

Inter-Domain Routing
Internet-Draft
Intended status: Standards Track
Expires: 19 November 2026

S. Krier, Ed.
J. Horn
Cisco Systems
M. Ciurea
Swisscom AG
J. Tantsura
Nvidia
K. Patel
Arrcus, Inc.
18 May 2026

BGP Unreachable Prefix Announcement (UPA)
draft-krierhorn-idr-upa-02

Abstract

Summarization is often used in multi-domain networks to improve network efficiency and scalability. With summarization in place, there is a need to signal loss of reachability to an individual prefix covered by the summary. This enables fast convergence by steering traffic away from the node which owns the prefix and is no longer reachable.

This document specifies the mechanism, referred to as Unreachable Prefix Announcement (UPA), for networks where BGP is used to carry summary routes. It is also equally beneficial for operators to share the unreachable prefixes.

About This Document

This note is to be removed before publishing as an RFC.

Status information for this document may be found at
<https://datatracker.ietf.org/doc/draft-krierhorn-idr-upa/>.

Discussion of this document takes place on the Inter-Domain Routing Working Group mailing list (<mailto:idr@ietf.org>), which is archived at <https://mailarchive.ietf.org/arch/browse/idr/>. Subscribe at <https://www.ietf.org/mailman/listinfo/idr/>.

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 19 November 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. Conventions and Definitions	3
3. Terminology	3
4. Applicability	4
5. BGP UPA Signaling	4
5.1. UPA Extended Community	4
6. Trigger for UPA Origination in BGP	5
6.1. Scenario A: IGP Redistribution of Summary into BGP	6
6.2. Scenario B: BGP Aggregation/Summarization	6
7. UPA Origination in BGP	6
8. UPA Propagation in BGP	6
9. UPA Processing in BGP	7
10. Backwards Compatibility	7
11. Operational Considerations	7
12. Security Considerations	8
13. IANA Considerations	8
13.1. BGP Transitive Opaque Extended Community Sub-Type	8
13.2. UPA Flags Registry	9
14. References	9
14.1. Normative References	9

14.2. Informative References	9
Acknowledgments	10
Contributors	10
Authors' Addresses	10

1. Introduction

In modern networks, route summarization is a common practice to reduce routing table size and improve scalability. However, summarization can mask the loss of reachability of specific prefixes covered by the summary route, leading to slower convergence times.

To address this, an Unreachable Prefix Announcement (UPA) mechanism [RFC9929] has been specified for Interior Gateway Protocols (IGPs) to explicitly signal the loss of specific prefixes, enabling fast convergence mechanisms like BGP Prefix Independent Convergence (PIC) [I-D.ietf-rtgwg-bgp-pic] on ingress devices.

This document proposes a similar UPA mechanism for BGP. In multi-AS networks where IGP is not running end-to-end, a BGP-based UPA is crucial. It ensures that the loss of reachability for a specific prefix which might be part of a summarized route, can be quickly signaled across AS boundaries, thereby enabling fast convergence without compromising on the scaling benefits of route summarization.

2. Conventions and Definitions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Terminology

- * UPA: Unreachable Prefix Announcement.
- * BGP PIC: BGP Prefix Independent Convergence.
- * PE: Provider Edge router.
- * AS: Autonomous System.
- * RIB: Routing Information Base.
- * MP_REACH: Multiprotocol Reachable NLRI.
- * MP_UNREACH: Multiprotocol Unreachable NLRI.

- * ExtCom: Extended Community.
- * AFI: Address Family Identifier.
- * SAFI: Subsequent Address Family Identifier.

4. Applicability

BGP UPA applies to any operator network where BGP carries reachability information and summarization is performed. When a specific prefix within a summary becomes unreachable, the UPA mechanism is needed to signal this loss of reachability across the network to edge or leaf routers to trigger fast reconvergence (e.g., to trigger BGP-PIC at the ingress PEs).

A typical deployment scenario is a multi-AS network where BGP is used to carry IP Prefixes using either AFI=1/2, SAFI=1. However, the mechanism described in this document is equally applicable to any BGP address family that uses route summarization, such as BGP CAR for SRv6 [RFC9871] or Ethernet VPN (EVPN) Route Type 5 [RFC9136].

The BGP UPA mechanism is intended to be enabled and deployed in a network under the administration of a single operator or cooperative operators (either a single AS or multi-AS).

5. BGP UPA Signaling

A BGP UPA is signaled as a BGP UPDATE used to indicate the loss of reachability of a specific prefix.

The specific prefix whose reachability is lost is encoded in the MP_REACH_NLRI attribute [RFC4760]. The next-hop is set following standard BGP procedures.

The UPA Extended Community (as defined in Section 5.1) MUST be attached to the route to identify it as a UPA.

An Update message carrying a UPA MUST only contain UPA prefixes (i.e., no other reachability advertisements or withdrawals) due to the presence of the UPA Extended Community.

To withdraw a previously announced UPA, a BGP speaker sends an MP_UNREACH_NLRI [RFC4760] for the UPA prefix.

5.1. UPA Extended Community

A new Transitive Opaque Extended Community is defined for UPA.

The structure of this Extended Community is as follows:

```

      0               1               2               3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|      Type      | Sub-Type  |      Flags   |      Reserved   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     BGP Router ID                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

- * Type Field: 0x03 (Transitive Opaque Extended Community, per [RFC7153]).
- * Sub-Type Field: TBD (assigned by IANA).
- * Flags Field (1 byte): See below.
- * Reserved (1 byte): MUST be set to zero on transmission and ignored on receipt.
- * BGP Router ID (4 bytes): This field carries the BGP Router-ID of the node originating the UPA in BGP. This is helpful for the identification of the originator in a multi-domain network since the deployment is intended within a single administrative domain. It is assumed that BGP Router-IDs are unique within the operator's managed ASes. It does not influence route selection or forwarding behavior.

The Flags field is 8 bits in length and is defined as follows:

```

    0 1 2 3 4 5 6 7
+-----+-----+
|D|   Reserved   |
+-----+-----+

```

- * Bit 0 (D bit): If set, the forwarding entry corresponding to the route SHOULD be set to drop the traffic received for it. The receiver MAY skip installing such a forwarding entry due to local policy or when it is a route-reflector that is not in the forwarding path.
- * Bits 1-7: These bits MUST be set to zero and ignored on receipt.

6. Trigger for UPA Origination in BGP

UPA origination in BGP can be triggered by multiple scenarios and following are two scenarios:

6.1. Scenario A: IGP Redistribution of Summary into BGP

When an IGP summary route is redistributed into BGP, and a specific component prefix within that summary loses reachability in the IGP, the UPA indication is conveyed from IGP to BGP. The details of this mechanism is implementation specific and outside the scope of this document.

6.2. Scenario B: BGP Aggregation/Summarization

When BGP itself is performing summarization, and a constituent specific route goes away, the UPA is triggered internally within BGP.

Implementations SHOULD provide a configurable option to specify which types of specific prefixes trigger UPA (e.g., only /48 prefixes for SRv6 locators).

7. UPA Origination in BGP

UPA origination is triggered by BGP only in the absence of a valid reachable route for that specific prefix. The origination of UPA indication involves the update generation of the BGP UPA message as specified in Section 5.

The UPA route MUST be withdrawn when the specific prefix becomes reachable. The originator MAY withdraw the UPA route before the specific prefix becomes reachable if the use of the UPA in the deployment was only to serve as an event notification of unreachability that does not require the unreachability state to be maintained until that prefix becomes reachable.

8. UPA Propagation in BGP

The propagation of UPA messages in BGP follows the same principles as UPA origination. BGP speakers receiving a UPA will process it (refer to Section 9) and propagate it to their peers as appropriate.

When a BGP speaker receives routes with the UPA ExtCom from multiple peers, it MUST aggregate the UPA ExtComs from all UPA routes when propagating the UPA message. This ensures that the receiver gets all the originators of the UPA when this is being triggered from multiple routers.

9. UPA Processing in BGP

A BGP speaker processes UPA messages only for those prefixes for which it does not have a valid reachable route. A route carrying the UPA Extended Community is not considered a valid reachable route for this purpose. If a valid reachable route for the prefix is subsequently received, it **MUST** take precedence over the UPA route.

In the absence of a valid reachable route for a prefix, the UPA route for that prefix becomes the best route. However, since it is not a reachable route, it **MUST NOT** be considered as valid for forwarding traffic or redistribution into other protocols except as unreachable. A forwarding entry **SHOULD** be programmed with indication to drop traffic matching that prefix, if the UPA ExtCom has the drop D-bit flag set.

The processing of UPA message on certain BGP speakers would also involve notification of unreachability depending on the deployment use case. The details of this mechanism are outside the scope of this document.

10. Backwards Compatibility

A BGP speaker that does not understand the UPA Extended Community will treat the UPA as a standard route advertisement for the prefix. Due to longest-prefix-match, this may attract traffic toward a next-hop that cannot deliver it to the unreachable destination. However, this would not be significantly different than when the forwarding happens using the summary. This can be addressed by mechanisms in Section 11.

Also, the UPA ExtCom being transitive will be carried as unknown ExtCom across BGP Speakers which do not understand it and get treated correctly at routers that understand UPA. To prevent this, a BGP speaker **MUST** only send UPA messages to peers that are known to support UPA processing.

Implementations **SHOULD** provide a configuration knob to enable UPA propagation to specific neighbors. The default **MUST** be to not propagate UPA messages. This ensures that UPA propagation can be limited to the desired domain or network boundary.

11. Operational Considerations

By default, UPA origination **MUST** be disabled. Implementations **SHOULD** provide a configurable option to enable UPA origination on a per-address-family basis.

By default, the propagation of UPA MUST be disabled across AS boundaries. Implementations SHOULD provide a configuration knob to enable UPA propagation to specific neighbors. This ensures that UPA propagation can be limited to the desired domain or network boundary.

Implementations SHOULD provide configuration to allow operators to enforce limits on the number of UPA routes that can be originated at any given time. This prevents excessive UPA generation during large-scale failures which could overwhelm BGP speakers and degrade network stability. The specific limits are implementation-defined and SHOULD be configurable by the operator.

12. Security Considerations

The primary security consideration relates to the potential for leakage of internal infrastructure details into the public Internet if filtering route policies are misconfigured. The explicit signaling of unreachable prefixes via UPA could reveal more granular internal network topology information if not properly contained.

Operators SHOULD ensure robust filtering policies are in place at AS boundaries. The operational considerations in this document can serve as a mitigation strategy to limit the scope of UPA messages to trusted domains.

13. IANA Considerations

This document requests IANA to perform the following actions.

13.1. BGP Transitive Opaque Extended Community Sub-Type

IANA is requested to assign a Sub-Type for the "UPA Extended Community" (defined in this document) from the "BGP Extended Communities" registry, specifically within the "Transitive Opaque Extended Community Sub-Types" sub-registry.

Sub-Type Value	Description	Reference
TBD	UPA Extended Community	[This-Doc]

Table 1: New Transitive Opaque Extended Community Sub-Type

Note to IANA: The assignment should be made from the First Come First Served (FCFS) range (0x00-0xBF) as per [RFC7153].

13.2. UPA Flags Registry

IANA is requested to create a new registry for the 8-bit "UPA Flags" field of the UPA Extended Community.

The registration policy for this registry is "IETF Review" or "Expert Review" as per [RFC8126].

Bit	Description	Reference
0	D-bit (Drop bit)	[This-Doc]
1-7	Unassigned	

Table 2: UPA Flags

14. References

14.1. 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/rfc/rfc2119>>.
- [RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, DOI 10.17487/RFC4760, January 2007, <<https://www.rfc-editor.org/rfc/rfc4760>>.
- [RFC7153] Rosen, E., "IANA Registries for BGP Extended Communities", RFC 7153, DOI 10.17487/RFC7153, January 2014, <<https://www.rfc-editor.org/info/rfc7153>>.
- [RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", RFC 8126, DOI 10.17487/RFC8126, June 2017, <<https://www.rfc-editor.org/info/rfc8126>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/rfc/rfc8174>>.

14.2. Informative References

[I-D.ietf-rtgwg-bgp-pic]

Bashandy, A., Filsfils, C., and P. Mohapatra, "BGP Prefix Independent Convergence", Work in Progress, Internet-Draft, draft-ietf-rtgwg-bgp-pic-22, 20 April 2025, <<https://datatracker.ietf.org/doc/html/draft-ietf-rtgwg-bgp-pic-22>>.

[RFC9136] Rabadan, J., Henderickx, W., Drake, J., Lin, W., and A. Sajassi, "IP Prefix Advertisement in Ethernet VPN (EVPN)", RFC 9136, DOI 10.17487/RFC9136, October 2021, <<https://www.rfc-editor.org/rfc/rfc9136>>.

[RFC9871] Rao, D., Talaulikar, K., Raszuk, R., Decraene, B., and L. Jalil, "BGP Color-Aware Routing (CAR)", RFC 9871, DOI 10.17487/RFC9871, 20 November 2025, <<https://www.rfc-editor.org/rfc/rfc9871>>.

[RFC9929] Psenak, P., Filsfils, C., Voyer, D., Hegde, S., and G. S. Mishra, "IGP Unreachable Prefix Announcement", RFC 9929, DOI 10.17487/RFC9929, March 2026, <<https://www.rfc-editor.org/rfc/rfc9929>>.

Acknowledgments

The authors would like to acknowledge the contribution of Ketan Talaulikar, Clarence Filsfils for their valuable input and review of this document. The authors would like also to recognize Swadesh Agrawal and Dhananjaya Rao for the initial idea.

Contributors

Krishnaswamy Ananthamurthy
Cisco Systems
170 W. Tasman Drive
San Jose, CA, 95134
United States of America
Email: kriswamy@cisco.com

Authors' Addresses

Serge Krier (editor)
Cisco Systems
De Kleetlaan 6a
1831 Diegem
Belgium
Email: sekrier@cisco.com

Jakub Horn
Cisco Systems
Milpitas, CA 95035
United States of America
Email: jakuhorn@cisco.com

Mihai Ciurea
Swisscom AG
Alte Tiefenaustrasse 6
CH-3048 Worblaufen
Switzerland
Email: mihai.ciurea@swisscom.com

Jeff Tantsura
Nvidia
United States of America
Email: jefftant.ietf@gmail.com

Keyur Patel
Arrcus, Inc.
2077 Gateway Pl
San Jose, CA, 95110
United States of America
Email: keyur@arrcus.com