

Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: 30 August 2026

ZB. ZOUNDI, Ed.
26 February 2026

BGP FLEX
draft-bonaventure-bgp-flex-01

Abstract

The Border Gateway Protocol (BGP) is a very important protocol in the Internet to exchange routing information between network domains. BGP populates the routing table with valid routes, sometimes the routes populated by BGP are not the best choices specially when a ISP provides IP Transit service to an other ISP, and the two ISP are peering at an Internet Exchange Point.

Let us assume that ISP A provides IP Transit to ISP B, and the two are peering at an Internet Exchange Point. The return traffic from Internet to ISP B via ISP A will be routed over the Internet Exchange Point instead of the IP Transit link between the two ISP, because of BGP Local preference which has usually a higher value on Internet Exchange links. This will lead to an issue because the return traffic IP Transit service is now passing over the Internet Exchange link.

An other issue faced in Internet Exchange Point, is the sub-optimal routing caused by the advertisement of the most specific routes over Internet by ISP.

Let us assume that ISP A wants to advertise most specific routes to Internet for traffic Engineering, the peers of ISP A at the Internet Exchange Point could also receive the most specific routes over Internet .

This will mean that traffic from the peers of ISP A towards ISP A will now go through Internet because of most specific routes which is not supposed to be the case.

To solve that ISP A will have to send the same specific routes over the Internet Exchange Point which will increase complexity.

This document provides alternative solutions that can be used to solve the above problems

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 30 August 2026.

Copyright Notice

Copyright (c) 2026 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 (<https://trustee.ietf.org/license-info>) in effect on the date of 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 Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction	3
2. Terminology	3
2.1. Acronyms	3
3. BGP Flex	3
4. BGP Flex Concepts	4
4.1. BGP Local Source IP (BLSI)	4
4.2. BGP Remote Source IP (BRSI)	4
4.3. BGP Remote Peer IP (BRPI)	4
4.4. Source IP Router Advertiser (SIRA)	4
5. Routing mode in BGP Flex	5
6. BGP vs BGP Flex	5
7. IANA Considerations	6
8. Security Considerations	6
9. References	6
9.1. Normative References	6
Author's Address	6

1. Introduction

In the past years BGP has been operating by relying on the same route table for Internet Exchange and Internet prefixes.

As explained above there are issues caused by this BGP behaviour.

To solve these issues, we would like to propose protocol based on BGP called BGP Flex. BGP Flex will create two different routing tables like VRF concepts.

One route table will be populated with prefixes received over Internet Exchange links and the other one with prefixes received over Internet. The decision to use a particular route table will depend on some criterias that will be discussed in this document.

2. Terminology

2.1. Acronyms

1. IX : Internet Exchange
2. IXP : Internet Exchange Point
3. BLSI : BGP Local Source IP
4. BRSI : BGP Remote Source IP
5. SIRA : Source IP Router Advertiser
6. BRPI : BGP Remote Peer IP
7. CISP : Customer ISP refers the ISP buying service from an other ISP
8. SISP : Seller ISP refers to the ISP selling service to customer ISP

3. BGP Flex

As explained in the introduction, the routing issues mentionned there can be solved by BGP Flex.

BGP Flex is an extension of BGP that will create a dedicated IX route table different from the default routing table (main route table).

When BGP will advertise the routes over the IX, these routes will be marked as IX routes on the receiving router and will be inserted into the IX route table.

The decision to choose a particular routing table will depend on the classification of the packet.

This has been explained in the section 5 of this document.

The fact that IX route table is separated from Internet route table will solve routing issues and the sub-optimal routing described in abstract section, because each route table is independant of the other.

In the next section, we will define the key concepts of BGP Flex

4. BGP Flex Concepts

4.1. BGP Local Source IP (BLSI)

The BLSI refer to the public IP owned by the ISP. These IP are important as they determine how the packet will be routed, the ISP should explicitily define the BLSI on at least one router.

4.2. BGP Remote Source IP (BRSI)

In some situations the ISP announces other ISP public IP over the IXP, in this situation, the main ISP needs to declare the other ISP block as BGP Remote Source IP (BRSI).

This will make a difference between ISP own blocks and other ISP public IP.

There are two ways to define the BRSI as stated below :

1. Static mode defined locally on the router from the SISP side
2. Dynamic mode by receiving via the IP routes from CISP via BGP

In the static mode, the BRSI are mapped to CISP BGP ASN.

In the dynamic mode, when the SISP receives the routes from CISP, the SISP will flag the received routes as BRSI and maps the routes with CISP BGP ASN.

4.3. BGP Remote Peer IP (BRPI)

These are the IP received from other peers at the IXP.

4.4. Source IP Router Advertiser (SIRA)

Up to now, we have seen how to define the BLSI and BRSI, but what if they are more routers in ISP core backbone ?

This is where the Source IP Router Advertiser (SIRA) plays a role, the BLSI and BRSI will be configured on the SIRA which will advertise the BLSI and BRSI to other routers in the ISP core backbone.

It is advisable to define at least two nodes as SIRA for redundancy purpose.

BGP will be extended to support the advertisement of BLSI and BRSI. Note that routes received by SIRA are not used for routing, they are only used to identify the BLSI and BRSI required by BGP Flex for routing.

5. Routing mode in BGP Flex

In this section, we will discuss how to classify a packet received by a router.

For BGP Flex to know how to route the packets, BGP needs to classify the packets.

There are two main categories : IX packets and Internet packets.

The IX packets refer to the IP packet with source IP from BLSI, BRSI or BRPI described in the BGP Flex concepts section.

Any other packets will be treated as Internet packets.

If the packet is classified as IX then IX route table will be checked first, if there is a match the packet is forwarded using the IX routing table otherwise the packet will be routed based on the Internet table .

If the packet is not classified as IX then the packet will be routed based on the Internet route table, in this case only the Internet routing table is checked.

To summarize if the packet is classified as IX packet, the routes tables will be checked following an order (IX routing table first, then Internet routing table).

If the packet is not IX packet it is an Internet packet, in this case only Internet routing table is used.

6. BGP vs BGP Flex

Let us take the issue described in the Abstract section regarding the return traffic from Internet to ISP B (CISP) via ISP A (SISP) to see how BGP Flex will solve the problem .

It is important to note that ISP A will receive ISP B IP blocks in both route table (IX and default), because they are peering for IP Transit and Internet Exchange.

In the IX route table the next-hop to ISP B will be the Internet Exchange link, while in the default route it will be the IP Transit link.

One of the key aspect with BGP Flex is to classify the packets as either IX or Internet packet.

Here, the return traffic from Internet will have a source IP not part of BLSI, BRSI, BRPI, then the packet will be classified as Internet. As per BGP Flex rule, any packet classified as Internet should be routed based on the default route table (main route table), then the return traffic will use the IP Transit link instead of the Internet Exchange link.

BGP Flex will solve the sub-optimal routing.

When ISP A advertised most specific routes over Internet, the peers of ISP A like an ISP called ISP B at the Internet Exchange might receive these most specific routes, but the routes will be received in the Internet route table.

ISP B will have most specific routes in the Internet route table and less specific routes in the IX route table.

Because the traffic to ISP A from ISP B will be classified as IX packet, the IX route table will be used to route the packets even if it does not have the most specific routes.

7. IANA Considerations

BGP Flex supports both IPv4 and IPV6, it will require two SAFI

1. specific SAFI to identify the Network Layer Reachability Information (NLRI) exchanged via IX
2. specific SAFI to identify the BLSI and BRSI advertised by SIRA

8. Security Considerations

BGP Flex is an extension of BGP, so all security considerations for BGP can be applied.

It also advisable to apply all best practices at an Internet Exchange Point

9. References

9.1. Normative References

- [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>>.

Author's Address

Internet-Draft

Abbreviated Title

February 2026

ZOUNDI Bonaventure (editor)
Abidjan
Côte d'Ivoire
Phone: +2250586757570
Email: zoundibona@yahoo.fr