

RTGWG
Internet-Draft
Intended status: Informational
Expires: 3 September 2026

A. Abdelsalam, Ed.
C. Filsfils
Cisco Systems, Inc.
E. Ruan
D. Cai
Alibaba Cloud
2 March 2026

IP Fast Reroute for BGP-Only Networks
draft-abdelsalam-rtgwg-ipfrr-bgp-only-network-00

Abstract

IP Fast Reroute (IPFRR) [RFC5714] is widely deployed in IP networks to provide protection against failures by invoking locally determined repair paths. Traditional IPFRR deployments require the use of a link-state IGP to provide locally computed repair paths. This document provides the problem statement and outlines a solution to enable IPFRR computation and deployment in BGP-only networks. The solution maintains standard BGP routing operations while providing fast reroute capabilities comparable to IGP-based deployments.

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 3 September 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	2
2. Problem Statement	2
3. IPFRR for BGP-only networks	3
3.1. Topology Collection	3
3.2. The IPFRR Agent	4
4. Considerations for BGP Policies	5
5. Security Considerations	5
6. References	5
6.1. Normative References	5
6.2. Informative References	5
Authors' Addresses	6

1. Introduction

IP Fast Reroute (IPFRR) [RFC5714] is widely deployed in IP networks to provide protection against failures by invoking locally determined repair paths.

In Data Center (DC) networks with extensive Equal-Cost Multi-Path (ECMP), operators often rely on local fast ECMP re-hashing to protect against failures of ECMP members. However, some failures are not addressed by this ECMP re-hashing and require additional fast reroute solutions for comprehensive protection. The requirements and mechanisms for IPFRR in Artificial Intelligence (AI) DC networks and DCI environments are described in [I-D.draft-clad-rtgwg-ipfrr-aiml].

Deploying IPFRR in networks traditionally requires running a link-state IGP. This document describes an FRR solution for BGP-only networks where no IGP is present, like Massively Scalable Data Centers (MSDCs), AI DCs, and DCI networks.

2. Problem Statement

[RFC5714] specifies the IPFRR framework assuming link-state IGP deployment. This framework relies on two fundamental capabilities:

1. Complete topology visibility through Link-State Databases (LSDBs)
2. Infrastructure to compute paths using the LSDB and install repair paths in the forwarding table

BGP-only networks face the following challenges in deploying IPFRR:

- * **Lack of Topology Information:** As a path-vector protocol, BGP advertises reachability information without topology details. BGP routers maintain no equivalent to an IGP LSDB.
- * **No Native Topology-Based Computation:** BGP's best-path selection process [RFC4271] operates on path attributes, not topology. BGP has no inherent mechanism to compute paths based on network topology.
- * **Policy Complexity:** Unlike IGPs where forwarding follows shortest paths, BGP policies at intermediate hops can alter forwarding decisions, complicating repair path computation. See Section 4 for more details.

These limitations prevent direct application of IPFRR techniques in BGP-only environments.

3. IPFRR for BGP-only networks

This document describes a solution that enables IPFRR computation and installation of repair paths in BGP-only networks by leveraging topology information obtained via BGP-LS.

The solution does not modify base BGP routing protocol operations in the BGP-only network fabric, which continues to provide routing using the BGP best path selection process [RFC4271].

3.1. Topology Collection

[I-D.ietf-idr-bgp-ls-bgp-only-fabric] defines a solution for obtaining a detailed topology view, similar to that available when using link-state routing protocols, in networks where BGP is used as the only routing protocol. It specifies extensions to the BGP Link-State (BGP-LS) address family and procedures for advertising topology information in BGP-only networks.

With this topology information, IPFRR computations can be performed on the BGP-only network, using the same principles as in link-state IGPs (OSPF, IS-IS). The primary difference lies in how the topology database is populated.

The extensions in [I-D.ietf-idr-bgp-ls-bgp-only-fabric] have been implemented and available as open-source in the FRR Routing Stack [FRRouting].

3.2. The IPFRR Agent

The solution introduces an independent logical entity, referred to as the "IPFRR Agent" which is responsible for IPFRR computation and programming of repair paths on behalf of BGP routers that have registered to it. Depending on the deployment model, an IPFRR Agent may service one or multiple BGP routers.

The IPFRR Agent is an application that can be located within a network node, on an out-of-network server, within a controller, or elsewhere.

The IPFRR Agent continuously collects the real time network topology information from a centralized BGP-LS speaker. On behalf of each of the BGP router that has registered to it, the agent runs the IPFRR computations, and programs the resulting repair paths on the corresponding BGP router.

The workflow to deploy IPFRR in a BGP-only network via the IPFRR Agent involves the following steps:

1. Topology collection: All BGP routers report their local topology view to one or multiple centralized BGP-LS speakers. BGP Route Reflectors (RRs) are typically used here to scale the distribution of this information.
2. Topology consolidation: The centralized BGP-LS speakers consolidate the BGP-LS topology information received from all BGP routers to build a database representing the full network topology.
3. Topology retrieval: The IPFRR Agent retrieves the full real time network topology view from the centralized BGP-LS speaker. The IPFRR Agent updates its topology database upon receiving BGP-LS advertisements. The IPFRR Agent may be co-located with or external to the centralized BGP-LS speaker.
4. IPFRR computation: The IPFRR Agent computes the IPFRR repair paths for each of its registered BGP routers, using its topology database.
5. Programming: The IPFRR Agent installs the computed IPFRR repair paths on the corresponding registered BGP router using a standard or well-known southbound API (e.g., NETCONF, RESTCONF, gRPC, or local APIs).

The details of the IPFRR computation algorithm are outside the scope of this document.

4. Considerations for BGP Policies

In link-state IGP networks, IPFRR computation relies solely on the network topology, as forwarding decisions are determined by the shortest path algorithm without local policy modifications. However, in eBGP hop-by-hop environments, BGP policies applied on intermediate nodes may affect forwarding decisions, and therefore IPFRR computation outcomes. These policies MUST be accounted for during IPFRR computation to ensure repair paths are viable and loop-free.

5. Security Considerations

The solution relies on BGP-LS for topology population. Standard BGP-LS security considerations apply. Additional care must be taken to secure the interface between the IPFRR Agent and BGP speakers or controllers, especially when using external southbound APIs.

6. References

6.1. Normative References

- [I-D.ietf-idr-bgp-ls-bgp-only-fabric]
Talaulikar, K., MahendraBabu, A. B., Filsfils, C.,
Ananthamurthy, K., Zandi, S., Dawra, G., and M. Durrani,
"BGP Link-State Extensions for BGP-only Networks", Work in
Progress, Internet-Draft, draft-ietf-idr-bgp-ls-bgp-only-
fabric-04, 20 October 2025,
<<https://datatracker.ietf.org/doc/html/draft-ietf-idr-bgp-ls-bgp-only-fabric-04>>.
- [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>>.
- [RFC5714] Shand, M. and S. Bryant, "IP Fast Reroute Framework",
RFC 5714, DOI 10.17487/RFC5714, January 2010,
<<https://www.rfc-editor.org/info/rfc5714>>.

6.2. Informative References

- [FRRouting]
"Add support for BGP-LS for BGP fabric", 2026,
<<https://github.com/FRRouting/frr/pull/20726>>.

[I-D.draft-clad-rtgwg-ipfrr-aiml]

"IP Fast Reroute for AI/ML Fabrics", 2026,
<<https://datatracker.ietf.org/doc/draft-clad-rtgwg-ipfrr-aiml>>.

Authors' Addresses

Ahmed Abdelsalam (editor)
Cisco Systems, Inc.
Italy
Email: ahabdels@cisco.com

Clarence Filsfils
Cisco Systems, Inc.
Belgium
Email: cf@cisco.com

Eddie Ruan
Alibaba Cloud
United States of America
Email: eddie.ruan@alibaba-inc.com

Dennis Cai
Alibaba Cloud
United States of America
Email: d.cai@alibaba-inc.com