Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: 9 June 2026

L. Dunbar
Futurewei
C. Li
Huawei Technologies

6 December 2025

BGP Edge Metadata Path Applicability
draft-dunbar-cats-5g-metadata-applicability-00

Abstract

This document analyzes the applicability of the Edge Metadata Path Attribute specified in (ietf-idr-5g-edge-service-metadata) to the computing and service related metrics defined by the IETF CATS Working Group.

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 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

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 9 June 2026.

Copyright Notice

Table of Contents

1. Introduction

The Edge Metadata Path Attribute defined in [I-D.ietf-idr-5g-edge-service-metadata] allows egress routers at edge data centers to advertise edge service related metadata (e.g., availability, preference, delay prediction, and resource status) to ingress routers.

CATS (Computing-Aware Traffic Steering) optimizes traffic steering to service instances by considering both network and compute resource state. The CATS WG defines raw and normalized compute/communication/delay metrics in [I-D.ietf-cats-metric-definition] and describes BGP distribution for service and shared-resource metrics in [I-D.ll-idr-cats-bgp-extension].

This document evaluates how far the existing Edge Metadata mechanisms can carry, approximate, or be extended to support CATS metrics and their operational requirements.

2. Conventions used in this document

The following conventions are used in this document.

Edge DC: Edge Data Center, which provides the hosting environment for the edge services. An Edge DC might host 5G core functions in addition to the frequently used edge services.

gNB: next generation Node B [TS.23.501-3GPP]

RTT: Round-trip Time

PSA: PDU Session Anchor (UPF) [TS.23.501-3GPP]

UE: User Equipment

UPF: User Plane Function [TS.23.501-3GPP]

This document uses the terms "Service Contact Instance (SCI)", "Service", "Compute Resource", "Shared Resource", and "CATS Metric Levels (L0/L1/L2)" as defined in [I-D.ietf-cats-framework] and [I-D.ietf-cats-metric-definition].

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 [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Summary of 5G Edge Service Metadata Mechanism

3.1. Attribute Structure

The Edge Metadata Path Attribute is an optional, non-transitive BGP Path Attribute consisting of a set of Sub-TLVs. Each Sub-TLV carries one metric or an abstracted value about the running environment or capability of an edge service or its hosting site.

The base document defines Sub-TLVs including (but not limited to): Site Preference Index, Site Physical Availability Index, Service Delay Prediction, Raw Measurement, Service-Oriented Capability, and Service-Oriented Available Resource.

The attribute is carried with selected service routes, typically anycast service prefixes, and can also be used in standalone UPDATEs for site-wide values.

3.2. High-Level Distribution Behavior

Edge Metadata is intended for limited-domain distribution. The attribute is non-transitive and is expected to be filtered at administrative boundaries. Route Reflectors (RRs) can further constrain propagation using NO-ADVERTISE or an in-attribute AS-Scope Sub-TLV.

The metadata is propagated in BGP UPDATEs, subject to a minimum advertisement interval (default 30 seconds) to control churn.

3.3. Mapping Sub-TLVs to CATS Metric Categories (Overview)

CATS defines metrics at three abstraction levels: L0 (raw resource and network measures), L1 (normalized per-resource-type metrics), and L2 (composite or policy-derived metrics). The 5G Edge Metadata Path Attribute primarily supports L1 and L2-style abstractions, which aligns with the CATS guidance to avoid flooding highly dynamic raw (L0) metrics in BGP.

The Sub-TLVs defined in [I-D.ietf-idr-5g-edge-service-metadata] naturally map to CATS metric classes as follows:

Service-Oriented Capability -> CATS L1/L2 compute capability:
Represents a normalized measure of compute availability for a service or set of instances.

Service-Oriented Available Resource -> CATS L1 compute availability: Similar to CPU/memory/storage availability, but abstracted into a comparable index.

Service Delay Prediction -> CATS L1 normalized delay or L0 raw delay: Provides a predictive or measured service-instance delay, supporting both raw and normalized forms depending on encoding.

Raw Measurement -> CATS L0 telemetry: Byte/packet counters can be used to construct utilization and load factors that feed into higher-level metrics.

Site Preference Index -> CATS L2 policy composite: Not a standard CATS metric, but fits naturally into the L2 category used for operator-defined policy overrides.

Site Physical Availability Index (SPAI) -> CATS Shared Resource metric: Directly expresses physical-site-level constraints that affect all SCIs hosted at that site.

4. Applicability of 5G Edge Metadata to CATS Metrics

The 5G Edge Metadata Path Attribute provides a ready-made container to carry most CATS-defined compute and service metrics, especially those that must be synchronized across ingress nodes for service-instance selection. The mechanisms align with CATS in several dimensions:

4.1. Compute/Resource Capability (CATS L1/L2)

The Service-Oriented Capability Sub-TLV is directly usable as a normalized compute capability indicator.

It already conforms to the CATS requirement that compute metrics be abstracted and comparable across heterogeneous hardware.

4.2. Compute/Resource Availability (CATS L1)

The Available Resource Sub-TLV corresponds to CATS L1 availability metrics such as CPU headroom, memory availability, or accelerator capacity.

Where CATS defines a family of resource-type-specific primitives, the 5G mechanism provides a unified normalized format suitable for routing-level decisions.

4.3. Service Processing Delay (CATS L0 or L1)

The Service Delay Prediction SubTLV supports both, raw time values (L0), when expressed in milliseconds; and normalized delay scores (L1/L2), when expressed as a 0-100 index.

This flexibility matches CATS' need for both precise measurements and abstracted comparative metrics.

4.4. Traffic Load / Utilization (CATS L0-L1/L2

The Raw Measurement Sub-TLV offers input signals for evaluating congestion, load, and resource pressure.

These serve as L0 primitives from which CATS-compliant L1 or L2 metrics can be derived.

4.5. Site Preference (CATS L2 policy metric)

Although not defined by CATS, the Site Preference Index aligns with the CATS concept of operator-defined L2 composite metrics (e.g., policy-driven affinity, regulatory guidance, cost-based biasing).

4.6. Site Physical Availability (CATS Shared Resource)

Site Physical Availability Index (SPAI) Sub-TLV provides a site/pod/row/floor/DC availability percentage that applies to a group of routes sharing the same physical characteristic. This directly aligns with the "Shared Resource" concept in [I.D.ll-idr-cats-bgp-extension], where a single metric update should influence multiple SCIs without per-service replication.

The SPAI is naturally suited for use as a CATS Shared Resource metric because it reflects conditions that affect all service instances within the same facility. SPAI represents substrate-level constraints such as power redundancy, cooling capacity, rack/row availability, and environmental health - factors that are not tied to any individual service instance. When the physical state of a site changes, all hosted services are simultaneously impacted, meaning a single SPAI update conveys a facility wide condition without requiring per-service updates.

This behavior matches the CATS Shared Resource semantics defined in [I-D.ll-idr-cats-bgp-extension], which aims to avoid churn by enabling one update to apply to many SCIs. SPAI's value is also normalized and abstracted (0-100), satisfying CATS's preference for L1/L2 metrics that can be compared uniformly across egress nodes. Additionally, SPAI is disseminated using the same limited domain, non-transitive BGP propagation model recommended for CATS, without requiring new identifiers or mapping constructs because the site itself serves as the implicit shared resource boundary.

Therefore, SPAI can be used directly as a Shared Resource metric in CATS deployments, helping reduce protocol overhead, avoiding per service metric replication, and improving responsiveness to physical site level events.

4.7. Expressiveness, Granularity, Update Semantics

CATS recommends distributing normalized L1/L2 metrics due to their stability and comparability.

The 5G Edge Metadata model already restricts itself to abstracted values, with raw L0 metrics not directly exposed except through Raw Measurement.

Thus, the overall expressiveness aligns well with CATS' guidance and operational constraints.

5. Distribution Model and Leakage Prevention

6. Metric Lifetimes

The Edge Metadata model uses BGP UPDATE propagation, with default minimum advertisement interval of 30 seconds in iBGP domains. This is consistent with typical CATS control plane expectations for "moderately dynamic" metrics (seconds-to-minutes scale), but may be insufficient for highly bursty L0 measurements unless operators apply dampening and aggregation.

For L2-style normalized values (capability/availability/delay prediction), BGP timeliness is generally adequate in a limited domain where topology diameter is small. For raw L0-style signals (if added later), asynchronous collection and local aggregation before BGP export is RECOMMENDED.

7. Granularity Semantics Issues

CATS separates metrics by scope: per-SCI (service instance), per-service (aggregate over instances), per-site, and shared-resource. The Edge Metadata mechanisms currently mix these scopes:

* Service-Oriented Capability and Available Resource are per-service at a site (implicitly per-service aggregate over instances).

* Service Delay Prediction is per-service at a site, but can represent either a normalized score or raw time; consistent semantic interpretation across ingress nodes is required.

* Site Physical Availability Index is per-site or shared-resource and supports standalone updates. This matches CATS shared resource needs.

* Raw Measurement is per-service route signal, not a standardized CATS L0 set.

If CATS requires explicit per-SCI differentiation (e.g., for multiple instances behind one anycast service prefix), the 5G Edge Metadata encoding needs augmentation to carry an instance identifier or to bind distinct metrics to distinct NLRIs.

8. Scalability and Churn Considerations

The 5G Edge Metadata draft deliberately limits scope to a small subset of services in a limited domain, and applies a minimum advertisement interval to avoid route churn. This is aligned with CATS concerns about frequent metric updates causing cascading BGP UPDATEs. Operators SHOULD:

* Prefer L1/L2 normalized metrics for distribution.

* Aggregate L0 signals locally and export only on meaningful change.

* Use shared-resource/site-wide updates (e.g., Site Physical Availability Index) whenever possible.

In large deployments with many services per site, the per-service Sub-TLV updates could still create noticeable churn. Mapping to the generic CATS BGP metric container may reduce the need for new Sub-TLV types and simplify extensibility.

9. Gaps and Extensions Needed for CATS Compliance

The following additions would allow the 5G Edge Metadata mechanisms to serve as a fully compliant CATS metric container:

Per-SCI Differentiation: Add an Instance-ID field or tie individual metric Sub-TLVs to distinct NLRIs so ingress routers can distinguish metrics for different service instances behind an anycast prefix.

Explicit Resource-Type Identification (for L0->L1 transitions): CATS defines specific raw metrics (CPU load, GPU usage, memory pressure). Adding optional resource-type enumerations would allow Raw Measurement or Capability TLVs to expose more structured CATS L0 primitives when needed.

Metric Lifetime / Validity Interval: CATS defines explicit metric-freshness semantics. The 5G Edge Metadata mechanism can be extended with: TTL or "valid until" timestamps, change-threshold indicators, or explicit "stale" flags.

Scope Identifier for Shared Resource Metrics: Allow SPAI-like metrics to reference a named resource pool (e.g., "DC-Room-A", "GPU-Cluster-2") when a site contains multiple independent shared resources.

Support for More Dynamic Metric Classes: For highly dynamic CATS L0 metrics, support for sub-second updates is generally not appropriate in BGP. Guidance should mandate local smoothing/ aggregation before exporting updates into the control plane.

10. Security Considerations

This document does not introduce new security mechanisms beyond those in [I-D.ietf-idr-5g-edge-service-metadata] and [I-D.ll-idr-cats-bgp-extension]. Operators MUST ensure metadata remains within the trusted limited domain, and SHOULD apply policy-based filtering and AS-scope controls.

11. IANA Considerations

This document makes no IANA requests.

12. Contributors

13. Acknowledgements

14. References

14.1. 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>.

[RFC4360] Sangli, S., Tappan, D., and Y. Rekhter, "BGP Extended Communities Attribute", RFC 4360, DOI 10.17487/RFC4360, February 2006, <https://www.rfc-editor.org/info/rfc4360>.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <https://www.rfc-editor.org/info/rfc4364>.

[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>.

[RFC4761] Kompella, K., Ed. and Y. Rekhter, Ed., "Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling", RFC 4761, DOI 10.17487/RFC4761, January 2007, <https://www.rfc-editor.org/info/rfc4761>.

[RFC4786] Abley, J. and K. Lindqvist, "Operation of Anycast Services", BCP 126, RFC 4786, DOI 10.17487/RFC4786, December 2006, <https://www.rfc-editor.org/info/rfc4786>.

[RFC5291] Chen, E. and Y. Rekhter, "Outbound Route Filtering Capability for BGP-4", RFC 5291, DOI 10.17487/RFC5291, August 2008, <https://www.rfc-editor.org/info/rfc5291>.

[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>.

[RFC5905] Mills, D., Martin, J., Ed., Burbank, J., and W. Kasch, "Network Time Protocol Version 4: Protocol and Algorithms Specification", RFC 5905, DOI 10.17487/RFC5905, June 2010, <https://www.rfc-editor.org/info/rfc5905>.

[RFC6513] Rosen, E., Ed. and R. Aggarwal, Ed., "Multicast in MPLS/BGP IP VPNs", RFC 6513, DOI 10.17487/RFC6513, February 2012, <https://www.rfc-editor.org/info/rfc6513>.

[RFC7120] Cotton, M., "Early IANA Allocation of Standards Track Code Points", BCP 100, RFC 7120, DOI 10.17487/RFC7120, January 2014, <https://www.rfc-editor.org/info/rfc7120>.

[RFC7432] Sajassi, A., Ed., Aggarwal, R., Bitar, N., Isaac, A., Uttaro, J., Drake, J., and W. Henderickx, "BGP MPLS-Based Ethernet VPN", RFC 7432, DOI 10.17487/RFC7432, February 2015, <https://www.rfc-editor.org/info/rfc7432>.

[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/info/rfc8126>.

[RFC8277] Rosen, E., "Using BGP to Bind MPLS Labels to Address Prefixes", RFC 8277, DOI 10.17487/RFC8277, October 2017, <https://www.rfc-editor.org/info/rfc8277>.

[RFC9012] Patel, K., Van de Velde, G., Sangli, S., and J. Scudder, "The BGP Tunnel Encapsulation Attribute", RFC 9012, DOI 10.17487/RFC9012, April 2021, <https://www.rfc-editor.org/info/rfc9012>.

14.2. Informative References

[I-D.ietf-cats-framework]
C. Li, et al, "A Framework for Computing-Aware Traffic Steering (CATS)", November 2025, <https://datatracker.ietf.org/doc/draft-ietf-cats-framework/>.

[I-D.ietf-cats-metric-definition]
C. Li, et al, "CATS Metrics Definition", October 2025, <https://datatracker.ietf.org/doc/draft-ietf-cats-metric-definition/>.

[I-D.ietf-idr-5g-edge-service-metadata]
L. Dunbar, et al, "BGP Extension for 5G Edge Service Metadata", September 2025, <https://datatracker.ietf.org/doc/draft-ietf-idr-5g-edge-service-metadata/>.

[I-D.ll-idr-cats-bgp-extension]
C. Li and P. Liu, "CATS BGP Extension", October 2025, <https://datatracker.ietf.org/doc/draft-ll-idr-cats-bgp-extension/>.

[IANA-BGP-PARAMS]
IANA, "BGP Path Attributes", BGP Path Attributes https://www.iana.org/assignments/bgp-parameters/.

[RFC1136] Hares, S. and D. Katz, "Administrative Domains and Routing Domains: A model for routing in the Internet", RFC 1136, DOI 10.17487/RFC1136, December 1989, <https://www.rfc-editor.org/info/rfc1136>.

[RFC4456] Bates, T., Chen, E., and R. Chandra, "BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)", RFC 4456, DOI 10.17487/RFC4456, April 2006, <https://www.rfc-editor.org/info/rfc4456>.

[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>.

[RFC8799] Carpenter, B. and B. Liu, "Limited Domains and Internet Protocols", RFC 8799, DOI 10.17487/RFC8799, July 2020, <https://www.rfc-editor.org/info/rfc8799>.

[TS.23.501-3GPP] 3rd Generation Partnership Project (3GPP), "System Architecture for 5G System; Stage 2, 3GPP TS 23.501 v2.0.1", December 2017.

Appendix A. Service Delay Prediction Based on Load Measurement

When data centers detailed running status are not exposed to the network operator, historic traffic patterns through the egress routers can be utilized to predict the load to a specific service. For example, when traffic volume to one service at one data center suddenly increases a huge percentage compared with the past 24 hours average, it is likely caused by a larger than normal demand for the service. When this happens, another data center with lower-than-average traffic volume for the same service might have a shorter processing time for the same service.

Here are some measurements that can be utilized to derive the Service Delay Predication for a service ID:

- Total number of packets to the attached service instance (ToPackets);

- Total number of packets from the attached service instance (FromPackets);

- Total number of bytes to the attached service instance (ToBytes);

- Total number of bytes from the attached service instance (FromBytes);

- The actual load measurement to the service instance attached to an egress router can be based on one of the metrics above or including all four metrics with different weights applied to each, such as:

  LoadIndex = w1*ToPackets+w2*FromPackes+w3*ToBytes+w4*FromBytes

  Where w1/w2/w3/w4 are between 0-1. w1+ w2+ w3+ w4 = 1;

The weights of each metric contributing to the index of the service instance attached to an egress router can be configured or learned by self-adjusting based on user feedbacks.

The Service Delay Prediction Index can be derived from LoadIndex/24Hour-Average. A higher value means a longer delay prediction. The egress router can use the ServiceDelayPred sub-TLV to indicate to the ingress routers of the delay prediction derived from the traffic pattern.

Note: The proposed IP layer load measurement is only an estimate based on the amount of traffic through the egress router, which might not truly reflect the load of the servers attached to the egress routers. They are listed here only for some special deployments where those metrics are helpful to the ingress routers in selecting the optimal paths.

Appendix B. Service Metadata Influenced Decision Process

B.1. Egress Router Behavior

Multiple instances of the same service could be attached to one egress router. When all instances of the same service are grouped behind one application layer load balancer, they appear as one single route to the egress router, i.e., the application loader balancer's prefix. Under this scenario, the compute metrics for all those instances behind one application layer balancer are aggregated under the application load balancer's prefix. In this case, the compute metrics aggregated by the Load Balancer are visible to the egress router as associated with the Load Balancer's prefix. However, how the application layer Load Balancers distribute the traffic among different instances is out of the scope of this document. When multiple instances of the same service have different paths or links reachable from the egress router, multiple groups of metrics from respective paths could be exposed to the egress router. The egress router can have preconfigured policies on aggregating various metrics from different paths and the corresponding policies in selecting a path for forwarding the packets received from ingress routers. The aggregated metrics can be carried in the BGP UPDATE messages instead of detailed measurements to reduce the entries advertised by the control plane and dampen the routes update in the forwarding plane. Upon receiving packets from ingress routers, the egress router can use its policies to choose an optimal path to one service instance. It is out of the scope of this document how the measurements are aggregated on egress routers and how ingress routers are configured with the algorithms to integrate the aggregated metrics with network layer metrics.

Many measurements could impact and correspondingly reflect service performance. In order to simplify an optimal selection process, egress routers can have preconfigured policies or algorithms to aggregate multiple metrics into one simple one to ingress routers. Though out of the scope of this document, an egress router can also have an algorithm to convert multiple metrics to network metrics, an IGP cost for each instance, to pass to ingress nodes. This decision-making process integrates network metrics computed by traditional IGP/BGP and the service delay metrics from egress routers to achieve a well-informed and adaptive routing approach. This intelligent orchestration at the edge enhances the service's overall performance and optimizes resource utilization across the distributed infrastructure. When the egress has merged the compute metrics from the local sites behind it, it can include one or more aggregated compute metrics in the Metadata Path Attribute in the BGP UPDATE to the Ingress. Also, an identifier or flag can be carried to indicate that the metrics are merged ones. After receiving the routes for the Service ID with the identifier, the ingress would do the route selection based on pre-configured algorithms (see Section 3 of this document).

B.2. Integrating Network Delay with the Service Metrics

As the service metrics and network delays are in different units, here is an exemplary algorithm for an ingress router to compare the cost to reach the service instances at Site-i or Site-j.

$$Cost\text{-}i = \min\left(w * \left(\frac{ServD\text{-}i * CP\text{-}j}{ServD\text{-}j * CP\text{-}i}\right) + (1-w) * \left(\frac{Pref\text{-}j * NetD\text{-}i}{Pref\text{-}i * NetD\text{-}j}\right)\right)$$

CP-i: Capacity Availability Index at Site-i. A higher value means higher capacity available.

NetD-i: Network latency measurement (RTT) to the Egress Router at the site-i.

Pref-i: Preference Index for Site-i, a higher value means higher preference.

ServD-i: Service Delay Predication Index at Site-i for the service, i.e., the ANYCAST address [RFC4786] for the service.

w: Weight is a value between 0 and 1. If smaller than 0.5, Network latency and the site Preference have more influence; otherwise, Service Delay and capacity availability have more influence.

When a set of service Metadata is converted to a simple metric, a decision process is determined by the metric semantics and deployment situations. The goal is to integrate the conventional network decision process with the service Metadata into a unified decision-making process for path selection.

B.3. Integrating with BGP Route Selection

Not all metadata attributes specified in this document are intended for use in every deployment. Each deployment may choose to consider only a subset of the available metadata attributes based on its specific service requirements.

- Deployment-Specific Attribute Selection:

A deployment may prioritize only certain metadata attributes relevant to its operational needs. For example, one deployment might only use the Service Delay Prediction Index for latency-sensitive applications, while another might focus solely on the Capacity Availability Index to manage resource availability.

- Influence on BGP Decision Process:

The edge service Metadata influences next-hop selection differently from traditional BGP metrics (e.g., Local Preference, MED). Unlike a general next-hop metric that can affect many routes, edge service Metadata selectively impacts optimal next-hop selection for specific routes configured to consider these service-specific attributes. This targeted influence allows for optimized path selection without disrupting broader route decisions.

- Handling Degraded Metrics (Policy-Based):

If a service-specific metric degrades beyond a configured threshold (e.g., the Service Delay Prediction Index exceeds an acceptable delay threshold or the Capacity Availability Index drops below a required level), the ingress router will treat that route as ineligible for traffic steering. This is similar to a BGP route withdrawal, where the degraded route is deprioritized or ignored, even if traditional BGP attributes would otherwise favor it. This ensures that traffic is directed only to service instances that meet the defined performance criteria.

- Fallback to Non-Metadata Routes:

If no suitable routes with the required metadata are available, the BGP decision process defaults to traditional attribute evaluation [RFC4271], ensuring consistent routing even when metadata-specific paths are absent.

This approach provides flexibility and adaptability in routing decisions, allowing each deployment to apply relevant metadata attributes and enforce performance thresholds for improved service quality.

Authors' Addresses

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

Cheng Li
Huawei Technologies
Beijing
China
Email: c.l@huawei.com