IDR
Internet-Draft
Intended status: Standards Track
Expires: 1 September 2026

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

BGP Failure Propagation (BGP-FP) for Enhancing Control-Plane Convergence
draft-li-idr-bgp-failure-propagation-convergence-01

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 inter-AS reachability changes and to configure local routers, (2) a logically centralized Repository to store and selectively forward AS reachability state, and (3) BGP Large Communities as a "route freshness" marker. Agents validate and apply Repository updates to filter routes that traverse AS pairs whose reachability has been lost or that violate the originating AS’s forwarding intent, reducing route-flap propagation in the control plane.

This document clarifies that "AS reachability" refers to the reachability between two ASes, not to the state of individual physical or logical links within an AS. If multiple links exist between two ASes, the failure of a single link that does not break overall AS-to-AS reachability does not trigger the BGP-FP mechanism.

A new Repository deployment model is introduced, suggesting that the Repository be operated by a newly established organization composed of Tier-1 ASes and Regional Internet Registries (RIRs), using a distributed deployment with Byzantine fault-tolerant consensus and rotating leadership.

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/.

Copyright Notice

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 the reachability between two ASes is lost) can persist and propagate until withdrawals and path exploration complete.

BGP-FP addresses this by disseminating authenticated AS reachability changes and by marking routes with a per-AS freshness indicator. Upon learning that reachability between AS X and AS Y has been lost, downstream ASes can proactively reject stale routes whose AS_PATH traverses that AS pair, 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].

"AS reachability" in this document refers to the ability of one AS to reach another AS via at least one eBGP session, not to the state of individual links or routers inside an AS. If multiple parallel links or sessions exist between two ASes, the failure of a single component that does not break the overall AS-to-AS reachability does not trigger the BGP-FP mechanisms.

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.

AS Reachability: The property that two ASes have at least one usable eBGP session and forwarding relationship between them. It is a pairwise AS-level property, not a link-level property.

AS Reachability State: The status of reachability between a pair of ASes, represented as a tuple (Local ASN, Peer ASN, Scope, Status). Scope and Status are defined in Section 3.2.

Invalid (Stale) BGP Route: A route whose AS_PATH includes an AS pair whose reachability has been lost, or whose propagation violates a policy/prefix condition expressed by the originating AS. For example, if the AS reachability between AS X and AS Y is down, but a route's AS_PATH still contains "... X Y ...", the 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 AS reachability changes, 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 AS reachability state and forwards BGP-FP updates to ASes that have subscribed to the source AS. The Repository is logically centralized but can be physically distributed.

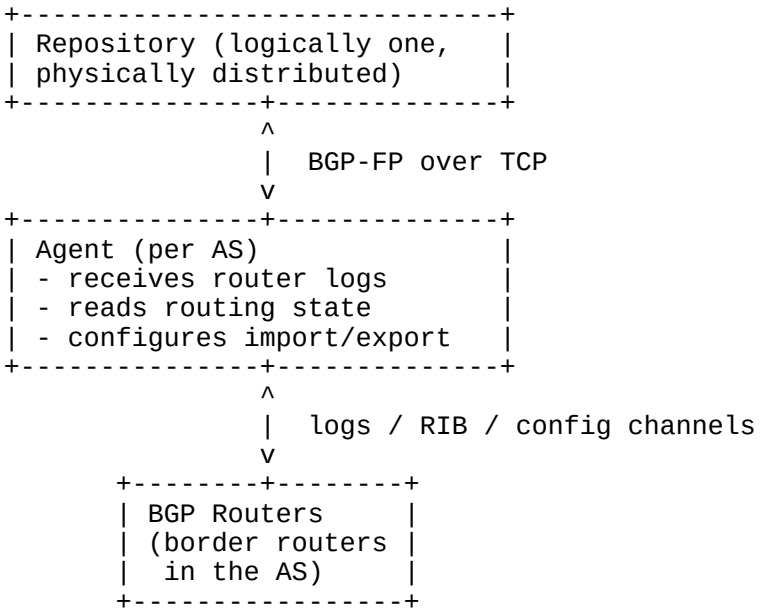


Figure 1: The Architecture of BGP-FP

3.2. Information Model: AS Reachability State Entries

An Agent reports "AS Reachability 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.

* 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:

* (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) AS reachability failure (all prefixes to a peer): "PeerASN:*:Failure" . The Local AS indicates that its overall reachability to PeerASN is lost (i.e., no usable eBGP session or forwarding relationship exists 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).

If multiple parallel links or sessions exist between two ASes, the failure of a single link or session that does not break overall AS-to-AS reachability MUST NOT be reported as a "PeerASN:*:Failure" entry. Only when the Local AS determines that it has no usable path to PeerASN for any prefixes does it report such a failure.

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 (loss of AS reachability, policy change, prefix withdrawal/restore), an Agent increments its Batch Counter and reports updated AS Reachability 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.

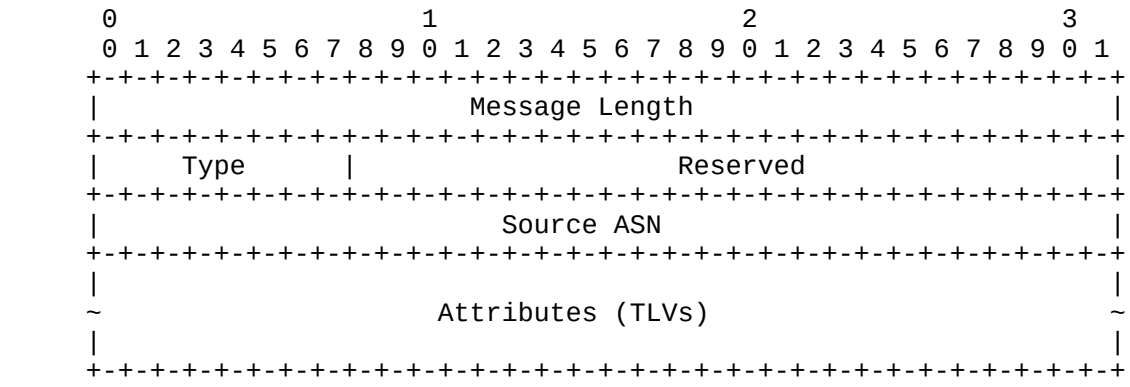
* Filtering: Receiving Agents validate updates and install import policy rules to reject stale routes that traverse AS pairs with lost reachability or that violate reported policy/prefix constraints.

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.

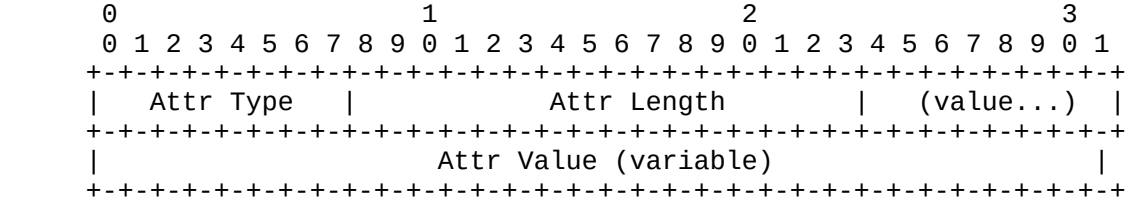
* Type (8 bits): Message type (Agent-to-Repository in Section 6; Repository-to-Agent in Section 7).

* 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 message.

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 AS Reachability 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 change affecting AS reachability (including loss of overall AS-to-AS reachability, policy changes, and 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) AS Reachability 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. AS Reachability State List (0x05)

Encodes a list of AS reachability 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:

```
as-reachability-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 a Update message, 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: AS Reachability State List (0x05), Subscribed ASes (0x06)

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

6.3. Agent Procedures

Initialization:

* Set Batch Counter to 0.

* Choose Function (default example: 1234).

* Collect current AS Reachability 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.

AS Reachability/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 AS Reachability 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: AS Reachability 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_AS_Reachability 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 (in Section 12), 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 AS pair whose reachability is lost), 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. AS-Internal Propagation of Failure Information

Within an AS, there may be multiple BGP routers, including multiple border routers participating in eBGP. The Agent is responsible for propagating AS reachability changes and related filtering policies to all relevant BGP routers inside the AS.

The Agent connects to the AS's border BGP routers using existing management and configuration channels, such as:

* BGP Monitoring Protocol (BMP) [RFC7854] to receive BGP session state updates and routing information.

* Network Configuration Protocol (NETCONF) [RFC6241] or RESTCONF with YANG [RFC7950] models to configure import/export policies.

When the Agent detects a change in AS reachability, it:

1. Updates the Batch Counter and refreshes outbound marking on all eBGP sessions (Section 8.2).

2. Computes new import filtering rules based on the received Repository updates (Section 8.1).

3. Installs these rules on all border routers using the configured management interface.

Internal routers that are not directly connected to the Agent learn about stale routes via normal BGP mechanisms:

* If a route is rejected at a border router's Adj-RIB-In, the border router will not advertise it via iBGP to internal routers. In topologies where route reflectors or Confederations are used, internal routers will simply not see the stale route or will see its withdrawal.

* If the filtering is applied at the point of import, internal routers receive fewer or no stale routes, reducing churn in the AS-internal control plane.

The Agent itself does not directly propagate BGP withdrawals or updates; it only influences the policies applied by the border routers. This design minimizes changes to existing BGP implementations and operational practices.

10. Deployment Considerations

The introduction of a logically centralized Repository raises questions about its deployment, operation, and governance. This section provides guidance on how BGP-FP could be deployed in the global Internet.

10.1. Repository Organization

It is RECOMMENDED that a new organization be established to operate the Repository, with participation from:

* Tier-1 ASes, which have broad visibility into inter-domain routing , play a crucial role in global route convergence..

* Regional Internet Registries (RIRs), which already operate critical infrastructure such as RPKI repositories and have established trust relationships with network operators.

This organization would be responsible for:

* Operating the distributed Repository infrastructure.

* Defining policies for access control, data retention, and privacy.

* Coordinating with existing RPKI and routing security efforts.

10.2. Distributed Deployment

The Repository SHOULD be deployed in a distributed manner across multiple locations and administrative domains. This distribution serves several purposes:

* Reducing latency for Agents in different regions.

* Improving resilience against failures and attacks.

* Avoiding a single point of failure or control.

A possible architecture is similar to that of RPKI repositories [RFC6480] or distributed logs in Certificate Transparency [RFC9162], but tailored for AS reachability state.

10.3. Consensus and Trust

To ensure consistency and integrity across distributed Repository nodes, it is RECOMMENDED to use a Byzantine fault-tolerant (BFT) consensus protocol, such as variants of Practical Byzantine Fault Tolerance (PBFT) or similar algorithms. This allows the Repository to tolerate malicious or faulty nodes while maintaining a consistent view of the AS reachability state.

A rotating leadership model can be used to:

* Distribute the load of ordering and committing updates.

* Reduce the risk of any single party gaining long-term control over the consensus process. The precise choice of consensus protocol and governance structure is left to the organization operating the Repository and may be specified in separate deployment documents.

10.4. Integration with RPKI Infrastructure

The architecture of BGP-FP shares significant similarities with the Resource Public Key Infrastructure (RPKI) [RFC6480]. RPKI employs origin validation to filter routes with invalid source assertions, whereas BGP-FP disseminates AS reachability failure states to filter stale routes. Both mechanisms rely on cryptographic validation to secure routing information. Consequently, it is RECOMMENDED that BGP-FP deployment leverage the existing RPKI infrastructure to minimize deployment overhead and capitalize on established operational trust models. This approach requires specific extensions to existing RPKI components:

* *Repository Enhancements*: RPKI Repository systems SHOULD be extended to support BGP-FP data objects. This includes implementing capabilities for the active push of AS reachability updates to subscribed local Relying Party (RP) software. Furthermore, repositories MUST support Merkle verification mechanisms and the generation of Merkle proofs to ensure the integrity and authenticity of the propagated state.

* *RP Software Extensions*: Local RPKI Relying Party software SHOULD be extended to assume the functionality of the BGP-FP Agent. This involves monitoring local BGP routers to detect the reachability status between the Local AS and its neighbors. The RP software MUST be capable of uploading this state information to the repository and performing local Merkle root calculation and verification.

* *Protocol Extensions*: The RPKI-to-Router (RTR) protocol [RFC8210] MUST be extended to convey AS reachability failure lists to routers. This allows routers to filter stale routes based on BGP-FP data, similar to how they currently filter routes with invalid RPKI Route Origin Authorizations (ROAs).

10.5. Incremental Deployment

BGP-FP is designed to allow incremental deployment:

* An AS can deploy BGP-FP unilaterally to filter stale routes from other participating ASes, without requiring global adoption.

* Early deployment can focus on monitoring and verification of AS reachability state, with optional enforcement of route filtering.

* As more ASes adopt BGP-FP, the benefits of faster convergence and reduced route-flap propagation increase.

11. Example of Operation

This section illustrates the operation of BGP-FP using a simplified topology with six ASes. The example demonstrates how loss of AS reachability between AS1 and AS2 leads to multiple rounds of path exploration in the worst case, and how BGP-FP could accelerate convergence by preventing the propagation of stale routes.

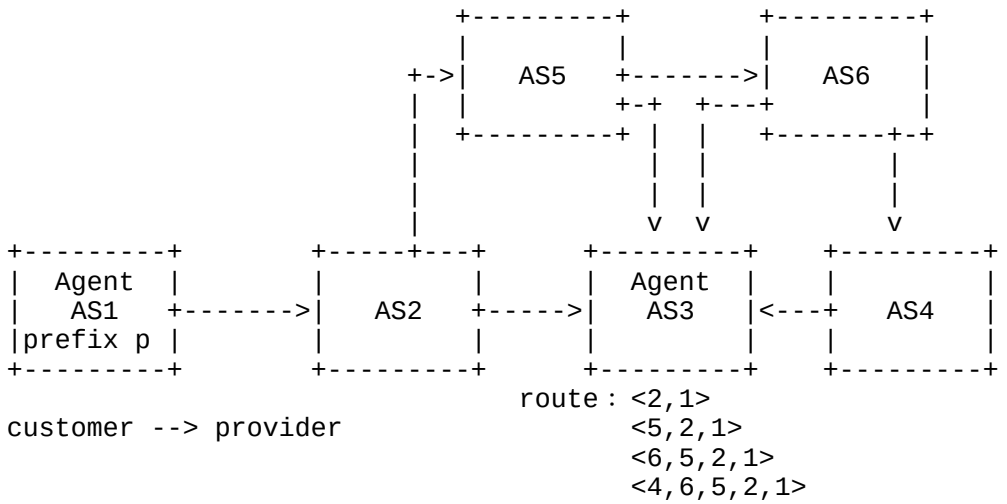


Figure 2: An Example topology with 6 ASes

11.1. Topology and Customer-Provider Relationships

Consider the following ASes: AS1, AS2, AS3, AS4, AS5,AS6.

The customer-provider relationships are:

* AS1 -> AS2 (AS1 is a customer of AS2)

* AS2 -> AS3 , AS2 -> AS5

* AS4->AS3

* AS5->AS6, AS5->AS3

* AS6->AS4, AS6->AS3

These relationships determine the allowed directions of route advertisement. For simplicity, this example assumes that AS1 announces a single prefix P to its providers. Under normal conditions, AS3 would learn the following paths to prefix P, ordered by AS_PATH length:

* Path 1: <2,1>

* Path 2: <5,2,1>

* Path 3: <6,5,2,1>

* Path 4: <4,6,5,2,1>

11.2. Loss of AS Reachability Between AS1 and AS2

Suppose the overall AS reachability between AS1 and AS2 is lost. In a plain BGP deployment without BGP-FP, AS2 will withdraw its route to prefix P from AS3 and AS5. However, routes learned via other ASes (AS5, AS6, AS4) may still be considered valid until their own withdrawals propagate.

In the worst case, AS3 would perform path exploration as follows:

1. *Path 1 Withdrawal*: AS2 withdraws the path <2,1> from AS3. AS3 removes this path.

2. *Path 2 Exploration*: AS3 may then select the next best path <5,2,1>. It sends a withdrawal for this path to AS5. AS5, upon processing the withdrawal, removes the route and propagates it to AS6 and AS3.

3. *Path 3 Exploration*: AS3 may then select the next best path <6,5,2,1>. It sends a withdrawal for this path to AS6. AS6 processes and propagates the withdrawal to AS4 and AS3.

4. *Path 4 Exploration*: Finally, AS3 may select the last known path <4,6,5,2,1>. It sends a withdrawal to AS4. AS4 processes it.

After multiple rounds of propagation and MRAI timers, AS3 converges to having no valid route to prefix P. The root cause is that the critical information, “AS1 is no longer reachable from AS2”, did not propagate in time to AS3. AS3 continued to accept routes whose AS_PATH traversed the failed AS pair (AS1, AS2), assuming they were still valid.

11.3. BGP-FP Operation

With BGP-FP deployed:

* *Failure Detection & Reporting*: The Agent in AS1 detects the loss of reachability to AS2. It increments its Batch Counter and reports an AS Reachability State entry “AS2:*:Failure” to the Repository.

* *State Dissemination*: The Repository securely forwards this authenticated state to all Agents subscribed to AS1, including the Agent in AS3.

* *Proactive Filtering*: Upon receiving and validating the update, the Agent in AS3 installs an import policy rule on its border routers. This rule rejects *any* route whose:

  1. AS_PATH contains the adjacent pair “… AS1 AS2 …” , AND

  2. Carries a freshness community from AS1 with a Batch Counter value older than the new one reported in the Repository update.

Consequently, when AS3 receives the updates or withdrawals for paths like <5,2,1> or <6,5,2,1>, it can immediately recognize them as stale because they traverse the failed link between AS1 and AS2. The routes are filtered at the edge, preventing them from being considered in the BGP decision process.

This mechanism breaks the chain of path exploration. AS3 converges to “no route” state as soon as the withdrawal for the direct path <2,1> is processed, without exploring alternative stale paths. This example highlights how BGP-FP’s authenticated AS reachability propagation can drastically reduce control-plane churn and convergence time.

12. 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.

13. 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.

14. IANA Considerations

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

14.1. 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.

14.2. "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.

14.3. "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 AS Reachability State List
* 0x06 Subscribed ASes
* 0x07 Merkle Proof

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

15. 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.

16. References

16.1. 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>.

[RFC8210] Bush, R. and R. Austein, "The Resource Public Key Infrastructure (RPKI) to Router Protocol, Version 1", RFC 8210, DOI 10.17487/RFC8210, September 2017, <https://www.rfc-editor.org/rfc/rfc8210>.

16.2. 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>.

[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/rfc/rfc7854>.

[RFC6241] Enns, R., Ed., Bjorklund, M., Ed., Schoenwaelder, J., Ed., and A. Bierman, Ed., "Network Configuration Protocol (NETCONF)", RFC 6241, DOI 10.17487/RFC6241, June 2011, <https://www.rfc-editor.org/rfc/rfc6241>.

[RFC7950] Bjorklund, M., Ed., "The YANG 1.1 Data Modeling Language", RFC 7950, DOI 10.17487/RFC7950, August 2016, <https://www.rfc-editor.org/rfc/rfc7950>.

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