

Network Working Group
Internet Draft
Intended status: Standard
Expires: January 6, 2024

L. Dunbar
Futurewei
K. Majumdar
Microsoft
G. Mishra
H. Wang
Huawei
Verizon
H. Song
Futurewei
July 6, 2023

5G Edge Services Use Cases
[draft-dunbar-cats-edge-service-metrics-01](#)

Abstract

This draft describes the 5G Edge computing use cases for CATS and how BGP can be used to propagate additional IP layer detectable information about the 5G edge data centers so that the ingress routers in the 5G Local Data Network can make path selections based on not only the routing distance but also the IP Layer relevant metrics of the destinations. The goal is to improve latency and performance for 5G Edge Computing (EC) services even when the detailed servers running status are unavailable.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#). This document may not be modified, and derivative works of it may not be created, except to publish it as an RFC and to translate it into languages other than English.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

5G Edge Service Use Cases

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."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/1id-abstracts.txt>

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>

This Internet-Draft will expire on April 7, 2021.

Copyright Notice

Copyright (c) 2023 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 (<http://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 Simplified BSD License text as described in [Section 4.e](#) of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction.....	3
2.	Conventions used in this document.....	3
3.	5G Edge Computing Background.....	7
4.	Low Latency Service Instances Selection.....	9
5.	Unbalanced Traffic Distribution by Mobility.....	11
6.	5G EC Service ID.....	11
7.	Site Availability Index.....	11
8.	Site Preference Index.....	12
9.	Network Delay to an ANYCAST Address in 5G EC.....	12
10.	Metrics for Predicting Service Delays.....	13

5G Edge Service Use Cases

10.1. Service Delay Prediction.....	14
10.2. IP-Layer metrics for Service Delay Predication.....	14
11. Algorithm in Selecting the optimal Target Location.....	15
12. Scope of Service Metrics Advertisement.....	16
13. Manageability Considerations.....	17
14. Security Considerations.....	17
15. IANA Considerations.....	17
16. References.....	17
16.1. Normative References.....	17
16.2. Informative References.....	18
17. Acknowledgments.....	18

1. Introduction

This document describes the 5G Edge Computing use cases for CATS and how BGP can be used to propagate additional IP-layer relevant information about the destination so that the ingress routers in the 5G Local Data Network can make path selections based on not only the routing distance but also the IP Layer relevant metrics of the destinations. The goal is to improve latency and performance for 5G Edge Computing (EC) services even when the detailed servers running status are unavailable, as most applications and their hosting servers/VMs' detailed status are not exposed to network operators. Their communications are generally encrypted and do not respond to PING or ICMP messages initiated by routers or network elements.

This draft specifies the IP Layer metrics and algorithms that enable the 5G Local Data Networks (LDN) to dynamically optimize the forwarding of low latency EC services without any knowledge above the IP layer.

2. Conventions used in this document

CATS: Computing-Aware Traffic Steering takes into account the dynamic nature of computing resource metrics and network state metrics to steer service traffic to a service instance.

Service: A monolithic function. A composite service can be built by orchestrating monolithic services.

5G Edge Service Use Cases

Service instance: A run-time environment (e.g., a server or a process on a server) that makes the functionality of a service available. One service can have multiple instances running at the same or different network locations.

CS-ID: The CATS Service ID is an identifier representing a service, which the clients use to access said service. Such an identifier identifies all of the instances of the same service, no matter on where they are actually running. The CS-ID is independent of which service instance serves the service demand. Usually multiple instances provide a (logically) single service, and service demands are dispatched to the different instance by choosing one instance among all available instances.

CB-ID: The CATS Binding ID is an identifier of a single service instance of a given CS-ID. Different service instances provide the same service identified through a single CS-ID, but with different CATS Binding IDs.

Service request: The request for a specific service instance.

CATS-router: A network device (usually at the edge of the network) that makes forwarding decisions based on CATS information to steer traffic belonging to the same service demand to the same chosen service instance.

Ingress CATS-Router: A network edge router that serves as a service access point for CATS clients. It steers the service packets onto an overlay path to an Egress CAN-Router linked to the most suitable edge site to access a service instance.

CATS-ER: CATS-ER is an egress CATS-Router, i.e., the egress endpoint of an overlay path to a service instance. CATS-ER is used to describe the last router that the service instances are attached.

5G Edge Service Use Cases

In a 5G EC environment, the CATS-ER can be the gateway router to the Edge Computing Data Center.

C-SMA: The CATS Service Metric Agent responsible for collecting service capabilities and status, and for reporting them to the C-PS.

NOTE: The above terminologies are the same as those used in 3GPP TR 23.758

C-NMA: The CATS Network Metric Agent responsible for collecting network capabilities and status, and for reporting them to the C-PS

C-PS: The CATS Path Selector determines the path toward the appropriate service location and service instances to meet a service demand given the service status and network status information.

C-TC: The CATS Traffic Classifier is responsible for determining which packets belong to a traffic flow for a particular service demand, and for steering them on the path to the service instance as determined by the C-PS.

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 application servers.

gNB next generation Node B

PSA: PDU Session Anchor (UPF)

SSC: Session and Service Continuity

UE: User Equipment

UPF: User Plane Function

5G Edge Service Use Cases

ANYCAST Instance: refer to the service instance at a specific location which is reachable by the ANYCAST address.

Service Instance Location: Represent a cluster of servers at one location serving the same Service. One service may have a Layer 7 Load balancer, whose address(es) are reachable from external IP network, in front of a set of service instances. From the IP network perspective, this whole group of instances are considered as one service instance at the location.

EC: Edge Computing

Edge Computing Hosting Environment: An environment, such as physical or virtual machines, host the service instances.

NOTE: The above terminologies are the same as those used in 3GPP TR 23.758

Edge DC: Edge Data Center, which provides the Edge Hosting Environment. It might be co-located with 5G Base Station and not only host 5G core functions, but also host frequently used Edge server instances.

LDN: 5G Local Data Network

PSA: PDU Session Anchor (UPF)

RTT: Round Trip Time

RTT-ANYCAST: A list of Round trip times to a group of routers that have the ANYCAST instances directly attached.

5G Edge Service Use Cases

SSC: Session and Service Continuity
UE: User Equipment
UPF: User Plane Function

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. 5G Edge Computing Background

One of the 5G key features is the ultra-low latency services, which are enabled by instantiating one application or service in multiple edge data centers nearby [[3GPP-EdgeComputing](#)]. Those Edge Computing (EC) mini data centers are usually very close to, or co-located with, 5G base stations to minimize the latency. The 5G Local Data Networks (LDN), a.k.a. N6 interface from 3GPP 5G perspective, connect the edge data centers with the 5G User Plane Functions (UPF) with a small number of dedicated routers. The ultra-low latency 5G EC services are registered premium services that require super-low latency and very high SLA. Most UE service requests, such as internet browsing, are not part of the registered ultra-low latency services.

When a UE (User Equipment) initiates the packets using the destination address from a DNS reply or its own cache, the packets from the UE are carried in a PDU session through the 5G Core [5GC] to the 5G UPF-PSA (User Plan Function - PDU Session Anchor). The UPF-PSA decapsulates the 5G GTP outer header and forwards the packets from the UEs to the Ingress router of the 5G LDN. The LDN, the IP Network from 3GPP's 5G Core perspective, is responsible for forwarding the packets to the intended destinations.

When the UE moves out of coverage of its current gNB (next-generation Node B) and anchors to a new gNB, the 5G SMF

5G Edge Service Use Cases

(Session Management Function) could select the same UPF or a new UPF for the UE per standard handover procedures described in 3GPP TS 23.501 and TS 23.502. If the UE is anchored to a new UPF-PSA when the handover process is complete, the packets to/from the UE is carried by a GTP tunnel to the new UPF-PSA. Per TS 23.501-h20 [Section 5.8.2](#), the UE may maintain its IP address when anchored to a new UPF-PSA unless the new UPF-PSA belongs to different mobile operators. 5GC may maintain a path from the old UPF to the new UPF for a short time for the SSC [Session and Service Continuity] mode 3 to make the handover process more seamless.

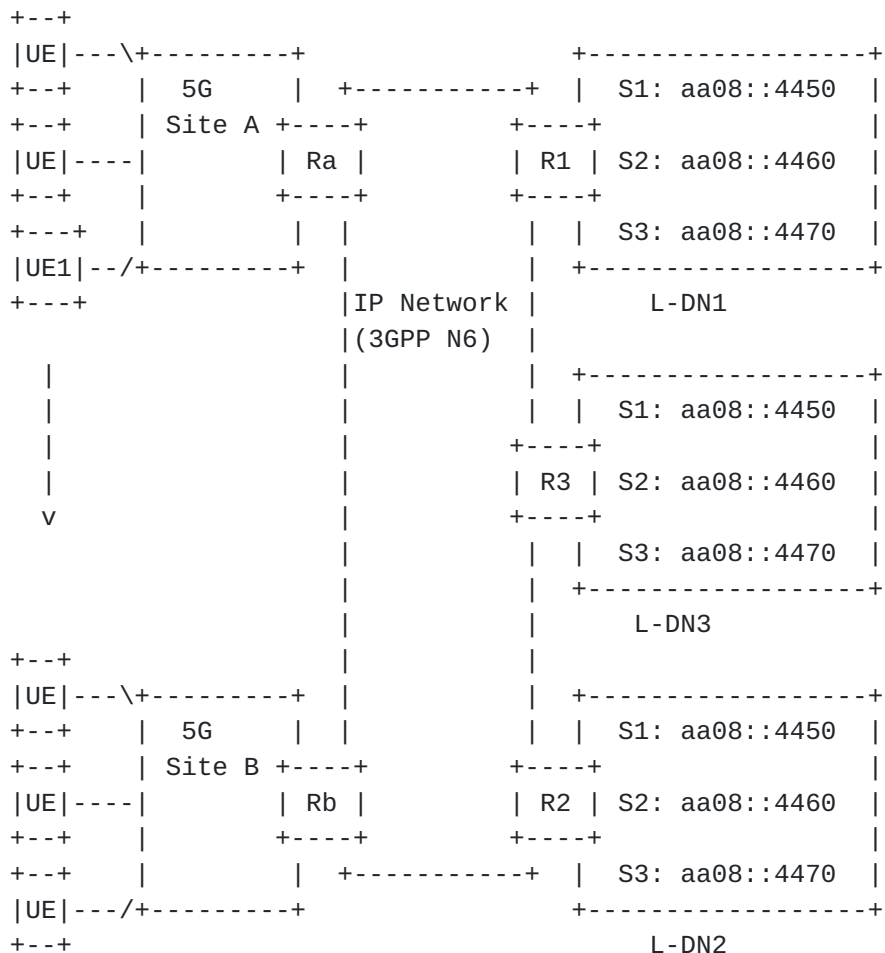


Figure 1: multiple ANYCAST instances in different edge DCs

4. Low Latency Service Instances Selection

Having one application/service instantiated in multiple locations closer to UEs can greatly improve the user experience. But selecting an optimal location for the service requests from a UE may not be that simple.

Using DNS to reply with the address of the service instance location closest to the requesting UE can encounter issues like:

- UE can cache results indefinitely. When the UE moves to a 5G cell site very far away, the cached address may still be used, which can incur a large network delay.
- The service instance at a specific location, directed by the DNS, might be heavily loaded, causing slow or no response when there are available low utilized service instances for the same service at locations very close in proximity.
- No inherent leverage of proximity information present in the network (routing) layer, resulting in performance loss.
- Local DNS resolver becomes the unit of traffic management.

Increasingly, ANYCAST is used to provide better and faster resiliency to failover events. Anycast address leverages the proximity information present in the network (routing) layer. It eliminates the single point of failure and bottleneck at the DNS resolvers. Anycast address can be assigned to instances in multiple data centers to leverage network conditions for balanced forwarding. Another benefit of using the ANYCAST address is removing the dependency on UEs refreshing their cached IP addresses.

Using a Virtual IP address is another method to scale dynamic changes of application instances, a common practice in Cloud Native networking, e.g., Kubernetes. Virtual IP requires the destination gateway node to perform address translation for return traffic, which is unsuitable for underlay network nodes with millions of packets passing by.

5G Edge Service Use Cases

Having multiple locations of the same IP address in the 5G EC LDN can be problematic if path selection is solely based on routing cost as the routing cost differences to reach different EC data centers can be very small. This list elaborates the issues in detail:

- Path Selection: When a new flow comes to an ingress node (Ra in Figure 1), avoiding instability with ANYCAST flipping among paths to the same address can be an issue. The problem also exists in the BGP multipath environment, with the optimal path selected based on routing cost metrics.

The ingress node needs to forward the packets from one flow to the same service instance, a.k.a. Flow Affinity or Flow-based load balancing. The ingress node (Ra/Rb in Figure 1) can use Flow ID (in IPv6 header), or UDP/TCP port number combined with the source address to enforce packets in one flow being placed in one tunnel to one egress router.

- When a UE moves to a new 5G site in the middle of a communication session with an EC service instance, a method is needed to stick the flow to the same EC service instance, which is required by 5G Edge Computing [3GPP TR 23.748]. [5g-edge-compute-sticky-service] describes several approaches to achieve stickiness in the IPv6 domain.

Note: most EC services have shorter sessions, e.g., shorter TCP sessions. Most likely, when a UE is moving to a new 5G site, the TCP session via the old UPF to an EC service instance is already finished. Only a very small percentage of registered EC services need to stick to the original service instance when handover to a new cell tower.

From BGP perspective, the multiple service instances with the same IP address (ANYCAST) attached to different egress routers is the same as multiple next hops for the IP address.

5. Unbalanced Traffic Distribution by Mobility

It is common to have higher capacity EC service instances placed in a metro data center to accommodate more UEs in proximity and fewer placed in remote sites. Sometimes, UEs swarm to a specific site unexpectedly, e.g., a special event at a remote site for a short period, e.g., 1~2 days. The EC service instances in the remote site might be heavily utilized. In contrast, the EC service instances of the same app in the metro DC can be under-utilized. Since the condition can be short-lived or unexpected, it might not make business sense to adjust EC capacity among DCs.

6. 5G EC Service ID

From the network perspective, a service identifier, or IP Layer Service ID, is an ANYCAST address shared by multiple service instances at different locations. Here are some assumptions about the 5G EC services:

- Only the registered EC services, which are only a small portion of the services, need to incorporate the destination related metrics for optimal forwarding.
- The 5G EC controller or management system can send those EC service identifiers to relevant routers.
- The ingress routers' local BGP path compute algorithm has a special plugin that considers both the destination service metrics and traditional BGP path metrics in computing the path to the optimal Next Hop (egress router).

7. Site Availability Index

Site Availability Index is a numeric number representing the percentage of the site being functional, e.g., 100%, 50%, or 0%. When a site goes dark, the Index is set to 0. 50 means 50% capacity functioning. When a data center goes dark (i.e., the Site Availability index goes to 0 caused by a power outage), a large number of service instances are impacted. Instead of sending many BGP route withdrawal messages for

5G Edge Service Use Cases

many address families impacted, the egress router can send one single message to indicate all the routes associated with a site are impacted. The ingress routers can switch all or a portion of the instances associated with the site depending on how much the site is degraded.

Cloud Site/Pod failures and degradation can be caused by a variety of reasons, such as fiber cut connecting to the site or among pods within one site, cooling failures, insufficient backup power, cyber threats attacks, too many changes outside of the maintenance window, etc. Fiber-cut is not uncommon within a Cloud site or between sites.

When those failure events happen, the Edge (egress) router visible to the ingress routers can be running fine. Therefore, the ingress routers with paths to the egress routers can't use BFD to detect the failures.

8. Site Preference Index

As described in [IPv6-StickyService] and [ISPF-EXT-EC], an EC sticky service needs to connect a UE to the service instance that has been serving the UE before the UE moves to a new 5G Site unless there is a failure to that location.

To achieve the goal of sticking a flow from one specific UE to a specific site, a "Site Preference Index" is created. The value of the Site Preference Index can be manipulated for packets of some flows to be steered towards an instance location farther away in routing distance. The "Site Preference Index" enables some sites to be more preferred for handling the UE traffic to an instance than others.

9. Network Delay to an ANYCAST Address in 5G EC

ANYCAST used in 5G EC environment is slightly different from the typical ANYCAST address being deployed. Typical ANYCAST address is used to represent instances in vast different geographical locations, such as different continents. ANYCAST address for "app.net" for Asia lead packets to a server instance of "app.net" hosted in Asia. Therefore, the RTT for "app.net" in Asia, is a single value that represent the round time trip to the server in Asia that host the "app.net".

5G Edge Service Use Cases

5G EC can have one service hosted in multiple EC DCs close in proximity. Routers, i.e., the ingress router to 5G LDN, can forward packets for the ANYCAST address of "app.net" to different egress routers that have "app.net" instances attached.

When "app.net" is hosted in four different 5G EC Data Centers, the RTT to "app.net" ANYCAST address need to be a group of values (instead of one RTT value to a unicast address). The RTT group value should include the CATS-ER router's specific unicast address (e.g., the loopback address) to which the service instance is attached.

RTT to "app.net" ANYCAST Address is represented as:

List of {Egress Router address, RTT value}

This list is called "RTT-ANYCAST".

In order to better optimize the ANYCAST traffic, each router adjacent to 5G PSA needs to periodically measure RTT to a list of CATS-ER routers that advertise the ANYCAST address. The RTT to egress router at Site-i is considered as the RTT to the ANYCAST instance at the Site-i.

10. Metrics for Predicting Service Delays

It is desirable for an ingress router to select a path with the least network delay to an EC data center that has the shortest processing time for the service request from a UE for ultra-low latency services. But it is not easy to predict which site has "the shortest processing time" for an incoming service request because EC data centers have different resources and different allocations of service instances to physical servers.

The Service Delay Index is a value that predicts the processing delays at the site for future service requests. The higher the value, the longer the delay.

5G Edge Service Use Cases

10.1. Service Delay Prediction

Intuitively, an EC data center with more resources (e.g., computing, storage, network bandwidth among servers) can process a service request faster than an EC data center with fewer resources.

A Service Delay Predication value can be assigned to a site based on the relative resource level of the site, e.g., 1-100. A higher Service Delay Predication value means it might take a longer time to process an incoming service. The Service Delay Predication value is just an estimate, not meant to be accurate, even if the value can be adjusted based on the EC data center's actual running status.

10.2. IP-Layer metrics for Service Delay Predication

When EC data centers detailed running status is not exposed to the 5G LDN operator, historic traffic patterns through the LDNs can be utilized to anticipate or 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 number of UEs roaming to the same 5G site needing the service. When this happens, another EC data center with lower-than-average traffic volume for the same service might have a shorter processing time for the same service.

Without knowledge of applications' internal logic, egress routers can measure the traffic patterns to/from the service instances at each location to predict the processing delay of the service at the location. Like the assigned processing delay value, processing delay prediction based on historic traffic patterns might not be accurate but at least reflect the current changes to the service request volume.

5G Edge Service Use Cases

Here are some measurements that can be utilized to compute 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 a CATS-ER can be based on one of the metrics above or including all four metrics with different weights applied to each, such as:

LoadIndex =
 $w1*ToPackets+w2*FromPackets+w3*ToBytes+w4*FromBytes$

Where $0 \leq w_i \leq 1$ and $w1+ w2+ w3+ w4 = 1$.

The weights of each metric contributing to the load index of the service instance attached to a CATS-ER can be configured or learned by self-adjusting based on user feedbacks. The Service Delay Prediction Index can be computed as $LoadIndex/24Hour\text{-}Average$. A higher value means a longer delay prediction.

11. Algorithm in Selecting the optimal Target Location

The goal of the algorithm is to equalize the traffic among multiple locations of the same Service ID.

This exemplary algorithm takes the following attributes into consideration to compare the cost to reach the service instances at Site-i vs. Site-j:

- Service Delay Predication (SerD-i) value,

5G Edge Service Use Cases

- Capacity Availability index (CP-i)
- Preference Index (Pref-i), and
- network delay [NetD-i].

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

w: Weight for load and site information, which is a value between 0 and 1. If smaller than 0.5, Network latency and the site Preference have more influence; otherwise, Server load and its capacity have more influence.

When comparing metrics from Site-j with itself, the value from the algorithm is 1. Cost-i >1 indicates Site-i costs more than Site-j. Therefore, the shortest path to Site-j should be chosen. Cost-i <1 indicates Site-i costs less than Site-j. Therefore, the shortest path to Site-i should be chosen.

12. Scope of Service Metrics Advertisement

Each ultra-low latency EC service might be requested by a small group of UEs. Therefore, an egress router doesn't need to advertise the service metrics to all other routers in the 5G LDN. Likewise, each EC Data Center may only host a small number of low-latency EC services.

"Service ID Bound Group Routers" refers to a group of routers interested in a specific Service ID. The Service Metrics for a specific service ID should be advertised among the routers in the "Service ID bound Group Routers".

5G Edge Service Use Cases

BGP RT Constrained Distribution [[RFC4684](#)] can be used to form the "Service ID Bound Group Routers" by using the "Service ID," which is an IP address prefix, as the Route Target. When an ingress router receives the first packet of a flow destined to a Service ID, the ingress router sends a BGP UPDATE that advertises the Route Target membership NLRI per [RFC4684](#). The ingress router must assign a Timer for the Service ID as the UE that uses the Service ID might move away. Upon receiving a packet destined for the Service ID, the ingress router must refresh the Timer. The ingress router must send a BGP Withdraw UPDATE for the Service ID upon expiration of the Timer.

13. Manageability Considerations

To be added.

14. Security Considerations

To be added.

15. IANA Considerations

To be added.

16. References

16.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC4364] E. Rosen, Y. Rekhter, "BGP/MPLS IP Virtual Private networks (VPNs)", Feb 2006.
- [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>>.

5G Edge Service Use Cases

[RFC8200] s. Deering R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", July 2017

16.2. Informative References

[3GPP-EdgeComputing] 3GPP TS 23.548 V18.1.1, "3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; 5G System Enhancements for Edge Computing; Stage 2", Release 18, April 2023.

[SDWAN-EDGE-Discovery] L. Dunbar, S. Hares, R. Raszuk, K. Majumdar, "BGP UPDATE for SDWAN Edge Discovery", [draft-ietf-idr-sdwan-edge-discovery-10](#), June 2023.

17. Acknowledgments

Acknowledgements to XXX for their review and contributions.

This document was prepared using 2-Word-v2.0.template.dot.

5G Edge Service Use Cases

Authors' Addresses

Linda Dunbar
Futurewei
Email: ldunbar@futurewei.com

Kausik Majumdar
Microsoft
Email: kmajumdar@microsoft.com

Gyan Mishra
Verizon
Email: gyan.s.mishra@verizon.com

Haibo Wang
Huawei
Email: rainsword.wang@huawei.com

HaoYu Song
Futurewei
Email: haoyu.song@futurewei.com