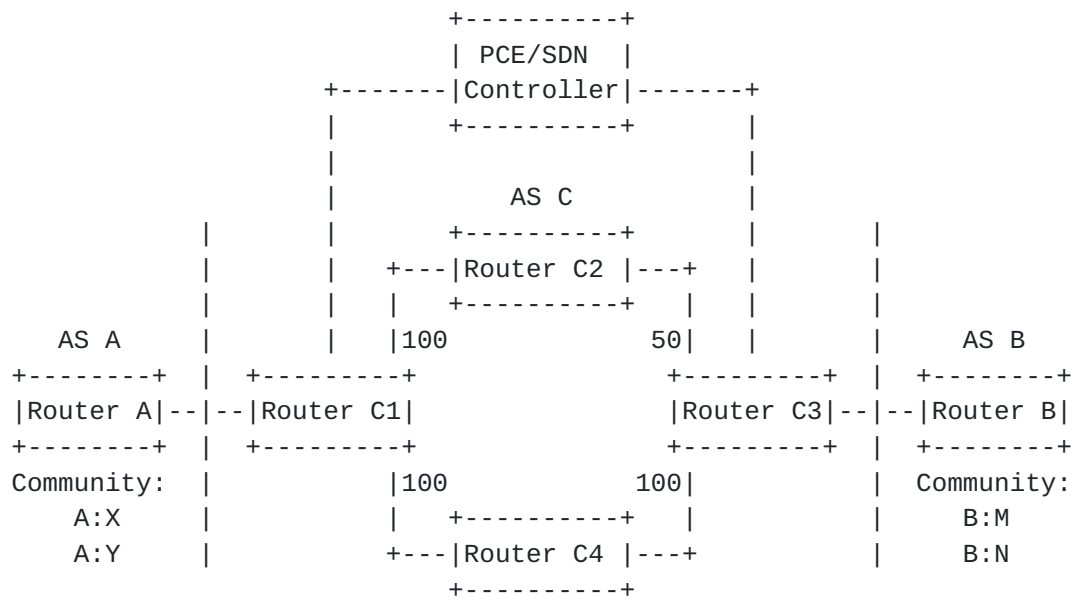


Figure 1: BGP Community based Traffic Collection

If the PCE/SDN controller in AS C can obtain the network traffic information at BGP community granularity, it can steer some traffic related to some BGP communities (when we consider only the source or destination of the traffic), or some BGP community pairs (when we consider both the source and the destination of the traffic) from Path-1 to Path-2 according to the utilization of different paths. For instance, steer the traffic generated by community A:X from Path-1 to Path-2 by deploying route policy at Router C1, or steer the traffic from community A:Y to community B:M from Path-1 to Path-2. Using the IEs defined in this document, IPFIX can export the BGP community information related to a specific traffic flow togecher with other flow information. The traffic information can then be accumulated at BGP community granularity and used by the PCE/SDN controller to steer the appropriate traffic from Path-1 to Path-2.

4. IEs for BGP Standard Community

[RFC1997] defines the BGP Communities attribute, called BGP Standard Community in this document, which describes a group of routes sharing some common properties. BGP Standard Communities are treated as 32 bit values as stated in[RFC1997].

In order to export BGP standard community information along with other flow information defined by IPFIX, three new IEs are introduced. One is bgpCommunity, which is used to identify that the value in this IE is a BGP standard community. The other two are bgpSourceCommunityList and bgpDestinationCommunityList, which are both basicList [RFC6313] of bgpCommunity, and are used to export BGP

standard community information corresponding to a specific flow's source IP and destination IP respectively.

The detailed information of the three new IEs are shown in the following sections.

4.1. bgpCommunity

ElementID	to be assigned by IANA
Name	bgpCommunity
Data Type	unsigned32
Data Type Semantics	identifier
Description	BGP community as defined in [RFC1997]
Units	none

Figure 2: bgpCommunity

4.2. bgpSourceCommunityList

ElementID	to be assigned by IANA
Name	bgpSourceCommunityList
Data Type	basicList, as specified in [RFC6313]
Data Type Semantics	list
Description	zero or more BGP communities corresponding with source IP address of a specific flow
Units	none

Figure 3: bgpSourceCommunityList

4.3. bgpDestinationCommunityList

ElementID	to be assigned by IANA
Name	bgpDestinationCommunityList
Data Type	basicList, as specified in [RFC6313]
Data Type Semantics	list
Description	zero or more BGP communities corresponding with destination IP address of a specific flow
Units	none

Figure 4: bgpDestinationCommunityList

5. IEs for BGP Extended Community

[RFC4360] defines the BGP Extended Communities attribute, which provides a mechanism for labeling the information carried in BGP. Each Extended Community is encoded as an 8-octet quantity with the format defined in [RFC4360].

In order to export BGP Extended Community information together with other flow information by IPFIX, three new IEs are introduced. The first one is bgpExtendedCommunity, which is used to identify that the value in this IE is a BGP Extended Community. The other two are bgpSourceExtendedCommunityList and bgpDestinationExtendedCommunityList, which are both basicList [RFC6313] of bgpExtendedCommunity, and are used to export the BGP Extended Community information corresponding to a specific flow's source IP and destination IP respectively.

The detailed information of the three new IEs are shown in the following sections.

5.1. bgpExtendedCommunity

ElementID	to be assigned by IANA
Name	bgpExtendedCommunity
Data Type	octetArray
Data Type Semantics	default
Description	BGP Extended Community as defined in [RFC4360] The size of this Information Element is 8 octets.
Units	none

Figure 5: bgpExtendedCommunity

5.2. bgpSourceExtendedCommunityList

ElementID	to be assigned by IANA
Name	bgpSourceExtendedCommunityList
Data Type	basicList, as specified in [RFC6313]
Data Type Semantics	list
Description	zero or more BGP Extended Communities corresponding with source IP address of a specific flow
Units	none

Figure 6: bgpSourceExtendedCommunityList

5.3. bgpDestinationExtendedCommunityList

ElementID	to be assigned by IANA
Name	bgpDestinationExtendedCommunityList
Data Type	basicList, as specified in [RFC6313]
Data Type Semantics	list
Description	zero or more BGP Extended communities corresponding with destination IP address of a specific flow
Units	none

Figure 7: bgpDestinationExtendedCommunityList

6. IEs for BGP Large Community

[RFC8092] defines the BGP Large Communities attribute, which is suitable for use with all Autonomous System Numbers (ASNs) including four-octet ASNs. Each BGP Large Community is encoded as a 12-octet quantity with the format defined in [RFC8092].

In order to export BGP Large Community information together with other flow information by IPFIX, three new IEs are introduced. The first one is bgpLargeCommunity, which is used to identify that the value in this IE is a BGP Large Community. The other two are bgpSourceLargeCommunityList and bgpDestinationLargeCommunityList, which are both basicList [RFC6313] of bgpLargeCommunity, and are used to export the BGP Large Community information corresponding to a specific flow's source IP and destination IP respectively.

The detailed information of the three new IEs are shown in the following sections.

6.1. bgpLargeCommunity

ElementID	to be assigned by IANA
Name	bgpLargeCommunity
Data Type	octetArray
Data Type Semantics	default
Description	BGP Large Community as defined in [RFC8092] The size of this Information Element is 12 octets.
Units	none

Figure 8: bgpLargeCommunity

6.2. bgpSourceLargeCommunityList

ElementID	to be assigned by IANA
Name	bgpSourceLargeCommunityList
Data Type	basicList, as specified in [RFC6313]
Data Type Semantics	list
Description	zero or more BGP Large Communities corresponding with source IP address of a specific flow
Units	none

Figure 9: bgpSourceLargeCommunityList

6.3. bgpDestinationLargeCommunityList

ElementID	to be assigned by IANA
Name	bgpDestinationLargeCommunityList
Data Type	basicList, as specified in [RFC6313]
Data Type Semantics	list
Description	zero or more BGP Large communities corresponding with destination IP address of a specific flow
Units	none

Figure 10: bgpDestinationLargeCommunityList

7. Operational Considerations

The maximum length of an IPFIX message is 65535 bytes as per [RFC7011] , and the maximum length of a normal BGP message is 4096 bytes as per [RFC4271]. Since BGP communities, including standard, extended, and large communities , are BGP path attributes carried in BGP Update messages, the total length of these attributes can not exceed the length of a BGP message, i.e. 4096 bytes. So one IPFIX message with maximum length of 65535 bytes has enough space to fit all the communities related to a specific flow, both the source IP and the destination IP related.

[I-D.ietf-idr-bgp-extended-messages] extends the maximum size of a BGP Update message to 65535 bytes. Then theoretically the BGP community information related to a specific flow may exceed the length one IPFIX message. However, according to the information about the networks in the field, the number of BGP communities in one BGP route is usually no more than 10. Nevertheless, BGP speakers that support the extended message SHOULD be careful to export the BGP communities in the IPFIX message properly, such as only convey as many communities as possible in the IPFIX message. The collector which receives an IPFIX message with maximum length and BGP communities contained in its data set SHOULD be aware that the BGP communities may be truncated due to limited message space. In this case, it is RECOMMENDED to configure export policy of BGP communities on the exporter to limit the BGP communities to be exported, so as to only export some specific communities,or not to export some specific communities.

If needed, we may consider to extend the message length of IPFIX [RFC7011] from 16 bits to 32 bits to solve this problem completely. The detailed mechanism is out of the scope of this document.

To align with the size of BGP extended community and large community, the size of IE bgpExtendedCommunity and bgpLargeCommunity is 8 octets and 12 octets respectively. In the event that the bgpExtendedCommunity or bgpLargeCommunity Elements are not of their expected sizes (8 and 12 octets, respectively), the IPFIX collector SHOULD ignore them. This is intended to protect implementations using BGP logic from calling their parsing routines with invalid lengths.

For the proper processing of the exporter, when it receives the template requesting to report the BGP community information (refer Appendix A for an example), the exporter SHOULD obtain the coressponding BGP community information through BGP lookup using the corresponding source or destination IP of the specific traffic flow. When exporting the IPFIX information to the collector, the exporter SHOULD include the corresponding BGP communities in the IPFIX message.

8. Security Considerations

This document only defines three new IEs for IPFIX. This document itself does not directly introduce security issues. The same security considerations as for the IPFIX Protocol Specification [RFC7011] and Information Model [RFC7012] apply.

As the BGP community information is deducible by other means, there are no increased privacy concerns, neither.

9. IANA Considerations

This draft specifies the following IPFIX IEs to export BGP community information along with other flow information.

The Element IDs for these IEs are solicited to be assigned by IANA. The following table is for IANA's reference to put in each field in the registry.

ElementID	Name	Data Type	Data Type Semantics
TBA1	bgpCommunity	unsigned32	identifier
TBA2	bgpSourceCommunityList	basicList	list

TBA3	bgpDestinationCommunityList	basicList	list
TBA4	bgpExtendedCommunity	octetArray	default
TBA5	bgpSourceExtended CommunityList	basicList	list
TBA6	bgpDestinationExtended CommunityList	basicList	list
TBA7	bgpLargeCommunity	octetArray	default
TBA8	bgpSourceLargeCommunityList	basicList	list
TBA9	bgpDestinationLarge CommunityList	basicList	list

ElementID	Description	Units
TBA1	BGP community as defined in [RFC1997]	
TBA2	zero or more BGP communities corresponding with source IP address of a specific flow	
TBA3	zero or more BGP communities corresponding with destination IP address of a specific flow	
TBA4	BGP Extended Community as defined in [RFC4360] The size of this IE is 8 octets	
TBA5	zero or more BGP Extended Communities corresponding with source IP address of a specific flow	
TBA6	zero or more BGP Extended communities corresponding with destination IP address of a specific flow	
TBA7	BGP Large Community as defined in [RFC8092] The size of this IE is 12 octets.	
TBA8	zero or more BGP Large Communities corresponding with source IP address of a specific flow	
	zero or more BGP Large communities	

| TBA9 | corresponding with destination IP address of a specific flow | | | | |

ElementID	Range	References	Requester	Revision	date
TBA1		RFC1997	this draft	0	
TBA2		RFC6313,RFC1997	this draft	0	
TBA3		RFC6313,RFC1997	this draft	0	
TBA4		RFC4360	this draft	0	
TBA5		RFC6313,RFC4360	this draft	0	
TBA6		RFC6313,RFC4360	this draft	0	
TBA7		RFC8092	this draft	0	
TBA8		RFC6313,RFC8092	this draft	0	
TBA9		RFC6313,RFC8092	this draft	0	

Figure 11: IANA Considerations

10. Acknowledgements

The authors would like to thank Benoit Claise and Paul Aitken for their comments and suggestions to promote this document. Appreciations are given to Tianran Zhou, Warren Kumari, Jeffrey Haas, Ignas Bagdonas, Stewart Bryant, Paolo Lucente, Job Snijders, Jared Mauch, Rudiger Volk, etc, for their discussion, comments and suggestions in the face to face meetings and in the mail list.

11. References

11.1. 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, <https://www.rfc-editor.org/info/rfc2119>.

[RFC6313] Claise, B., Dhandapani, G., Aitken, P., and S. Yates, "Export of Structured Data in IP Flow Information Export (IPFIX)", RFC 6313, DOI 10.17487/RFC6313, July 2011, <https://www.rfc-editor.org/info/rfc6313>.

[RFC7011] Claise, B., Ed., Trammell, B., Ed., and P. Aitken, "Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of Flow Information", STD 77, RFC 7011, DOI 10.17487/RFC7011, September 2013, <https://www.rfc-editor.org/info/rfc7011>.

[RFC7012] Claise, B., Ed. and B. Trammell, Ed., "Information Model for IP Flow Information Export (IPFIX)", RFC 7012, DOI 10.17487/RFC7012, September 2013, <https://www.rfc-editor.org/info/rfc7012>.

11.2. Informative References

[Community-TE] Shao, W., Devienne, F., Iannone, L., and JL. Rougier, "On the use of BGP communities for fine-grained inbound traffic engineering", Computer Science 27392(1):476-487, November 2015.

[I-D.ietf-idr-bgp-extended-messages] Bush, R., Patel, K., and D. Ward, "Extended Message support for BGP", draft-ietf-idr-bgp-extended-messages-24 (work in progress), November 2017.

[IANA-IPFIX] "IP Flow Information Export (IPFIX) Entities", <http://www.iana.org/assignments/ipfix/>.

[RFC1997] Chandra, R., Traina, P., and T. Li, "BGP Communities Attribute", RFC 1997, DOI 10.17487/RFC1997, August 1996, <https://www.rfc-editor.org/info/rfc1997>.

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI 10.17487/RFC4271, January 2006, <https://www.rfc-editor.org/info/rfc4271>.

[RFC4360] Sangli, S., Tappan, D., and Y. Rekhter, "BGP Extended Communities Attribute", RFC 4360, DOI 10.17487/RFC4360, February 2006, <https://www.rfc-editor.org/info/rfc4360>.

[RFC4384] Meyer, D., "BGP Communities for Data Collection", BCP 114, RFC 4384, DOI 10.17487/RFC4384, February 2006, <https://www.rfc-editor.org/info/rfc4384>.

[RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, DOI 10.17487/RFC4655, August 2006, <https://www.rfc-editor.org/info/rfc4655>.

[RFC5982] Kobayashi, A., Ed. and B. Claise, Ed., "IP Flow Information Export (IPFIX) Mediation: Problem Statement", RFC 5982, DOI 10.17487/RFC5982, August 2010, <https://www.rfc-editor.org/info/rfc5982>.

[RFC6183] Kobayashi, A., Claise, B., Muenz, G., and K. Ishibashi, "IP Flow Information Export (IPFIX) Mediation: Framework", RFC 6183, DOI 10.17487/RFC6183, April 2011, <https://www.rfc-editor.org/info/rfc6183>.

[RFC7854] Scudder, J., Ed., Fernando, R., and S. Stuart, "BGP Monitoring Protocol (BMP)", RFC 7854, DOI 10.17487/RFC7854, June 2016, <https://www.rfc-editor.org/info/rfc7854>.

[RFC8092] Heitz, J., Ed., Snijders, J., Ed., Patel, K., Bagdonas, I., and N. Hilliard, "BGP Large Communities Attribute", RFC 8092, DOI 10.17487/RFC8092, February 2017, <https://www.rfc-editor.org/info/rfc8092>.

[RFC8195] Snijders, J., Heasley, J., and M. Schmidt, "Use of BGP Large Communities", RFC 8195, DOI 10.17487/RFC8195, June 2017, <https://www.rfc-editor.org/info/rfc8195>.

Appendix A. Encoding Example

In this section, we give an example to show the encoding format for the new introduced IEs.

Flow information including BGP communities is shown in the below table. Suppose we want all the fields to be reported by IPFIX.

Source ip	Destination ip	Source BGP community	Destination BGP community
1.1.1.1	2.2.2.2	1:1001,1:1002,8:1001	2:1002,8:1001
3.3.3.3	4.4.4.4	3:1001,3:1002,8:1001	4:1001,8:1001

Figure 12: Flow information including BGP communities

A.1. Template Record

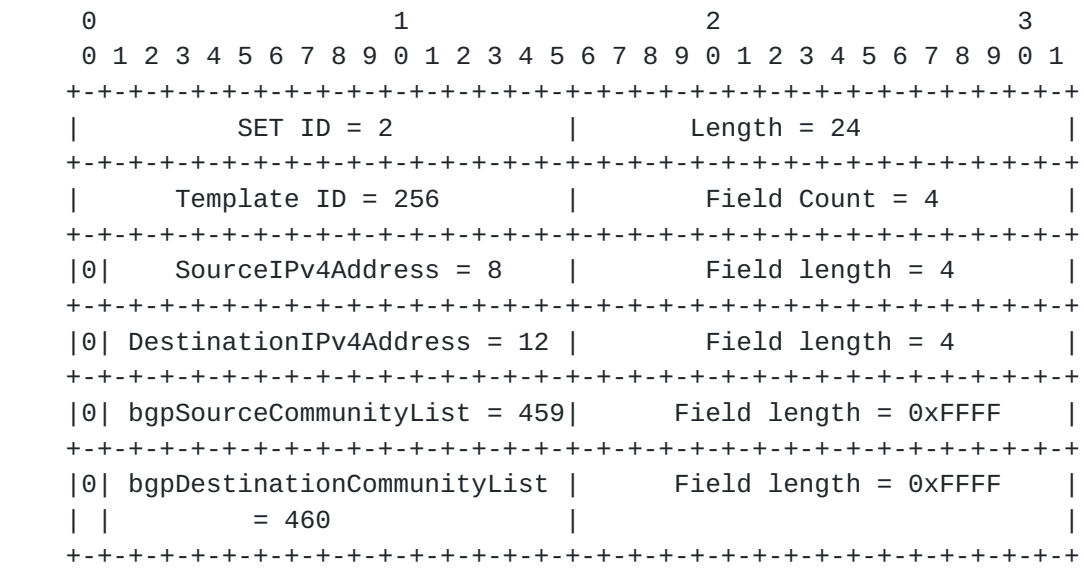
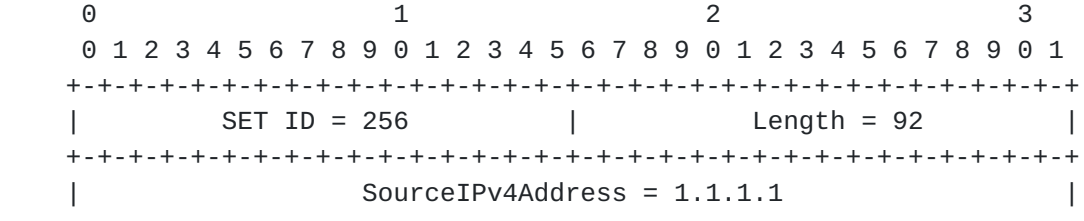


Figure 13: Template Record Encoding Format

In this example, the Template ID is 256, which will be used in the data record. The field length for bgpSourceCommunityList and bgpDestinationCommunityList is 0xFFFF, which means the length of this IE is variable, the actual length of this IE is indicated by the list length field in the basic list format as per [RFC6313].

A.2. Data Set

The data set is represented as follows:



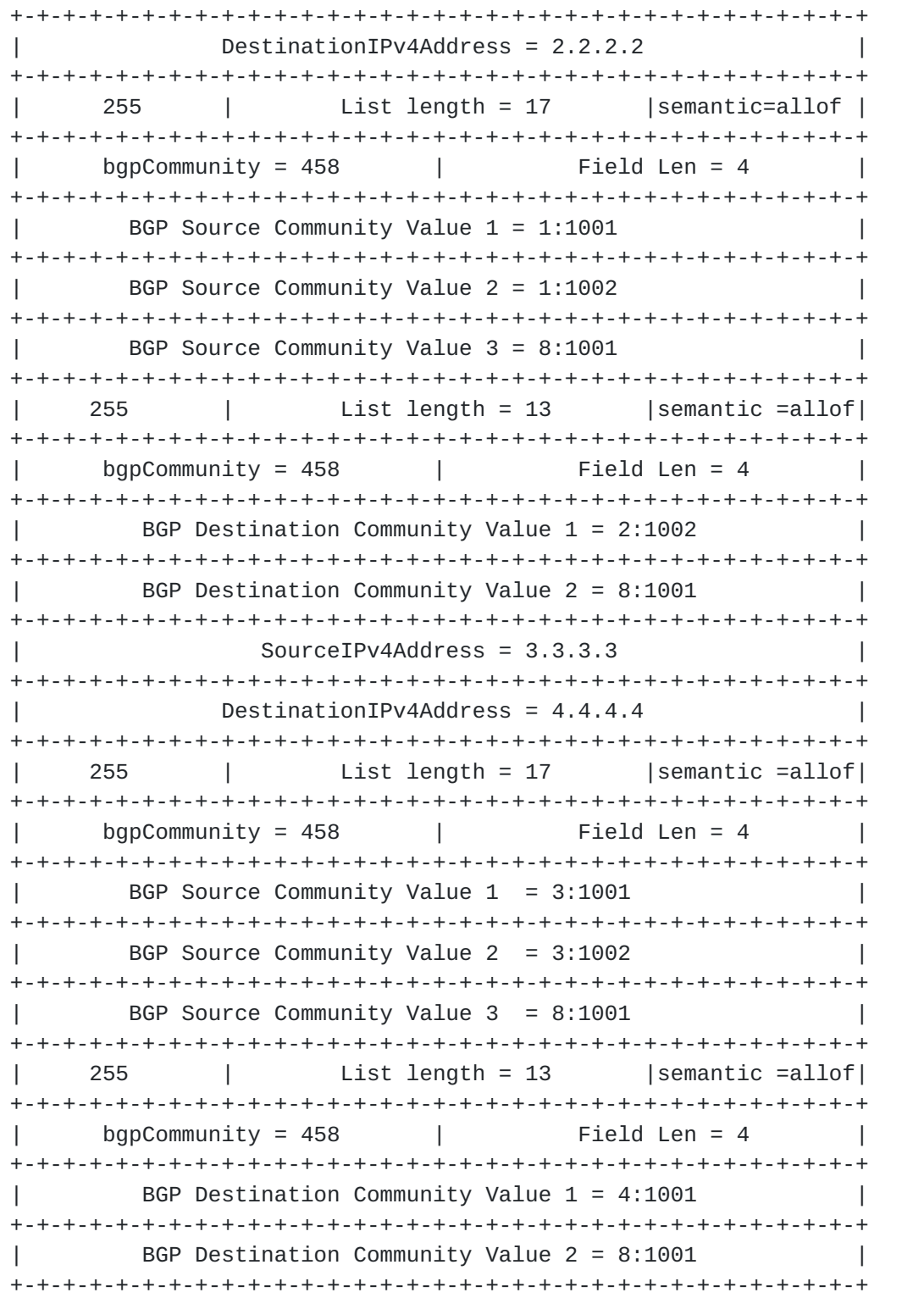


Figure 14: Data Set Encoding Format

Authors' Addresses

Zhenqiang Li
China Mobile
32 Xuanwumen West Ave, Xicheng District
Beijing 100053
China

Email: li_zhenqiang@hotmail.com

Rong Gu
China Mobile
32 Xuanwumen West Ave, Xicheng District
Beijing 100053
China

Email: gurong_cmcc@outlook.com

Jie Dong
Huawei Technologies
Huawei Campus, No. 156 Beiqing Rd.
Beijing 100095
China

Email: jie.dong@huawei.com