opsawg
Internet-Draft
Intended status: Standards Track
Expires: May 29, 2019

Z. Li
R. Gu
China Mobile
J. Dong
Huawei Technologies
November 25, 2018

Export BGP community information in IP Flow Information Export (IPFIX) draft-ietf-opsawg-ipfix-bgp-community-11

Abstract

By introducing new Information Elements (IEs), this draft extends the existing BGP-related IEs to enable IPFIX [RFC7011] to export BGP community information, including BGP standard communities [RFC1997], BGP extended communities [RFC4360], and BGP large communities [RFC8092]. Network traffic information can then be accumulated and analyzed at the BGP community granularity, which represents the traffic of different kinds of customers, services, or geographical regions according to the network operator's BGP community planning. Network traffic information at the BGP community granularity is useful for network traffic analysis and engineering.

To clarify, no new BGP community attribute is defined in this document and this document does not 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 network traffic at the BGP community granularity without needing to run the heavy BGP itself. When needed, the IPFIX Mediator or Collector can use the IEs introduced in this document to report the BGP community-related traffic flow information it gets either from Exporters or through local correlation to other IPFIX devices.

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 May 29, 2019.

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, and can then optimize their network when needed. For example, they can shift some flows

from congested links to low utilized links through an 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 Tier 1 or Tier 2 operators' backbone networks, traffic flow information based on these IEs may not be completely suitable or sufficient. For example, 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 that includes standard communities [RFC1997], extended communities [RFC4360], and large communities [RFC8092]. The BGP community attribute has a variety of use cases, one of which is to use BGP community with planned specific values to represent groups of customers, services, and geographical or topological regions, as used by operators in their networks. Detailed examples can be found in [RFC4384], [RFC8195] and Section 3 of this document. To understand the traffic generated by different kinds of customers, from different geographical or topological regions, by different kinds of customers in different regions, we need the corresponding community information related to the traffic flow information exported by IPFIX. Network traffic statistics at the BGP community granularity are useful not only for the traffic analyzing, but also can then be used by other applications, such as traffic optimization applications located in an IPFIX Collector, SDN controller or PCE. [Community-TE] also states that analyzing network traffic information at the BGP community granularity is preferred for inbound traffic engineering. However, [IANA-IPFIX] lacks IEs defined for the BGP community attribute.

Flow information based on BGP community may be collected by an IPFIX Mediator defined in [RFC6183]. IPFIX 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, the IPFIX Mediator needs to learn BGP routes and perform lookups in the BGP routing table to get the matching entry for a specific flow. Neither BGP route learning nor routing table lookup are trivial for an IPFIX Mediator. The IPFIX Mediator is mainly introduced to reduce the performance requirement for the Exporter [RFC5982]. In fact, to obtain the information for the already defined BGP related IEs, such as bgpSourceAsNumber, bgpDestinationAsNumber, and bgpNextHopIPv4Address, etc, the Exporter

has to hold the up-to-date BGP routing table and perform lookups in the table. The Exporter can obtain the BGP community information in the same procedure, thus the additional load added by exporting BGP community information is minimal if the Exporter is already exporting the existing BGP-related IEs. 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 lookups in the BGP routing table to correlate the matching entry for a specific flow, IPFIX Collectors and other applications, such as SDN controller or PCE, can determine the network traffic at the BGP community granularity. However, neither running BGP or BMP protocol nor routing table lookup are trivial for the IPFIX Collectors and other applications. Moreover, correlation between IPFIX flow information and the BGP RIB on the Exporter (such as a 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 receive the IPFIX flow information. And as stated above, the Exporter can obtain the BGP community information during the same procedure when it obtains other BGP related information. So exporting the BGP community information directly by the Exporter to the Collector is both efficient and accurate. If the IPFIX Collectors and other applications only want to determine the network traffic at the BGP community granularity, they do not need to run the full BGP or BMP protocols when the BGP community information can be obtained by IPFIX. However, the BMP protocol has its own application scenario, and the mechanism introduced in this document is not meant to replace it.

By introducing new IEs, this draft extends the existing BGP-related IEs to enable IPFIX [RFC7011] to export BGP community information, including the BGP standard communities [RFC1997], BGP extended communities [RFC4360], and BGP large communities [RFC8092]. Flow information, including packetDeltaCount, octetDeltaCount [RFC7012], etc., can then be accumulated and analyzed by the Collector or other applications, such as an SDN controller or PCE [RFC4655], at the BGP community granularity, which is useful for measuring the traffic generated by different kinds of customers, from different 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.

The IEs introduced in this document are applicable for both IPv4 and IPv6 traffic. Both the Exporter and the IPFIX Mediator can use these IEs to export BGP community information in IPFIX. When needed, the IPFIX Mediator or Collector can use these IEs to report BGP community

related traffic flow information it gets either from Exporters or through local correlation to other IPFIX devices.

As stated above, the method introduced in this document is not the definitive and the only one to obtain BGP community information related to a specific traffic flow, but a possible, efficient and accurate one.

No new BGP community attributes are defined in this document.

Note that this document does not update the IPFIX specification [RFC7011] and the Information Model [RFC7012]. Rather, IANA's IPFIX registry [IANA-IPFIX] contains the current complete Information Element reference, per Section 1 of [RFC7012].

Please refer to [IANA-IPFIX] for the complete list of BGP-related IEs.

Please refer to Appendix A of this document 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].

IPFIX-specific terminology used in this document is defined in Section 2 of [RFC7011] and Section 2 of [RFC6183].

BGP standard community: The BGP Communities attribute defined in [RFC1997]. In order to distinguish it from BGP extended communities [RFC4360], and large communities [RFC8092], BGP Communities attribute is called BGP standard community in this document.

3. BGP Community-based Traffic Collection

[RFC4384] introduces the mechanism of using BGP standard community and extended community to collect the geographical and topological related information in the BGP routing system. [RFC8195] gives some examples of the application of BGP large communities to represent the geographical regions. Since the network traffic at the BGP community granularity represents the traffic generated by different kinds of customers, from different geographical regions according to the network operator's BGP community plan, it is useful for 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.

Consider the following scenario, AS C provides a transit connection between ASes A and B. By tagging with different BGP communities, the routes of AS A and B are categorized into several groups respectively in the operator's plan. For example, communities A:X and A:Y are used for the routes originated from different geographical regions in AS A, and communities 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 originating from AS A and destined 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), because it will cause congestion in Path-2.

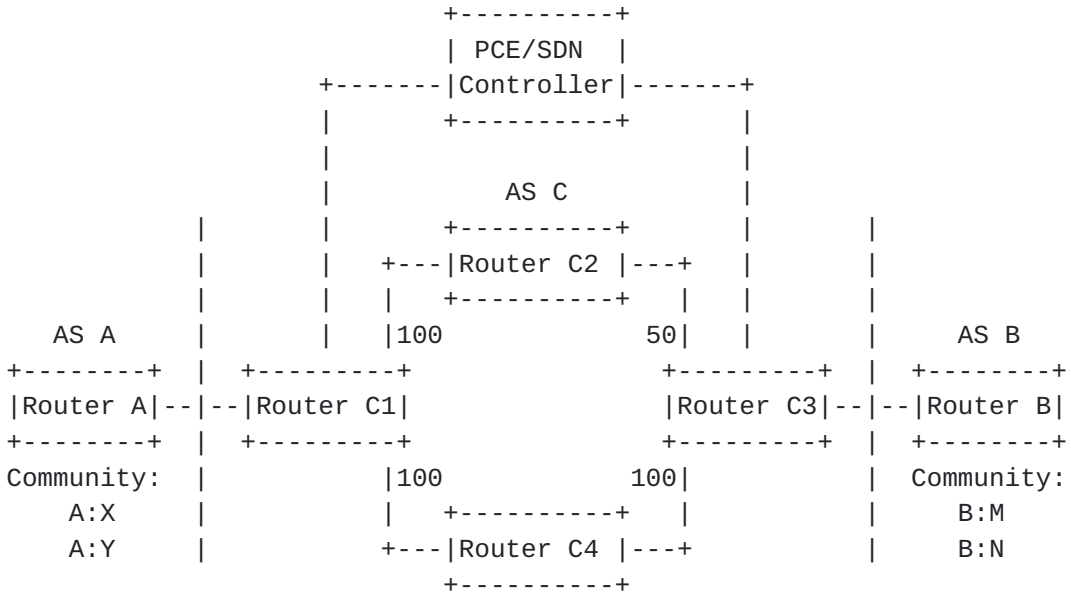


Figure 1: BGP Community based Traffic Collection

If the PCE/SDN controller in AS C can obtain the network traffic information at the 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 a 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 together

with other flow information. The traffic information can then be accumulated at the 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 Community is treated as 32 bit value 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 and destination IP address respectively.

The detailed information of the three new IEs are shown in Section 9, IANA Considerations.

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 and destination IP address respectively.

The detailed information of the three new IEs are shown in Section 9, IANA Considerations.

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 and destination IP address respectively.

The detailed information of the three new IEs are shown in Section 9, IANA Considerations.

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 a maximum length of 65535 bytes has enough space to fit all the communities related to a specific flow, relating to both the source and destination IP addresses.

[I-D.ietf-idr-bgp-extended-messages] extends the maximum size of a BGP Update message to 65535 bytes. In that case, the BGP community information related to a specific flow could theoretically exceed the length of one IPFIX message. However, according to information regarding actual networks in the field, the number of BGP communities in one BGP route is usually no more than ten. Nevertheless, BGP speakers that support the extended message SHOULD be careful to export the BGP communities in the IPFIX message properly, so that they 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 the export policy of BGP communities to limit the BGP communities by including or excluding specific communities.

If needed, the IPFIX message length could be extended from 16 bits to 32 bits to solve this problem completely. The details of increasing the IPFIX message length is out of scope of this document.

To align with the size of the BGP extended community and large community attributes, the size of IE bgpExtendedCommunity and bgpLargeCommunity is 8 octets and 12 octets respectively. In the event that the bgpExtendedCommunity or bgpLargeCommunity IE is not of its expected size, the IPFIX Collector SHOULD ignore it. 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 to Appendix A for an example), the Exporter SHOULD obtain the corresponding BGP community information through BGP lookup using the corresponding source or destination IP address 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 new IEs for IPFIX. This document itself does not directly introduce any new 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 as well.

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 requested to be assigned by IANA. The following table is for IANA's use to place 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	basicList of zero or more bgpCommunity IEs, containing the BGP communities corresponding with source IP address of a specific flow	
TBA3	basicList of zero or more bgpCommunity IEs, containing the BGP communities corresponding with destination IP address of a specific flow	
TBA4	BGP Extended Community as defined in [RFC4360] The size of this IE MUST be 8 octets	
TBA5	basicList of zero or more bgpExtendedCommunity IEs, containing the BGP Extended Communities corresponding with source IP address of a specific flow	
TBA6	basicList of zero or more bgpExtendedCommunity IEs, containing the 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 MUST be 12 octets.	
TBA8	basicList of zero or more bgpLargeCommunity IEs, containing the BGP Large Communities corresponding with source IP address of a specific flow	
	basicList of zero or more bgpLargeCommunity	

	IEs, containing the 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 2: IANA Considerations

10. Acknowledgements

The authors would like to thank Benoit Claise and Paul Aitken for their comments and suggestions to promote this document. We also thank Tianran Zhou, Warren Kumari, Jeffrey Haas, Ignas Bagdonas, Stewart Bryant, Paolo Lucente, Job Snijders, Jared Mauch, Rudiger Volk, and Andrew Malis for their discussion, comments, and suggestions to improve this document..

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>.

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-26 (work in progress), June 2018.

[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>.

[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>.

[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 provide an example to show the encoding format for the new introduced IEs.

Flow information, including BGP communities, is shown in the following table. In this example, all the fields are reported by IPFIX.

Source IP	Destination IP	BGP community corresponding with Source IP	BGP community corresponding with Destination IP
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 3: Flow information including BGP communities

A.1. Template Record

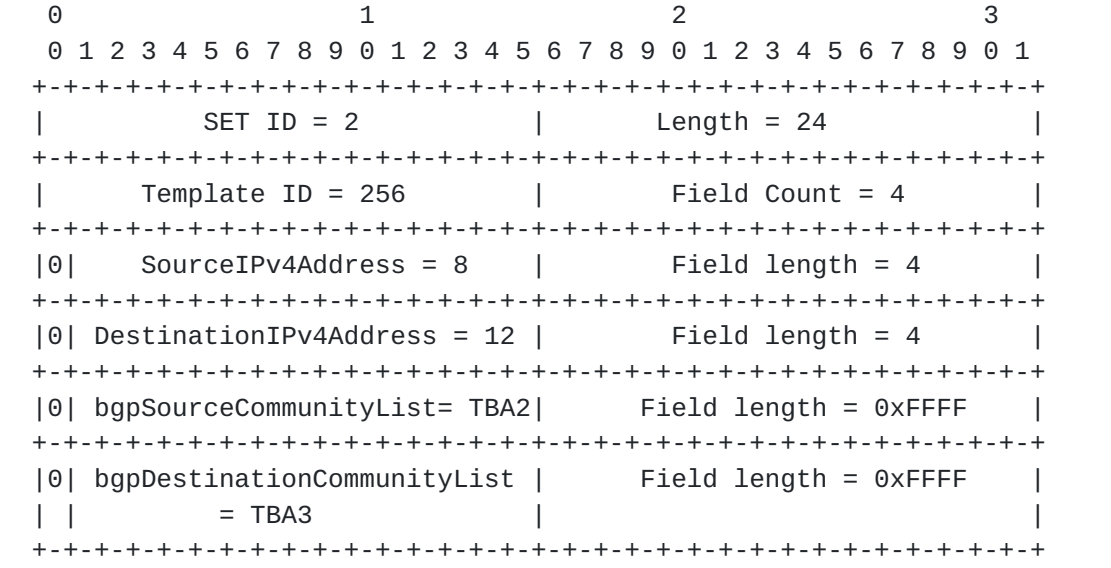
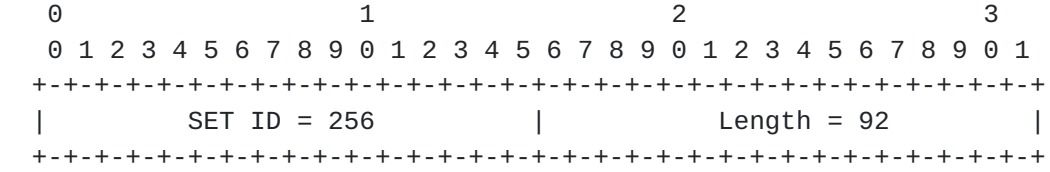


Figure 4: Template Record Encoding Format

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, and 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:



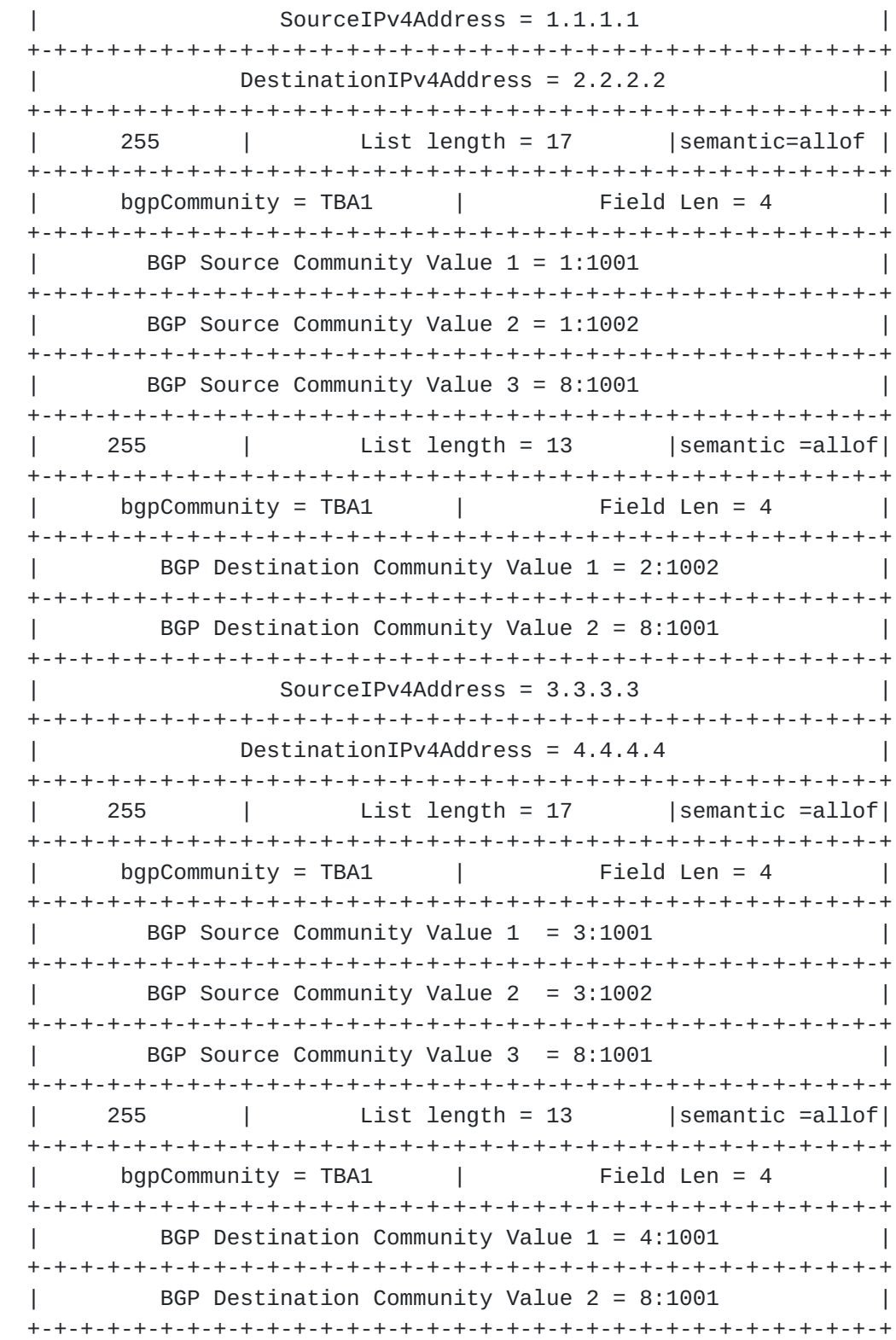


Figure 5: Data Set Encoding Format

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