IDR
Internet-Draft
Intended status: Standards Track
Expires: 14 August 2026

Y. Li
Y. Li
X. Yin
H. Zhang
X. Shi
Tsinghua University
10 February 2026

BGP Failure Propagation (BGP-FP) for Enhancing Control-Plane Convergence
draft-li-idr-bgp-failure-propagation-convergence-00

Abstract

This document specifies BGP Failure Propagation (BGP-FP), an infrastructure and protocol that improves inter-domain routing convergence by accelerating the removal of stale (invalid) routes. BGP-FP uses (1) an Agent deployed per Autonomous System (AS) to detect eBGP link/policy/prefix failures and to configure local routers, (2) a logically centralized (potentially distributed) Repository to store and selectively forward failure information, and (3) BGP Large Communities as a "route freshness" marker. Agents validate and apply Repository updates to filter routes that traverse failed eBGP links or otherwise violate the originating AS' s forwarding intent, reducing route flap propagation in the control 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 14 August 2026.

Table of Contents

1. Introduction

BGP-4 [RFC4271] is a path-vector inter-domain routing protocol. Internet operators have long observed BGP instability and route-flap propagation [RFC2439]. A key contributor is that routes that have become invalid (e.g., because an eBGP link failed) can persist and propagate until withdrawals and path exploration complete.

BGP-FP addresses this by disseminating authenticated failure information and by marking routes with a per-AS freshness indicator. Upon learning a failure, downstream ASes can proactively reject stale routes that are provably older than the failure event, reducing control-plane churn and improving eBGP convergence.

This document focuses on eBGP control-plane convergence. Terminology related to convergence benchmarking is aligned with [RFC4098].

2. Conventions and Definitions

2.1. Terminology

eBGP: As defined in [RFC4271].

Agent: A server deployed per AS that (1) receives router logs, (2) reads local routing state (e.g., RIB/Adj-RIB-In/Adj-RIB-Out), (3) communicates with a Repository using BGP-FP messages, and (4) configures local routers' import/export policies to apply BGP-FP.

Repository: A network service with a well-known reachability method (discovery is out of scope) that stores per-AS BGP-FP state and forwards updates on-demand to subscribed ASes.

BGP Link: An eBGP adjacency between two different ASes (i.e., an inter-AS BGP session and its corresponding forwarding relationship).

Invalid (Stale) BGP Route: A route whose AS_PATH includes an eBGP link that is no longer usable (explicitly or implicitly withdrawn), or whose propagation violates a policy/prefix condition expressed by the originating AS. For example, the eBGP relationship between AS X and AS Y is down, but a route' s AS_PATH still contains "... X Y ..."; such a route is considered stale and is expected to be withdrawn or replaced.

Local AS: The AS in which a given Agent is deployed.

Subscribed ASes: The set of ASNs that appear in AS_PATH attributes of routes in the Local AS routing state (e.g., Loc-RIB). The Local AS subscribes to BGP-FP updates originating from those ASes.

NOTE: The terms "route", "BGP route", "AS_PATH", "BGP Neighbor", "BGP Peer", "MRAI", and "RIB" are used consistent with [RFC4098] where applicable.

2.2. 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.

3. Architecture and Operational Overview

3.1. Components

BGP-FP consists of:

* Agent (per AS): detects local failures, maintains a per-AS freshness counter, reports state to the Repository, and configures routers to (a) attach freshness markers to outbound routes and (b) filter stale routes upon receiving Repository updates.

* Repository: stores per-AS reported state and forwards BGP-FP updates to ASes that have subscribed to the source AS.

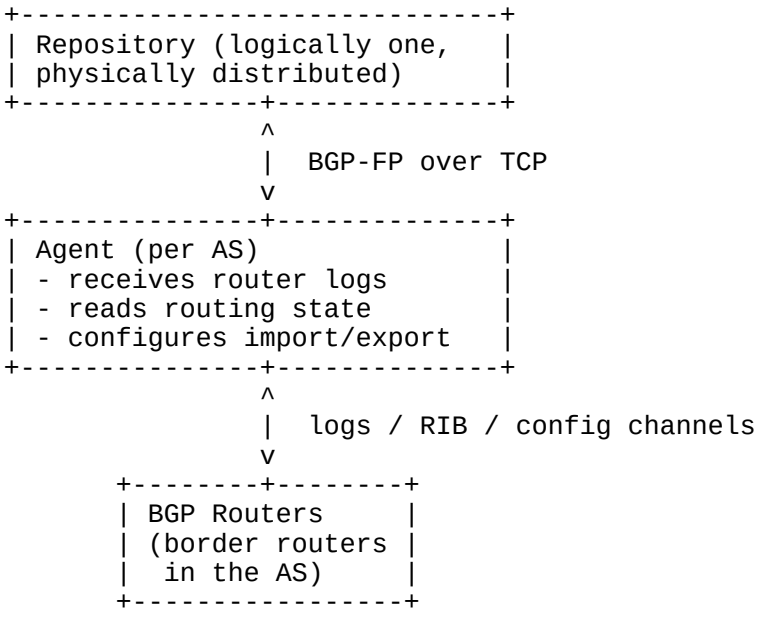


Figure 1: The Architecture of BGP-FP

3.2. Information Model: Link-State Entries

An Agent reports "BGP Link State" information as a list of entries. Each entry is a UTF-8 string formatted as: "peer-asn ":" scope ":" status" where:

* peer-asn: The 4-octet ASN of the eBGP neighbor (decimal).

* scope: Either "*" or a prefix/identifier that indicates the affected traffic scope. This document uses "prefix" in the broad sense and permits "*" to indicate all prefixes.

* status: "Established" or "Failure".

The following classes are distinguished (using the same textual form):

* (1) Policy failure (prefix scoped to a peer): "PeerASN:Prefix:Failure" The Local AS indicates that it will not forward the indicated prefix toward PeerASN.

* (2) Link failure (all prefixes to a peer): "PeerASN:*:Failure" The Local AS indicates that its eBGP link/forwarding relationship with PeerASN is not usable for any prefixes.

* (3) Prefix withdrawal/denial (all peers for a prefix): "*:Prefix:Failure" The Local AS indicates that it will not accept or propagate the indicated prefix (or has withdrawn it).

The semantics above are used by downstream Agents to derive filtering rules for stale routes (Section 8).

3.3. Route Freshness Marking with BGP Large Communities

BGP-FP uses the BGP Large Communities attribute [RFC8092]. Operational guidance for organizing Large Communities is discussed in [RFC8195].

Each AS maintains a monotonically increasing 32-bit Batch Counter. The Agent configures routers to attach a Large Community of the form.

RFC8092	this document
Global Administrator	Source ASN
Local Data Part 1	Function
Local Data Part 2	Batch Counter

Table 1: Field Mapping

The table above shows a mapping table between the fields in BGP Large Communities [RFC8092] and this document.

* Source ASN is the AS performing the marking (the Local AS).

* Function is an operator-chosen 32-bit value; this document uses 1234 as a default example.

* Batch Counter is the current freshness value for the Local AS.

A route MAY carry multiple Large Communities. BGP-FP requires that routers preserve existing Large Communities and that each AS that deploys BGP-FP adds (at least) its own freshness community.

4. Protocol Overview

BGP-FP defines a simple message protocol carried over TCP [RFC793] between Agents and the Repository. The protocol MUST support 4-octet ASNs as specified in [RFC6793]. High-level operation:

* Subscription: Each Agent computes its Subscribed ASes from local routing state and reports subscription deltas to the Repository.

* Failure reporting: Upon local events (eBGP down/up, policy change, prefix withdrawal/restore), an Agent increments its Batch Counter and reports updated Link-State Entries to the Repository.

* On-demand forwarding: The Repository forwards Source-AS updates only to Agents whose Local AS has subscribed to that Source ASN.

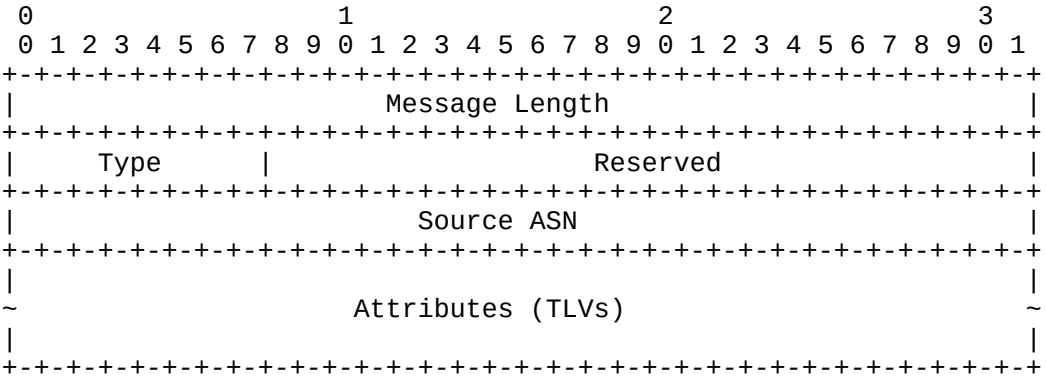
* Filtering: Receiving Agents validate updates and install import policy rules to reject stale routes that traverse failed links or violate reported policy/prefix constraints.

5. BGP-FP Message Encoding

All multi-octet fields are in network byte order (big-endian).

5.1. Common Header

Each BGP-FP message begins with the following header:



* Message Length (32 bits): Number of octets following this field (i.e., header remainder and TLVs). Implementations MUST reject messages whose length exceeds the remaining TCP segment stream availability.

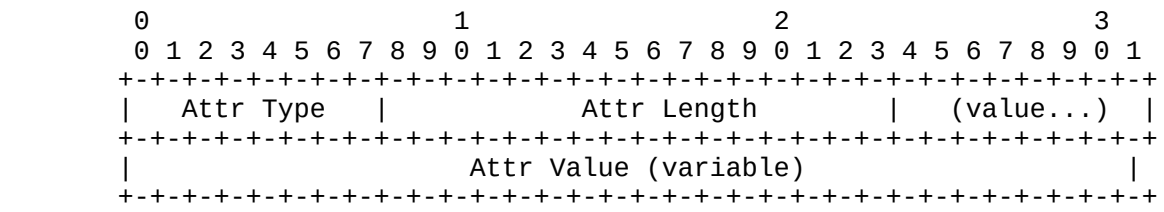
* Type (8 bits): Message type (Agent-to-Repository in Section 6; Repository-to-Agent in Section 7).

* Reserved (24 bits): Reserved for future use. The sender MUST set these bits to zero. The receiver MUST ignore them.

* Source ASN (32 bits): The originating AS for the state carried in the messag

5.2. Attributes (TLV)

Following the common header is a sequence of zero or more attributes, each encoded as a TLV:



* Attr Type (8 bits): Attribute identifier.

* Attr Length (16 bits): Length in octets of Attr Value.

* Attr Value (variable): Attribute content. Unknown Attr Type values MUST be ignored (skipped using Attr Length).

5.3. Defined Attributes

This document defines the following attribute types:

* 0x01 Batch Counter

* 0x02 Function

* 0x03 Merkle Root

* 0x04 Signature

* 0x05 BGP Link State List

* 0x06 Subscribed ASes

* 0x07 Merkle Proof

5.3.1. Batch Counter (0x01)

A 32-bit unsigned integer. Each Agent maintains its own counter. When an Agent reports any failure-affecting change (including link and policy/prefix events), it MUST increment the counter by 1.

5.3.2. Function (0x02)

A 32-bit unsigned integer chosen by operator policy. The default example value is 1234.

5.3.3. Merkle Root (0x03)

A 32-octet value computed using SHA-256 over the current dataset of (a) BGP Link State List and (b) Subscribed ASes. The Merkle tree and inclusion proof model is aligned with the Merkle tree approach used in Certificate Transparency [RFC9162]. The exact leaf canonicalization is to be specified in a future revision.

5.3.4. Signature (0x04)

A digital signature computed by the Source ASN's private key over the Merkle Root value. The signature algorithm and encoding follow the DigitallySigned conventions used in [RFC9162], unless otherwise profiled by an implementation profile. The Signature attribute MUST NOT appear unless a Merkle Root attribute is present in the same message.

5.3.5. BGP Link State List (0x05)

Encodes a list of link-state entries:

* Entry Count (16 bits), followed by Entry Count entries: Entry Length (16 bits) + Entry (UTF-8 string)

* Each Entry string MUST follow this ABNF:

```
link-state-entry = peer-asn ":" scope ":" status
peer-asn = 1*10DIGIT  ; decimal ASN, MUST fit in 32 bits
scope    = "*" / prefix
status   = "Established" / "Failure"
```

The "prefix" syntax is deployment-specific; implementations SHOULD use CIDR textual forms.

ABNF is specified per [RFC5234].

5.3.6. Subscribed ASes (0x06)

Carries two ASN lists:

* Add List: ASNs newly subscribed

* Remove List: ASNs unsubscribed

Encoding:

* Add Count (16 bits), followed by Add Count ASNs (each 32 bits)

* Remove Count (16 bits), followed by Remove Count ASNs (each 32 bits)

For an INIT message, Remove Count MUST be zero. For INIT, Add List SHOULD be the set of all ASNs appearing in AS_PATH attributes of BGP routes in the Local AS routing state.

5.3.7. Merkle Proof (0x07)

Provides hashes necessary to verify inclusion of disseminated data, aligned with [RFC9162]. Encoding:

* Proof Count (16 bits), followed by Proof Count tuples

* Each tuple: ASN (32 bits) + Hash (32 octets)

6. Agent-to-Repository Protocol

6.1. Message Types

Agent messages use the Type field as follows:

* Type = 0 KEEPALIVE Message

* Type = 1 INIT Message

* Type = 2 UPDATE Message

INIT is sent when an Agent is first deployed and informs the Repository that the Local AS participates in BGP-FP. INIT conveys a full snapshot. UPDATE conveys incremental changes after a successful INIT.

6.2. Mandatory/Optional Attributes

For Agent messages:

* KEEPALIVE (Type=0): no attributes.

* INIT (Type=1): MUST include: Batch Counter (0x01), Function (0x02), Merkle Root (0x03), Signature (0x04) MAY include: BGP Link State List (0x05), Subscribed ASes (0x06)

* UPDATE (Type=2): MUST include: Batch Counter (0x01), Merkle Root (0x03), Signature (0x04) MAY include: Function (0x02), BGP Link State List (0x05), Subscribed ASes (0x06)

6.3. Agent Procedures

Initialization:

* Set Batch Counter to 0.
* Choose Function (default example: 1234).
* Collect current BGP Link State from logs.
* Collect Subscribed ASes from routing state.
* Configure outbound marking (Section 8.2).
* Send INIT with required attributes.

Subscription change:

* Compute Add/Remove deltas.
* Batch Counter SHOULD NOT change for subscription-only updates.
* Send UPDATE carrying Subscribed ASes.

Link/policy/prefix events:

* Upon detecting a failure or restoration, Agent MUST increment the Batch Counter by 1 and MUST refresh outbound marking.
* Send UPDATE carrying affected Link State List entries.

7. Repository-to-Agent Protocol

7.1. Message Types

Repository messages use the Type field as follows:

* Type = 0 KEEPALIVE
* Type = 1 UPDATE

A Repository UPDATE broadcasts the BGP-FP state of a given Source ASN to ASes that have subscribed to that Source ASN.

7.2. Mandatory/Optional Attributes

For Repository messages:

KEEPALIVE (Type=0): no attributes.

UPDATE (Type=1): MUST include: Batch Counter (0x01), Function (0x02), Merkle Root (0x03), Signature (0x04), Merkle Proof (0x07) MAY include: BGP Link State List (0x05), Subscribed ASes (0x06)

The Repository MUST forward the Source ASN's Signature without modification.

7.3. Repository Procedures

Validation: Repository SHOULD validate Agent messages by verifying the signature over the Merkle Root and by recomputing the Merkle Root from the carried dataset.

Storage and maintenance: Maintain per-Source-ASN state: R_BatchCounter, R_Function, R_MerkleRoot, R_Signature, R_BGP_Link_State, and subscription mappings.

On-demand forwarding: Forward Source-ASN updates only to Local ASes that have subscribed to that Source ASN. Repository UPDATE messages MUST include a Merkle Proof sufficient for the receiver to validate inclusion of the forwarded elements.

8. Router Policy Behavior

8.1. Inbound Filtering Rule

Upon receiving a Repository UPDATE, an Agent MUST:

1. Obtain the Source ASN public key via the configured PKI mechanism (Section 9), verify Signature over Merkle Root, and verify the Merkle Proof.

2. For each failure entry affecting (Source ASN, Peer ASN, scope), install an import policy rule on routers such that a route r is rejected if ALL of the following hold:

* r's AS_PATH contains the adjacent pair "... SourceASN PeerASN ..." (i.e., r traverses the failed eBGP link), OR the failure scope indicates a prefix withdrawal that matches r.

* r contains a BGP Large Community whose Global Administrator is SourceASN and whose Local Data Part 1 matches Function.

* Let r_bc be Local Data Part 2 of that community. The route is considered stale if r_bc < BatchCounter carried in the Repository UPDATE for that Source ASN.

Filtering MUST be applied to Adj-RIB-In (or equivalently as close as possible to route import) so that the BGP decision process does not select stale routes.

8.2. Outbound Marking Rule

Routers in a BGP-FP-enabled AS MUST attach the Large Community (SourceASN, Function, BatchCounter) to all outbound eBGP announcements. The Large Communities attribute is an optional transitive attribute [RFC8092] ;implementations MUST NOT strip communities unrelated to the Local AS' s own marking.

When BatchCounter changes, the Agent MUST refresh the export policy so subsequent updates carry the new freshness value.

9. Privacy Considerations

Subscribed ASes may reveal portions of an AS' s observed AS_PATHs and thus aspects of routing visibility. Deployments SHOULD minimize unnecessary subscription disclosure (e.g., aggregate subscriptions, policy-based suppression) and secure transport and storage.

10. Operational Considerations

* Repository discovery, redundancy, and anycast/distribution are out of scope, but are expected to be critical for availability.

* Operators should carefully stage deployment, starting with monitoring-only mode before enforcing route rejection.

* Interaction with existing route flap damping [RFC2439] should be evaluated; BGP-FP is intended to reduce the need for aggressive damping by removing stale routes earlier.

11. IANA Considerations

Following the guidance of [RFC8126], this document requests IANA actions to support interoperable deployment:

11.1. TCP Service Name and Port

IANA is requested to allocate a TCP port for the "bgpfp" service (BGP Failure Propagation). The port number is TBD.

11.2. "BGP-FP Message Types" Registry

Create a new registry for the 8-bit Type field in the BGP-FP common header.

Initial allocations:

* 0 KEEPALIVE
* 1 INIT (Agent-to-Repository only)
* 2 UPDATE (Agent-to-Repository only)
* 3 UPDATE (Repository-to-Agent)

NOTE: If IANA prefers separate registries for Agent vs Repository message spaces, this can be revised.

Registration policy: IETF Review.

11.3. "BGP-FP Attribute Types" Registry

Create a new registry for the 8-bit Attr Type field.

Initial allocations:

* 0x01 Batch Counter
* 0x02 Function
* 0x03 Merkle Root
* 0x04 Signature
* 0x05 BGP Link State List
* 0x06 Subscribed ASes
* 0x07 Merkle Proof

Registration policy: IETF Review. Range 0x80-0xFF is RESERVED for Private Use.

12. Security Considerations

BGP-FP can cause large-scale route rejection if updates are forged or mishandled. The protocol therefore requires authenticity and integrity of Repository updates and of Source-AS-originated state.

Key requirements:

* Agents MUST verify signatures on Repository UPDATE messages.

* The system MUST provide a trustworthy mapping from ASN to public key. RPKI [RFC6480] is one candidate infrastructure for distributing authenticated resource-linked keys/certificates; the detailed certificate profile and distribution procedure are out of scope.

* Merkle proofs help receivers detect tampering with forwarded subsets of state.

Additional threats include replay of old updates, Repository compromise, denial of service against Agents/Repository, and privacy leakage via subscription information. Implementations SHOULD provide replay protection (e.g., track per-Source-ASN BatchCounter monotonicity) and rate limiting.

13. References

13.1. 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/rfc/rfc4271>.

[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/rfc/rfc8092>.

[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>.

[RFC5234] Crocker, D., Ed. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, DOI 10.17487/RFC5234, January 2008, <https://www.rfc-editor.org/rfc/rfc5234>.

[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>.

[RFC793] Postel, J., "Transmission Control Protocol", RFC 793, DOI 10.17487/RFC0793, September 1981, <https://www.rfc-editor.org/rfc/rfc793>.

[RFC6793] Vohra, Q. and E. Chen, "BGP Support for Four-Octet Autonomous System (AS) Number Space", RFC 6793, DOI 10.17487/RFC6793, December 2012, <https://www.rfc-editor.org/rfc/rfc6793>.

[RFC9162] Laurie, B., Messeri, E., and R. Stradling, "Certificate Transparency Version 2.0", RFC 9162, DOI 10.17487/RFC9162, December 2021, <https://www.rfc-editor.org/rfc/rfc9162>.

[RFC6480] Lepinski, M. and S. Kent, "An Infrastructure to Support Secure Internet Routing", RFC 6480, DOI 10.17487/RFC6480, February 2012, <https://www.rfc-editor.org/rfc/rfc6480>.

13.2. Informative References

[RFC2439] Villamizar, C., Chandra, R., and R. Govindan, "BGP Route Flap Damping", RFC 2439, DOI 10.17487/RFC2439, November 1998, <https://www.rfc-editor.org/rfc/rfc2439>.

[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/rfc/rfc8195>.

[RFC4098] Berkowitz, H., Davies, E., Ed., Hares, S., Krishnaswamy, P., and M. Lepp, "Terminology for Benchmarking BGP Device Convergence in the Control Plane", RFC 4098, DOI 10.17487/RFC4098, June 2005, <https://www.rfc-editor.org/rfc/rfc4098>.

[RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <https://www.rfc-editor.org/rfc/rfc8126>.

Authors' Addresses

Yuhang Li
Tsinghua University
Email: yh-li24@mails.tsinghua.edu.cn

Yahui Li
Tsinghua University
Email: liyahui@tsinghua.edu.cn

Xia Yin
Tsinghua University
Email: yxia@tsinghua.edu.cn

Han Zhang
Tsinghua University
Email: zhhan@tsinghua.edu.cn

Xingang Shi
Tsinghua University
Email: shixg@cernet.edu.cn