

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: 6 June 2026

L. Dunbar
A. Retana
Futurewei
K. Patel
Arcus
K. Majumdar
Oracle
3 December 2025

Metadata Constrained Distribution
draft-dunbar-idr-metadata-subscription-control-00

Abstract

This document specifies a receiver-driven Metadata Subscription (MDS) mechanism for BGP. A BGP speaker uses the new MDS NLRI to subscribe to specific service metadata attributes.

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 6 June 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 |
| 2. Conventions used in this document | 3 |
| 3. Summary of Operation | 3 |
| 4. Capability Negotiation | 4 |
| 5. Metadata Subscription (MDS) NLRI Format | 4 |
| 6. Error Handling | 5 |
| 7. Example | 5 |
| 8. Relationship to RTC (RFC 4684) | 5 |
| 9. Manageability and Operational Guidance | 6 |
| 9.1. Service Class RT Plan | 6 |
| 9.2. Interplay with RTC and Import/Export Policy | 6 |
| 9.3. Migration and Staging | 6 |
| 9.4. Operational Telemetry (Recommended) | 7 |
| 10. IANA Considerations | 7 |
| 11. Security Considerations | 7 |
| 12. Normative References | 8 |
| Appendix A. Using Metadata-Filter in 5G Environments | 9 |
| A.1. service class RTs and MDS | 10 |
| A.2. Operational Procedure (Example) | 10 |
| A.3. Benefits in 5G | 10 |
| Authors' Addresses | 11 |

1. Introduction

Service metadata can be attached to BGP UPDATES to enable path selection not only based on traditional routing cost but also on the running conditions of edge-hosted services as described in [I-D.ietf-idr-5g-edge-service-metadata]. These attributes may vary per service and change rapidly; distributing all such metadata to all ingress nodes can be overwhelming especially when some devices have limited processing capability, or when not all routes require consideration of both network cost and edge-environment cost to determine the optimal path.

In common iBGP topologies with Route Reflectors (RRs), advertising edge nodes attach service metadata to UPDATES without knowing which ingress nodes will receive and use the information. To enable

selective delivery, this document specifies a _Metadata Subscription (MDS)_ mechanism by which an ingress node explicitly signals, via a new MDS NLRI, the metadata types, and optionally the set of Route Targets (RTs), it wishes to receive. RRs remove metadata attributes when reflecting an UPDATE to a peer that has not subscribed. When the receiving peer has an active subscription, the Metadata Path Attribute is propagated accordingly. Peers without a subscription do not receive metadata, while reachability remain unchanged.

Unlike static filtering, MDS is dynamic: subscriptions can be installed, refined, and withdrawn as service placement or consumption needs evolve. This subscription-driven model limits control-plane churn and processing overhead while preserving normal BGP reachability semantics.

2. Conventions used in this document

The reader is expected to be familiar with the terminology defined in [I-D.ietf-idr-5g-edge-service-metadata].

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. Summary of Operation

An ingress node that wishes to receive service metadata for routes tagged with particular Route Targets (RTs) signals its interest by advertising one or more MDS NLRI (AFI=IPv4/IPv6, SAFI=TBD-MDS) that identify the relevant RTs. Nodes that do not advertise MDS NLRI are treated as having no subscription and therefore do not receive metadata.

BGP speakers remove metadata attributes when advertising an UPDATE to a peer that has not subscribed to the RT associated with that UPDATE. If the peer has subscribed, the metadata is propagated. Route reachability and other BGP attributes remain unaffected.

MDS state persists until the corresponding MDS NLRI is withdrawn or until the BGP session resets, unless configured otherwise.

A RR MAY advertise MDS NLRI on behalf of its clients to facilitate subscription aggregation, provided that it tracks the client subscriptions accurately.

Support for Enhanced Route Refresh [RFC7313] is RECOMMENDED to facilitate on-demand resynchronization.

4. Capability Negotiation

A BGP speaker that is able to receive and process MDS NLRI MUST advertise the corresponding (AFI, SAFI) pair (AFI = IPv4 or IPv6, SAFI = TBD-MDS) using the Multiprotocol Extensions Capability [RFC4760]. A speaker MUST NOT send MDS NLRI to a peer unless this capability has been successfully negotiated.

5. Metadata Subscription (MDS) NLRI Format

The MDS NLRI is encoded in MP_REACH_NLRI and MP_UNREACH_NLRI attributes [RFC4760] with (AFI = IPv4 or IPv6, SAFI = TBD-MDS). The Length of Next Hop Network Address MUST be set to 0.

Multiple MDS NLRIs MAY be advertised. Their effects are additive: if an RT is listed in any active MDS NLRI, the peer is considered subscribed for that RT. Withdrawal of an MDS NLRI removes only the corresponding RT entries.

The wire format of the MDS NLRI is shown below:

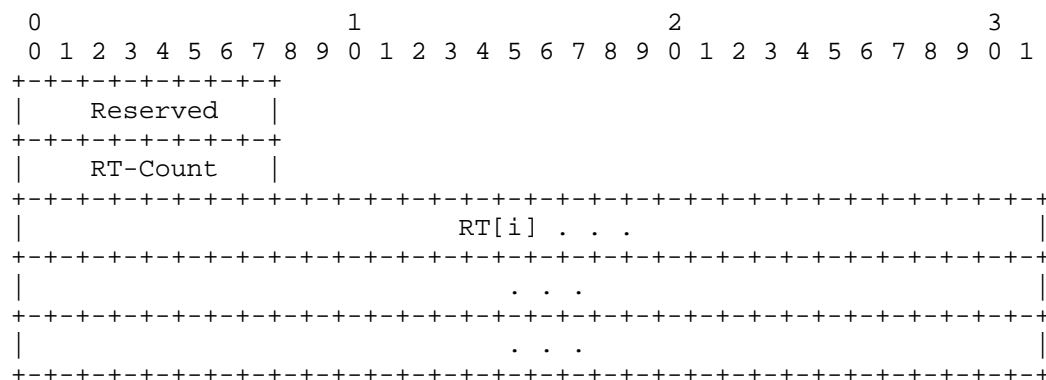


Figure 1: Metadata Subscription (MDS) NLRI Format

Reserved(1 octet): Reserved for future use and MUST be set to 0.

RT-Count (1 octet): This field indicates the number of Route Target entries.

RT[i]: Twelve octetfield containing RT entries encoded identically to the Route Target membership NLRI defined in [RFC4684].

6. Error Handling

Malformed MDS NLRI MUST be treated-as-withdraw, as specified in [RFC7606]. The BGP session MUST NOT be reset as a result of a malformed MDS NLRI. If errors persist, an implementation MAY use the AFI/SAFI disable procedure described in [RFC7606].

If a BGP speaker receives an MDS NLRI for a SAFI it has not advertised support for, it MUST ignore the NLRI.

7. Example

A peer signals “subscribe to metadata for RT = 64500:100” by advertising an MDS NLRI listing that RT. The receiving peer will propagate service metadata for routes carrying RT=64500:100 toward this subscriber. Routes carrying this RT sent to non-subscribed peers will have metadata removed, while reachability is still propagated.

Peer subscribes to metadata for RT 64500:100

UPDATE

Path Attributes:

ORIGIN: IGP

AS_PATH: (iBGP)

MP_REACH_NLRI (AFI=IPv4, SAFI=TBD-MDS)

NLRI:

MDS NLRI:

RT-Count: 1

RT[1]: 64500:100

8. Relationship to RTC (RFC 4684)

RTC [RFC4684] constrains which routes are propagated to a peer based on the RTs that the peer has expressed interest in. Metadata Subscription (MDS) applies a complementary control: it determines whether service metadata is included when those routes are advertised. Thus, RTC governs propagation of reachability information, while MDS governs propagation of associated service metadata. Both mechanisms may be deployed together.

9. Manageability and Operational Guidance

In this specification, Route Targets (RTs) are used to delineate service classes, not merely VPN membership. A group of routes may carry multiple RTs to identify VPN or customer groupings, indicate reachability or constrain membership, or identify routes carrying particular service characteristics. MDS then keys on the service class RTs to control whether metadata is delivered to a given peer. Nodes that do not subscribe to a service class RT receive reachability normally but not the corresponding metadata.

9.1. Service Class RT Plan

Operators SHOULD define a small, stable set of service classes per customer, application, or administrative domain. Advertised routes may be tagged with both:

- * base RT(s) that identify the VPN or customer membership and govern reachability, and
- * service class RT(s) that identify the class of service whose routes may carry service metadata.

Nodes that use metadata subscribe to the appropriate service class RT(s) via MDS NLRI so that service metadata is propagated. Nodes that do not use metadata simply do not subscribe; they still receive reachability for the routes but without service metadata attached.

Example (illustrative): A DC edge node advertises 192.0.2.0/24 with RT-VPN=64500:100 and RT-ULL=64500:200, attaching service metadata. An RR advertises the route with metadata to peers that have subscribed (using the MDS NLRI) to RT-ULL (64500:200), and advertises the same route without metadata to peers that have not subscribed. Reachability remains unchanged.

9.2. Interplay with RTC and Import/Export Policy

If multiple RTs are used (as above), RTC SHOULD remain keyed to the base RT(s) so that reachability distribution is unaffected by MDS usage. The service class RT is used for MDS subscription matching only.

9.3. Migration and Staging

A pragmatic introduction plan is:

1. _Define service class RTs_ and add them to exporter policy alongside existing RTs.
2. _Enable MDS on RRs_; verify that subscribed peers receive metadata unchanged.
3. _For ingress nodes that do not use metadata_, simply omit MDS subscription; validate that reachability persists and metadata is not delivered.
4. _Broaden coverage_ to additional service classes only as needed; keep the service class RT set small and well documented.

9.4. Operational Telemetry (Recommended)

Although MDS focuses on RT-based subscription of metadata, implementations should expose minimal telemetry for validation and troubleshooting. Useful telemetry includes: the number of active MDS entries per peer, the count of UPDATES where metadata was propagated or omitted due to subscription state, and timestamps of last subscription change. Visibility into these counters can help operators verify proper behavior at scale.

10. IANA Considerations

IANA is requested to allocate a new SAFI from the “SAFI Values” registry:

Name: Metadata-Subscription (MDS)
Reference: This document

No other IANA actions are requested by this document.

11. Security Considerations

This document introduces no new security vulnerabilities beyond those discussed in [RFC4684] and [I-D.ietf-idr-5g-edge-service-metadata].

MDS may reveal that a node is interested in service metadata for particular RTs, which could disclose policy intent or service-usage characteristics. To limit exposure, deployments should primarily use MDS within iBGP, and the set of peers permitted to advertise or receive MDS NLRI should be controlled.

MDS affects only whether metadata is propagated; route reachability is preserved regardless of subscription state. However, receiving metadata may influence path or service instance selection at the ingress node [I-D.ietf-idr-5g-edge-service-metadata]. Operators

therefore should audit subscription policy, and are encouraged to enable change logging to track subscription additions and withdrawals.

Ignoring MDS NLRI may result in receiving service metadata that the node does not intend to process, possibly consuming unnecessary memory or control plane resources. Conversely, misconfiguration that prevents a node from receiving metadata it expects could affect its service selection decisions. Operators should monitor subscription state and associated telemetry.

12. Normative References

- [I-D.ietf-idr-5g-edge-service-metadata]
Dunbar, L., Majumdar, K., Li, C., Mishra, G. S., and Z. Du, "BGP Extension for 5G Edge Service Metadata", Work in Progress, Internet-Draft, draft-ietf-idr-5g-edge-service-metadata-30, 18 September 2025, <<https://datatracker.ietf.org/doc/html/draft-ietf-idr-5g-edge-service-metadata-30>>.
- [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>>.
- [RFC4684] Marques, P., Bonica, R., Fang, L., Martini, L., Raszuk, R., Patel, K., and J. Guichard, "Constrained Route Distribution for Border Gateway Protocol/MultiProtocol Label Switching (BGP/MPLS) Internet Protocol (IP) Virtual Private Networks (VPNs)", RFC 4684, DOI 10.17487/RFC4684, November 2006, <<https://www.rfc-editor.org/info/rfc4684>>.
- [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/info/rfc4760>>.
- [RFC7313] Patel, K., Chen, E., and B. Venkatachalapathy, "Enhanced Route Refresh Capability for BGP-4", RFC 7313, DOI 10.17487/RFC7313, July 2014, <<https://www.rfc-editor.org/info/rfc7313>>.
- [RFC7606] Chen, E., Ed., Scudder, J., Ed., Mohapatra, P., and K. Patel, "Revised Error Handling for BGP UPDATE Messages", RFC 7606, DOI 10.17487/RFC7606, August 2015, <<https://www.rfc-editor.org/info/rfc7606>>.

[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>>.

Appendix A. Using Metadata-Filter in 5G Environments

In 5G deployments, multiple User Plane Functions (UPFs) or edge gateways anchor PDU sessions near users while distributed edge data centers host application workloads. BGP advertises reachability for these workloads, and may carry service metadata so ingress nodes can consider service conditions in addition to routing metrics when selecting paths.

Service metadata can churn rapidly and inflate UPDATES if flooded to all peers. Many ingress nodes (e.g., UPFs handling best-effort traffic) neither need nor can efficiently process such rapidly changing attributes. The MDS SAFI allows a receiver to signal its interest in metadata associated with specific Route Targets (RTs). Upon receiving an MDS NLRI, a Route Reflector (RR) propagates the matching metadata to that peer.

MDS signaling is dynamic: ingress nodes can add or withdraw MDS entries as service placement evolves, allowing networks to adapt metadata delivery without impacting route availability.

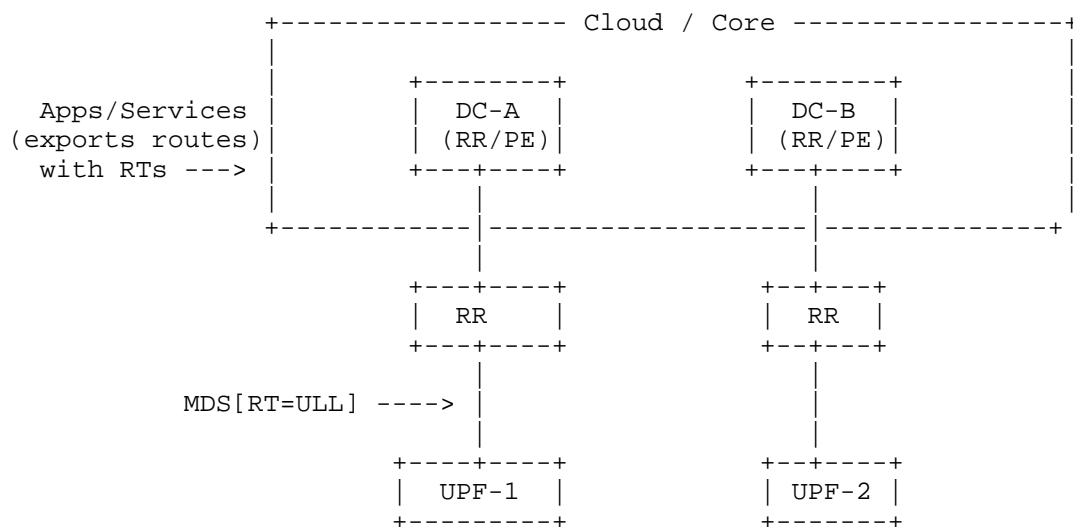


Figure 2: Illustrative 5G Topology with MDS-Scoped Metadata Delivery

- DC-A/DC-B advertise routes tagged with a base RT (VPN/customer) and a service class RT (RT=ULL).

- UPF-1 sends MDS NLRI for RT=ULL: RR propagates metadata on routes with RT=ULL when advertising to UPF-1.
- UPF-2 does not send MDS: it receives routes without metadata.

A.1. service class RTs and MDS

Operators define a small, stable set of `_service class RTs_` to delineate which groups of routes may carry service metadata (e.g., ultra-low-latency vs. best-effort). Exporters tag routes with both a base RT (for reachability/membership) and a service class RT. MDS then keys on the service class RT to control metadata delivery, while RTC (RFC 4684) and normal import/export policy remain keyed to the base RT so reachability is unaffected.

A.2. Operational Procedure (Example)

1. `*Define service classes:` Choose a minimal set of RTs representing classes that may carry metadata (e.g., RT-ULL, RT-VID, RT-ML). Document ownership and intended use.
2. `*Tag exports:` Data center exporters attach a base RT (VPN/customer) and a service class RT to the same NLRI when metadata may accompany the route.
3. `*Enable MDS on RRs:` RRs support the MDS SAFI and omit metadata per-peer when no subscription is present.
4. `*Ingress selection:` UPFs/ingress nodes that want to use metadata advertise MDS NLRIs for the relevant service class RTs. Nodes that don't need metadata do not send MDS.
5. `*Adjust dynamically:` As UE placement or service location changes, ingress nodes add/withdraw MDS entries to tune metadata reception over time.
6. `*Telemetry (recommended):` Expose per-peer MDS entry counts, "metadata omitted" counters, and last-change timestamps to validate behavior.

A.3. Benefits in 5G

- * `*Control-plane scale:` Limits fast-changing metadata propagation to UPFs and routers directly attached to UPFs, reducing UPDATE size and processing load on Route Reflectors and Provider Edge routers while preserving full reachability.
- * `*Service agility:` Supports dynamic changes in metadata subscription as new UEs (for example, drones or autonomous vehicles) roam into or away from UPFs. When a UE moves to a new UPF, that UPF can dynamically express interest in receiving metadata needed for optimal path selection; when the UE leaves, the UPF withdraws its interest, preventing unnecessary metadata distribution.

- * *Operational safety:* Receiver-driven and RT-scoped; enables incremental rollout without impacting route propagation for other peers or service classes.

Authors' Addresses

Linda Dunbar
Futurewei
Dallas, TX,
United States of America
Email: ldunbar@futurewei.com

Alvaro Retana
Futurewei
United States of America
Email: aretana@futurewei.com

Keyur Patel
Arrcus
United States of America
Email: keyur@arrcus.com

Kausik Majumdar
Oracle
United States of America
Email: kausik.majumdar@oracle.com