

IDR Working Group
Internet-Draft
Intended status: Standards Track
Expires: 15 May 2026

J. Tantsura
D. Sharp
V. Venkatraman
K. Muppalla
Nvidia
11 November 2025

BGP Unreachability Information SAFI
draft-tantsura-idr-unreachability-safi-00

Abstract

This document defines a new BGP Subsequent Address Family Identifier (SAFI) called "Unreachability Information" that allows the propagation of prefix unreachability information through BGP without affecting the installation or removal of routes in the Routing Information Base (RIB) or Forwarding Information Base (FIB). This mechanism enables network operators to share information about unreachable prefixes for monitoring, debugging, and coordination purposes while maintaining complete separation from the active routing plane.

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 15 May 2026.

Copyright Notice

Copyright (c) 2025 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	2
1.1. Requirements Language	3
1.2. Terminology	3
2. Protocol Extensions	3
2.1. Unreachability Information SAFI	3
2.2. Capability Advertisement	4
3. NLRI Format	5
3.1. IPv4 Unreachability NLRI (AFI=1, SAFI=86)	5
3.2. IPv6 Unreachability NLRI (AFI=2, SAFI=86)	5
3.3. TLV Format	5
4. Operation	6
4.1. NLRI Processing	6
4.2. Interaction with Route Reflection	7
4.3. Communities and Attributes	7
4.4. Preventing State Explosion	7
5. Deployment Considerations	7
5.1. Incremental Deployment	7
5.2. Use Cases	8
6. Security Considerations	8
7. IANA Considerations	9
8. Acknowledgements	9
9. References	9
9.1. Normative References	9
9.2. Informative References	10
Appendix A. Implementation Considerations	10
Appendix B. Examples	10
B.1. Example UPDATE Message	10
B.2. Configuration Example	11
Authors' Addresses	11

1. Introduction

BGP-4 [RFC4271] withdrawals are only propagated for prefixes that have been previously announced. This behavior, while preventing certain attack vectors, limits the ability of operators to share information about prefix unreachability for prefixes that were never announced in the first place.

There are several use cases where propagating unreachability information without affecting routing decisions would be valuable:

- * Debugging and troubleshooting routing issues across administrative domains
- * Sharing information about DDoS targets without null-routing traffic
- * Coordinating information about potentially hijacked prefixes
- * Monitoring and anomaly detection systems that need visibility into negative routing events
- * Providing telemetry about routing system health without affecting production traffic

This document defines a new SAFI that creates a parallel information plane for unreachability data, allowing BGP speakers to share this information while maintaining complete separation from the routing plane.

1.1. Requirements Language

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.

1.2. Terminology

UI-RIB: Unreachability Information RIB

NLRI: Network Layer Reachability Information

AFI: Address Family Identifier

SAFI: Subsequent Address Family Identifier

TLV: Type-Length-Value

2. Protocol Extensions

2.1. Unreachability Information SAFI

This document defines a new SAFI:

- * Value: 86 (to be assigned by IANA)
- * Name: Unreachability Information (UNREACH)

- * Applicable to AFI: 1 (IPv4) and 2 (IPv6)

2.2. Capability Advertisement

A BGP speaker that wishes to exchange Unreachability Information MUST advertise the corresponding AFI/SAFI capability as defined in [RFC5492].

The Capability Code for Multiprotocol Extensions is 1. The Capability Value field contains:

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	2	3	4	5	6	7	8	9
AFI										Reserved										SAFI																			

Where:

- * AFI = 1 (IPv4) or 2 (IPv6)
- * Reserved = 0 (MUST be set to 0 on transmit, ignored on receive)
- * SAFI = 86 (Unreachability Information)

Additionally, this document defines a new capability:

- * Capability Code: 86 (to be assigned by IANA)
- * Capability Name: Enhanced Unreachability Information
- * Capability Value: 1 octet flags field

0	1	2	3	4	5	6	7
T R				Reserved			

Where:

- * T bit: If set, speaker supports Timestamp TLV
- * R bit: If set, speaker supports Reason Code TLV
- * Reserved: MUST be zero on transmit, ignored on receive

3. NLRI Format

The NLRI is uniquely identified by the combination of Prefix Length and Prefix. TLVs are NOT part of the NLRI key and provide additional context about the unreachability information. The presence of an Unreachability Information NLRI for a prefix signifies that the prefix is unreachable. The withdrawal of such an NLRI indicates normal routing operation for that prefix.

3.1. IPv4 Unreachability NLRI (AFI=1, SAFI=86)

```

      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
+-----+-----+-----+-----+-----+-----+-----+-----+
| Prefix Length |           IPv4 Prefix (variable)           |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     TLVs (variable)                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

3.2. IPv6 Unreachability NLRI (AFI=2, SAFI=86)

```

      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
+-----+-----+-----+-----+-----+-----+-----+-----+
| Prefix Length |           IPv6 Prefix (variable)           |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     TLVs (variable)                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

3.3. TLV Format

TLVs provide additional context about the unreachability information:

```

      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   |           Length           |                   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Value (variable)                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Defined TLV Types:

Type 1: Original Reporter

* Length: 4 octets

* Value: BGP Identifier of the original reporting speaker

Type 2: Unreachability Reason Code

- * Length: 2 octets
- * Value: Detailed reason code (registry to be established)
- * 0: Unspecified
- * 1: Policy Blocked
- * 2: Security Filtered
- * 3: RPKI Invalid
- * 4-65535: Reserved for future use

Type 3: Timestamp

- * Length: 8 octets
- * Value: Unix timestamp (seconds since epoch) in network byte order, indicates when the unreachability event occurred or was detected

All TLV types except Type 1 (Original Reporter) are OPTIONAL. Type 1 MUST be present in all Unreachability Information NLRI. Implementations MUST ignore unknown TLV types to allow for future extensibility. Multiple TLVs of the same type SHOULD NOT be included; if present, only the first occurrence SHOULD be processed

4. Operation

4.1. NLRI Processing

When a BGP speaker receives an UPDATE message with Unreachability Information SAFI:

1. It MUST NOT install or remove any routes in the Loc-RIB based on this information
2. It MUST maintain a separate Unreachability Information RIB (UI-RIB) for this SAFI
3. It SHOULD apply standard BGP path selection to UI-RIB entries for consistency
4. It MAY propagate the information according to standard BGP rules and local policy

5. It MUST NOT mix Unreachability Information NLRI with other SAFIs in the same UPDATE message

4.2. Interaction with Route Reflection

Route Reflectors SHOULD treat Unreachability Information SAFI like any other AFI/SAFI combination, applying standard route reflection rules. The ORIGINATOR_ID and CLUSTER_LIST attributes apply normally.

4.3. Communities and Attributes

Standard BGP communities and attributes apply to Unreachability Information SAFI:

- * NO_EXPORT, NO_ADVERTISE, and NO_EXPORT_SUBCONFED work as defined
- * Large Communities [RFC8092] MAY be used for policy control
- * AS_PATH is constructed normally for loop prevention
- * ORIGIN SHOULD be set to IGP for locally generated information, EGP for information received via eBGP

4.4. Preventing State Explosion

To prevent unbounded growth of the UI-RIB:

1. Implementations SHOULD limit the number of unreachability entries per prefix
2. Implementations SHOULD age out entries based on the Timestamp TLV when present
3. Implementations MUST implement rate limiting for accepting new unreachability information
4. Default maximum UI-RIB size SHOULD be configurable with a reasonable default (e.g., 100,000 entries)

5. Deployment Considerations

5.1. Incremental Deployment

The Unreachability Information SAFI can be deployed incrementally:

- * Speakers that don't support it simply don't negotiate the capability

- * Mixed environments work correctly with normal BGP capability negotiation
- * Can be enabled on specific sessions for testing

5.2. Use Cases

Example deployment scenarios:

Inter-AS Debugging: Enable between cooperating ASes for troubleshooting

Route Collectors: Deploy on route collector sessions for enhanced telemetry

DDoS Coordination: Share attack target information without null-routing

Security Monitoring: Track suspicious unreachability patterns

6. Security Considerations

The Unreachability Information SAFI introduces new security considerations:

1. Information Leakage: Unreachability information might reveal network topology or operational issues. Operators SHOULD carefully consider peering policies for this SAFI.
2. State Exhaustion: Malicious peers could attempt to exhaust memory by advertising large numbers of unreachable prefixes. Implementations MUST enforce limits as described in Section 4.4.
3. False Information: Peers could advertise false unreachability information. This SAFI SHOULD only be enabled with trusted peers.
4. Prefix Hijacking: The SAFI could be used to claim prefixes are unreachable when they're not. Since this doesn't affect routing, the impact is limited to monitoring systems.

Operators SHOULD:

- * Use BGP TCP-AO [RFC5925] or MD5 for session protection
- * Implement prefix filtering for unreachability information
- * Monitor UI-RIB size and growth rate

- * Enable this SAFI only with explicitly trusted peers

7. IANA Considerations

IANA is requested to assign:

1. A new SAFI value (suggested: 86) for "Unreachability Information" applicable to AFI 1 and AFI 2 in the "Subsequent Address Family Identifiers (SAFI) Parameters" registry.
2. A new BGP Capability Code (suggested: 86) for "Enhanced Unreachability Information" in the "Capability Codes" registry.
3. Create a new registry called "BGP Unreachability Information TLV Types" under the "Border Gateway Protocol (BGP) Parameters" registry page. The registration procedure is "RFC Required". Initial values are defined in Section 3.3.
4. Create a new registry called "BGP Unreachability Reason Codes" under the "Border Gateway Protocol (BGP) Parameters" registry page. The registration procedure is "RFC Required". Initial values to be defined in a subsequent document.

8. Acknowledgements

The author would like to thank the IDR working group for their valuable feedback and suggestions on this proposal.

9. References

9.1. 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>>.
- [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>>.
- [RFC5492] Scudder, J. and R. Chandra, "Capabilities Advertisement with BGP-4", RFC 5492, DOI 10.17487/RFC5492, February 2009, <<https://www.rfc-editor.org/info/rfc5492>>.

- [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/info/rfc8174>>.

9.2. Informative References

- [RFC5925] Touch, J., Mankin, A., and R. Bonica, "The TCP Authentication Option", RFC 5925, DOI 10.17487/RFC5925, June 2010, <<https://www.rfc-editor.org/info/rfc5925>>.
- [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>>.

Appendix A. Implementation Considerations

Implementers should consider the following aspects when implementing the Unreachability Information SAFI:

1. UI-RIB should be stored separately from regular RIBs
2. Show commands should clearly differentiate unreachability info
3. MIB/YANG models need updating for monitoring
4. BMP [RFC7854] should be extended to support this SAFI

Appendix B. Examples

B.1. Example UPDATE Message

An UPDATE message advertising that 192.0.2.0/24 is unreachable due to RPKI validation failure:

```
* AFI: 1 (IPv4)
* SAFI: 86 (Unreachability Information)
* NLRI:
*      Prefix Length: 24
```

- * Prefix: 192.0.2.0
- * TLV Type 1 (Unreachability Type): Value 4 (RPKI Invalid)
- * TLV Type 2 (Timestamp): Current Unix timestamp
- * Path Attributes: Standard BGP attributes (AS_PATH, ORIGIN, etc.)

B.2. Configuration Example

Example BGP configuration (vendor-neutral):

```
router bgp 65001
  neighbor 192.0.2.1 remote-as 65002
  !
  address-family ipv4 unreachability
    neighbor 192.0.2.1 activate
    neighbor 192.0.2.1 maximum-prefix 50000
    neighbor 192.0.2.1 route-map UNREACH-IN in
    neighbor 192.0.2.1 route-map UNREACH-OUT out
  exit-address-family
  !
  address-family ipv6 unreachability
    neighbor 2001:db8::1 activate
  exit-address-family
  !
  route-map UNREACH-IN permit 10
    match community NO-EXPORT-UNREACH
    set local-preference 50
  !
```

Authors' Addresses

Jeff Tantsura
Nvidia
Email: jefftant.ietf@gmail.com

Donald Sharp
Nvidia
Email: sharpd@nvidia.com

Vivek Venkatraman
Nvidia
Email: vivek@nvidia.com

Karthikeya Venkat Muppalla
Nvidia
Email: kmuppalla@nvidia.com