

Network Working Group
Internet Draft
Intended status: Standard
Expires: August 23, 2022

L. Dunbar
Futurewei
K. Majumdar
CommScope
H. Wang
Huawei
G. Mishra
Verizon
February 23, 2022

BGP Update for 5G Edge Computing Service Metadata
draft-dunbar-idr-5g-edge-compute-app-meta-data-06

Abstract

This draft describes a new AppMetaData subTLV carried by Tunnel Encap[RFC9012] Path Attribute for egress router to advertise the running status and environment for the directly attached 5G Edge Computing (EC) servers. The AppMetaData can be used by the ingress routers in the 5G Local Data Network to make path selection not only based on the routing distance but also the running environment of the destinations. The goal is to improve latency and performance for 5G EC services.

The extension enables an EC server at one specific location to be more preferred than the others with the same IP address to receive data flows from a specific source (UE).

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.

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/lid-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) 2022 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
1.1.	5G Edge Computing Background.....	3
1.2.	5G Edge Computing Network Properties.....	4
1.3.	Problem#1: ANYCAST in 5G EC Environment.....	5
1.4.	Problem #2: Unbalanced Anycast Distribution due to UE Mobility.....	7
1.5.	Problem 3: Application Server Relocation.....	8
2.	Conventions used in this document.....	8
3.	Usage of AppMetaData for 5G Edge Computing.....	9
3.1.	Assumptions.....	9
3.2.	IP Layer Metrics to Gauge Application Behavior.....	10
3.3.	AppMetaData Constrained Optimal Path Selection.....	11

4.	BGP Protocol Extension to advertise Load & Capacity.....	12
4.1.	Ingress Node BGP Path Selection Behavior.....	12
4.1.1.	AppMetaData Influenced BGP Path Selection.....	12
4.1.2.	Ingress Router Forwarding Behavior.....	12
4.1.3.	Forwarding Behavior when UEs moving to new 5G Sites.....	14
5.	The Sub-TLVs for AppMetaData.....	14
5.1.	Load Measurement sub-TLV format.....	15
5.2.	Capacity Index sub-TLV format.....	16
5.3.	The Site Preference Index sub-TLV format.....	16
6.	AppMetaData Propagation Scope.....	17
7.	Minimum Interval for Metrics Change Advertisement.....	17
8.	Soft Anchoring of an ANYCAST Flow.....	17
9.	Manageability Considerations.....	19
10.	Security Considerations.....	19
11.	IANA Considerations.....	19
12.	References.....	20
12.1.	Normative References.....	20
12.2.	Informative References.....	20
13.	Acknowledgments.....	21

[1.](#) Introduction

This document describes a new subTLV, AppMetaData, for egress routers to advertise the running status and environment for the directly attached Edge Computing (EC) servers. The AppMetaData can be used by the ingress routers in the 5G Local Data Network to make path selection not only based on the routing distance but also the running environment of the destinations. The goal is to improve latency and performance for 5G Edge Computing services.

1.1. 5G Edge Computing Background

In 5G Edge Computing (EC), one Application can be hosted on multiple Servers in different EC data centers that are close in proximity. The 5G Local Data Networks (LDN) that connect the EC data centers with the 5G Base stations consist of a small number of dedicated routers.

When a User Equipment (UE) initiates application packets using the destination address from a DNS reply or its cache, the packets from the UE are carried in a PDU session through 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 its directly connected Ingress router of the 5G LDN. The LDN for 5G EC 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 (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 UFP-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.

1.2. 5G Edge Computing Network Properties

In this document, 5G Edge Computing Network refers to multiple Local IP Data Networks (LDN) in one region that interconnect the Edge Computing data centers. Those IP LDN networks are the N6 interfaces from 3GPP 5G perspective.

The ingress routers to the 5G Edge Computing Network are the routers directly connected to 5G UPFs. The egress routers to the 5G Edge Computing [EC] Network are the routers that have a direct link to the EC servers. The EC servers and the egress routers are co-located. Some of those Edge Computing Data centers may have virtual switches or Top of Rack [ToR] switches between the egress routers and the servers. But transmission delay between the egress routers and the EC servers is negligible, which is too small to be considered in this document.

When multiple EC servers are attached to one App Layer Load Balancer, only the IP addresses of the App Layer Load Balancer are visible to the 5G LDNs. How an App Layer Load balancer manages the individual servers is out of the scope of the network layer.

The 5G EC Services are registered premium services that require super-low latency and very high SLA. Most services by the UEs are not part of the registered 5G EC Services.

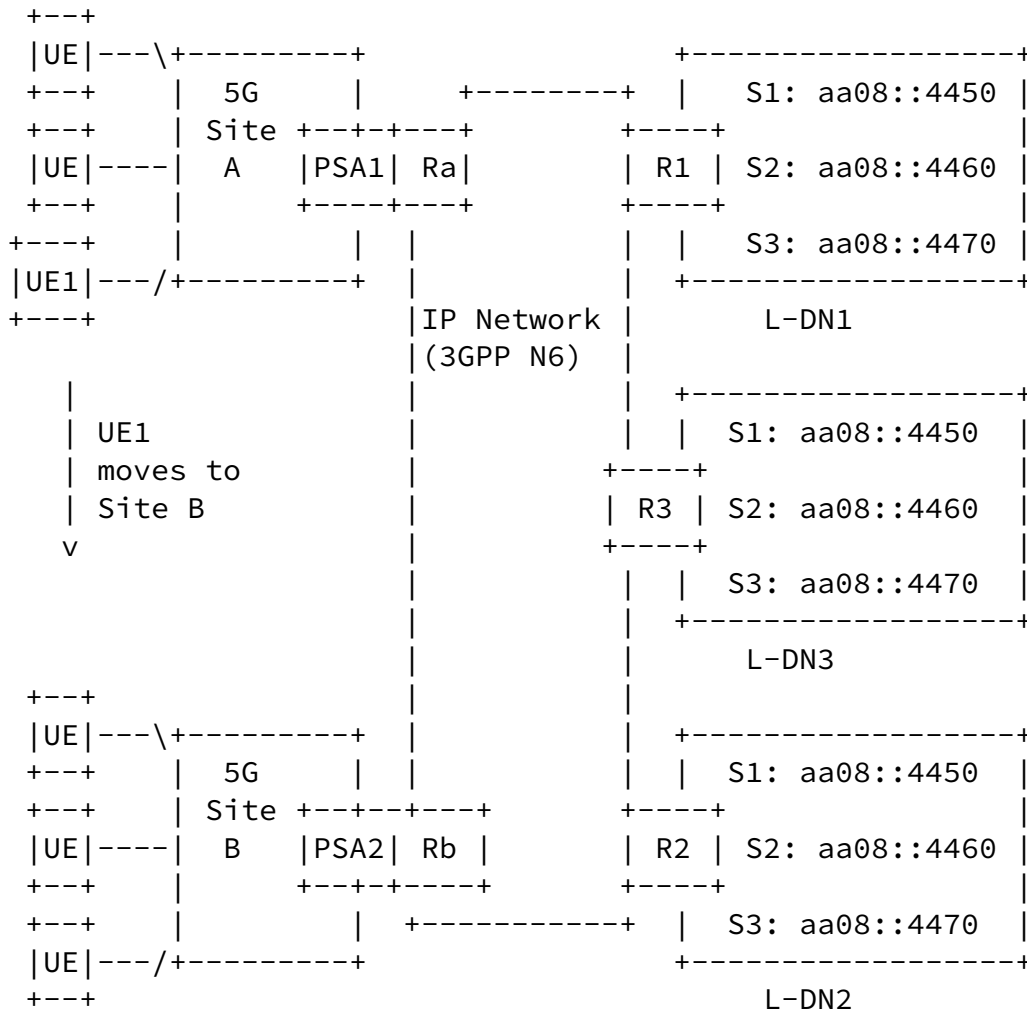


Figure 1: App Servers in different edge DCs

1.3. Problem#1: ANYCAST in 5G EC Environment

Increasingly, Anycast is used by various application providers and CDNs because Anycast provides better and faster resiliency to failover events than GEO database DNS-based load balancing, which relies on DNS to provide a different IP based on source address.

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 multiple app layer load balancers to leverage network condition for balanced forwarding. Another benefit of using the ANYCAST address is removing the dependency on UEs. Some UEs (or clients) might use their cached IP addresses for an extended period instead of querying DNS.

Client using Virtual IP address is a common practice in Cloud Native networking, e.g., Kubernetes, to scale dynamic changes of app servers' instantiations. Virtual IP requires the destination gateway node to perform address translation for return traffic, which is unsuitable for underlay network nodes with millions of flows passing by. The Cloud Native network can also leverage network condition to balance forwarding among multiple Cloud Gateway nodes by assigning the same virtual IP address (ANYCAST).

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 egress routers can be very small. This list elaborates the issues in detail:

- a) Path Selection: When a new flow comes to an ingress node (Ra), how to avoid instability with Anycast flipping between paths to the same address. The problem also exists in the BGP multipath environment, with the optimal path selected based on routing cost metrics.
- b) Ingress node forwards the packets from one flow to the same ANYCAST server.

a.k.a. Flow Affinity, or Flow-based load balancing.

Almost all vendors have supported flow or session based ECMP load balancing and not per packet to avoid out of order packets

for decades. When a flow is handled by an ECMP path, the flow remains on that path for the life of the flow until the flow ends.

The ingress node, (Ra/Rb), 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

Internet-Draft

AppMetaData for 5G EC Service

one tunnel to one Egress router. No new features are needed.

- c) When a UE moves to a new 5G site in the middle of a communication session with an EC server, a method is needed to stick the flow to the same EC server, 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 server is already finished. Only a very small percentage of registered EC services need to stick to the original server when handover to a new cell tower.

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

This draft describes the BGP UPDATE to enable ingress routers to take the App Server load, the capacity index, and the location preference into consideration when computing the optimal path to the egress routers.

1.4. Problem #2: Unbalanced Anycast Distribution due to UE Mobility

Usually, higher capacity EC servers are placed in a metro data center to accommodate more UEs in the proximity needing the services, and fewer are placed in remote sites. When there is a special event occurring at a remote site for a short period, e.g., 1~2 days, the EC servers in the remote site might be heavily utilized. In contrast, the EC servers of the same app in the metro DC can be very underutilized. Since the condition can be short-lived, it might not make business sense to adjust EC capacity among DCs. Sometimes, UEs swarming to a specific site are not anticipated.

Internet-Draft AppMetaData for 5G EC Service

1.5. Problem 3: Application Server Relocation

When an EC server is added to, moved, or deleted from a 5G EC Data Center, the routing protocol needs to propagate the changes to 5G PSA or the PSA adjacent routers. After the change, the cost associated with the site might change as well.

Note: for ease of description, the Edge Application Server and Application Server are used interchangeably throughout this document.

[2.](#) Conventions used in this document

A-ER: Egress Router to an Application Server, [A-ER] is used to describe the last router that the Application Server is attached. For a 5G EC environment, the A-ER can be the gateway router to a (mini) Edge Computing Data Center.

Application Server: An application server is a physical or virtual server that hosts the software system for the application.

Application Server Location: Represent a cluster of servers at one location serving the same Application. One application may have a Layer 7 Load balancer, whose address(es) are reachable from an external IP network, in front of a set of application servers. From an IP network perspective, this whole group of servers is considered as the Application server at the location.

Edge Application Server: used interchangeably with Application Server throughout this document.

EC: Edge Computing

Edge Hosting Environment: An environment providing the support required for Edge Application Server's execution.

Internet-Draft AppMetaData for 5G EC Service

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

Edge DC: Edge Data Center, which provides the Edge Computing Hosting Environment. An Edge DC might host 5G core functions in addition to the frequently used application servers.

gNB next generation Node B

L-DN: Local Data Network

PSA: PDU Session Anchor (UPF)

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.](#) Usage of AppMetaData for 5G Edge Computing

AppMetaData consists of metrics about the running environment at the egress routers to which EC servers are directly attached.

3.1. Assumptions

From the IP Layer, the EC servers or their respective load balancers are identified by their IP addresses. Those IP addresses are the identifiers to the EC servers throughout this document. 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 capacity metrics for optimal forwarding.

Internet-Draft

AppMetaData for 5G EC Service

- The 5G EC controller or management system can send those EC service identifiers to relevant routers.
- The ingress routers' local BGP path compute algorithm includes a special plugin that can compute the path to the optimal Next Hop (egress router) based on the BGP AppMetaData TLV received for the registered EC services.

The proposed solution is for the egress routers, a.k.a. A-ERs in this document, that have direct links to the EC Servers to collect various measurements about the Servers' running status [[5G-EC-Metrics](#)] and advertise the metrics to other routers in 5G EC LDN (Local Data Network).

3.2. IP Layer Metrics to Gauge Application Behavior

[5G-EC-Metrics] describes the IP Layer Metrics that can gauge the application servers running status and environment:

- IP-Layer Metric for App Server Load Measurement:
The Load Measurement to an App Server is a weighted combination of the number of packets/bytes to the App Server and the number of packets/bytes from the App Server which are collected by the A-ER to which the App Server is directly attached.
The A-ER is configured with an ACL that can filter out the packets for the Application Server.
- Capacity Index
a numeric number, configured on all A-ERs in the domain consistently, is used to represent the capacity of the application server attached to an A-ER. At some sites, the IP address exposed to the A-ER is the App Layer Load balancer that have many instances attached. At other sites, the IP address exposed is the server instance itself.
- Site preference index:
is used to describe some sites are more preferred than others. For example, a site with higher bandwidth has a higher preference number than other.

In this document, the term "Application Server Egress Router" [A-ER] is used to describe the last router that an Application Server is attached. For the 5G EC environment, the A-ER can be

Internet-Draft AppMetaData for 5G EC Service

the gateway router to the EC DC where multiple Application servers are hosted.

3.3. AppMetaData Constrained Optimal Path Selection

The main benefit of using ANYCAST is to leverage the network layer conditions to select an optimal path to the application instantiated in multiple locations.

When the ingress routers to the 5G LDN are informed of the Load and Capacity Index of the App Servers at different EC data centers, they can incorporate those metrics with the network path conditions for path selection.

Here is an algorithm that computes the cost to reach the App Servers attached to Site-i relative to another site, say Site-b. When the reference site, Site-b, is plugged in the formula, the cost is 1. So, if the formula returns a value less than 1, the cost to reach Site-i is less than reaching Site-b.

$$\text{Cost-i} = (w * \frac{\text{CP-b} * \text{Load-i}}{\text{CP-i} * \text{Load-b}}) + (1-w) * \frac{\text{Pref-b} * \text{Network-Delay-i}}{\text{Pref-i} * \text{Network-Delay-b}})$$

Load-i: Load Index at Site-i, it is the weighted combination of the total packets or/and bytes sent to and received from the Application Server at Site-i during a fixed time period.

CP-i: capacity index at Site-i, a higher value means higher capacity.

Delay-i: Network latency measurement (RTT) to the A-ER that has the Application Server attached at the site-i.

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

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.

[4.](#) BGP Protocol Extension to advertise Load & Capacity

The goal of the BGP extension is for egress routers to propagate the metrics about their running environment to ingress routers. Here are some examples of the metrics propagated by the egress routers:

- the Load Measurement Index for the attached EC Servers,
- the Capacity Index, and
- Site Preference Index.

This section specifies the Load Index Sub-TLV, Capacity Sub-TLV, and the Site Preference Sub-TLV that can be carried by the Tunnel Encap Path Attribute associated with the routes.

4.1. Ingress Node BGP Path Selection Behavior

4.1.1. AppMetaData Influenced BGP Path Selection

When an ingress router receives BGP updates for the same IP address from multiple egress routers, all those egress routers are considered the next hops for the IP address. For the selected EC services, the ingress router's BGP engine would call a Plugin function that can select paths based on the AppMetaData received. The Plugin function is called Load Compute Engine throughout this document.

Assume that both Ra and Rb in Figure-1 have BGP Multipath enabled. As a result, Dst Address: S1:aa08::4450 is resolved via multiple NextHop: R1, R2, R3.

Suppose the local BGP's Load Compute Engine identifies R1 as the optimal NextHop for the flow towards S1:aa08::4450. Then the Load Compute Engine can insert a higher weight for the path R1 so that BGP Best Path is locally influenced by the weight parameter based on the local decision.

4.1.2. Ingress Router Forwarding Behavior

When the ingress router receives a packet and lookup the FIB, it gets the destination prefix's whole path. It encapsulates the packet destined towards the optimal egress node.

For subsequent packets belonging to the same flow, the ingress router needs to forward them to the same egress router unless

the selected egress router is no longer reachable. Keeping packets from one flow to the same egress router, a.k.a. Flow Affinity, is supported by many commercial routers. Most registered EC services have relatively short flows.

How Flow Affinity is implemented is out of the scope for this document. Here is one example to illustrate how Flow Affinity can be achieved. This illustration is not to be standardized.

For the registered EC services, the ingress node keeps a table of

- Service ID (i.e., IP address)
- Flow-ID
- Sticky Egress ID (egress router loopback address)
- A timer

The Flow-ID in this table is to identify a flow, initialized to NULL. How Flow-ID is constructed is out of the scope for this document. Here is one example of constructing the Flow-ID:

- For IPv6, the Flow-ID can be the Flow-ID extracted from the IPv6 packet header with or without the source address.
- For IPv4, the Flow-ID can be the combination of the Source Address with or without the TCP/UDP Port number.

The Sticky Egress ID is the egress node address for the same flow. [5G-Sticky-Service] describes several methods to derive the Sticky Egress ID.

The Timer is always refreshed when a packet with the matching EC Service ID (IP address) is received by the node.

If there is no Stick Egress ID present in the table for the EC Service ID, the forwarding plane can select a NextHop

influenced by the Load Compute Engine. The forwarding plane encapsulates the packet with a tunnel to the chosen NextHop. The chosen NextHop and the Flow ID are recorded in the EC Service table entry.

When the selected optimal NextHop (egress router) is no longer reachable, refer to [Section 6](#) Soft Anchoring on how another path is selected.

4.1.3. Forwarding Behavior when UEs moving to new 5G Sites

When a UE moves to a new 5G eNB which is anchored to the same UPF, the packets from the UE traverse to the same ingress router. Path selection and forwarding behavior are same as before.

When the new eNB is anchored to a different UPF, the packets from the UE traverse a different ingress router. If the UE source IP address has been changed, indicating the new UPF might belong to a different administrative domain, the new ingress router treats the packets from the UE as a new flow and select the optimal path based on the configured policies. If the UE maintains the same IP address when anchored to a new UPF, the directly connected ingress router might use the pre-computed Egress Router, which is passed from a neighboring router. [\[5G-Edge-Sticky\]](#) describes methods for the ingress router connected to the UPF in the new site to consider the information passed from other ingress routers in selecting the optimal paths. The detailed algorithm is out of the scope of this document.

[5.](#) The Sub-TLVs for AppMetaData

The AppMetaData attribute is encoded in an optional subTLV within the Tunnel Encap [\[RFC9012\]](#) Path Attribute.

All values in the Sub-TLVs are unsigned 32 bits integers.

Internet-Draft AppMetaData for 5G EC Service

[5.1.](#) Load Measurement sub-TLV format

Two types of Load Measurement Sub-TLVs are specified. One is to carry the aggregated cost Index based on a weighted combination of the collected measurements; another one is to carry the raw measurements of packets/bytes to/from the App Server address. The raw measurement is useful when ingress routers have embedded analytics relying on the raw measurements.

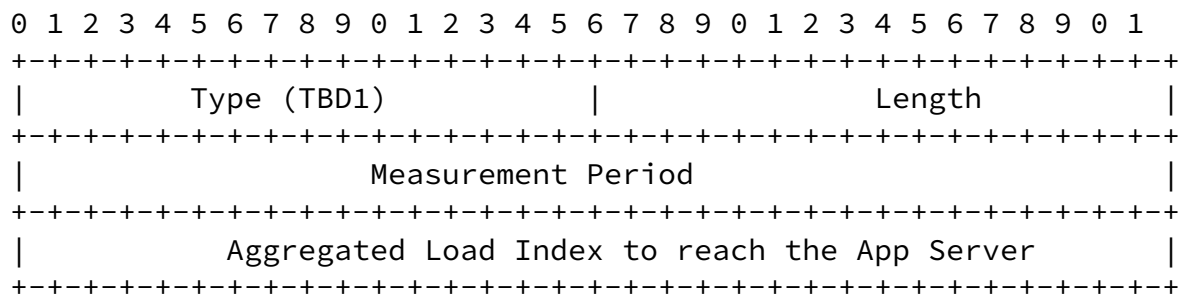
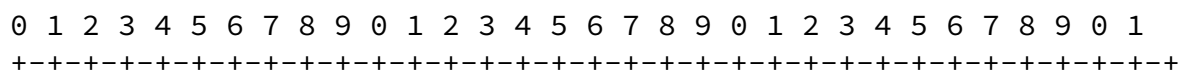


Figure 2: Aggregated Load Index Sub-TLV

Raw Load Measurement sub-TLV has the following format:



Type (TBD2)	Length
Measurement Period	
total number of packets to the AppServer	
total number of packets from the AppServer	
total number of bytes to the AppServer	
total number of bytes from the AppServer	

Figure 3: Raw Load Measurement Sub-TLV

Type =TBD1: Aggregated Load Measurement Index derived from the Weighted combination of bytes/packets sent to/received from the App server:

$$\text{Index} = w_1 * \text{ToPackets} + w_2 * \text{FromPackets} + w_3 * \text{ToBytes} + w_4 * \text{FromBytes}$$

Where w_i is a value between 0 and 1; $w_1 + w_2 + w_3 + w_4 = 1$.

Type= TBD2: Raw measurements of packets/bytes to/from the App Server address.

Measure Period: BGP Update period or user-specified period.

5.2. Capacity Index sub-TLV format

The Capacity Index sub-TLV has the following format:

Type (TBD3)	Length
Capacity Index	

Figure 4: Capacity Index Sub-TLV

Note: "Capacity Index" can be more stable for each site. If those values are configured to nodes, they might not need to be included in every BGP UPDATE.

5.3. The Site Preference Index sub-TLV format

The site Preference Index is used to achieve Soft Anchoring [Section 5] an application flow from a UE to a specific location when the UE moves from one 5G site to another.

The Preference Index sub-TLV has the following format:

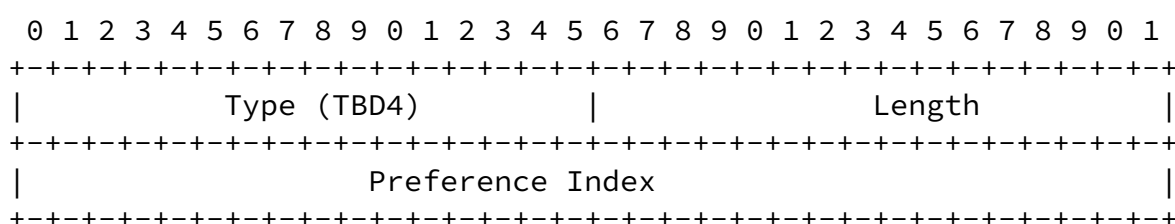


Figure 5: Preference Index Sub-TLV

Note: "Site Preference Index" can be more stable for each site. If those values are configured to nodes, they might not need to be included in every BGP UPDATE.

6. AppMetaData Propagation Scope

AppMetaData is only to be distributed to the relevant ingress nodes of the 5G EC local data networks. Only the ingress routers that are configured with the 5G EC services need to receive the AppMetaData for specific Service IDs.

For each registered EC service, a corresponding filter group can be formed on RR to represent the interested ingress routers that are interested in receiving the corresponding AppMetaData information.

7. Minimum Interval for Metrics Change Advertisement

As the metrics change can impact the path selection, the

Minimum Interval for Metrics Change Advertisement is configured to control the update frequency to avoid route oscillations. Default is 30s.

Significant load changes at EC data centers can be triggered by short-term gatherings of UEs, like conventions, lasting a few hours or days, which are too short to justify adjusting EC server capacities among DCs. Therefore, the load metrics change rate can be in the magnitude of hours or days.

8. Soft Anchoring of an ANYCAST Flow

"Sticky Service" in the 3GPP Edge Computing specification (3GPP TR 23.748) is about flows from a UE sticking to a specific location when the UE moves from one 5G Site to another.

"Soft Anchoring" is a mechanism for ingress routers to apply preference to the path towards the previous server location when the UE is anchored to a new UPF and continue using its cached IP for the EC server.

Let's assume one application "App.net" is instantiated on four servers that are attached to four different routers R1, R2, R3, and R4 respectively. It is desired for packets to the "App.net" from UE-1 to stick with one server, say the App Server attached to R1, even when the UE moves from one 5G site to another. However, when there is a failure reaching R1 or the Application Server attached to R1, the packets of the flow "App.net" from UE-1 need to be forwarded to the Application Server attached to R2, R3, or R4.

We call this kind of sticky service "Soft Anchoring", meaning that anchoring to the site of R1 is preferred, but other sites can be chosen when the preferred site encounters a failure.

Here is a mechanism to achieve Soft Anchoring:

- Assign a group of ANYCAST addresses to one application. For example, "App.net" is assigned with 4 ANYCAST addresses, L1, L2, L3, and L4. L1/L2/L3/L4 represents the location preferred ANYCAST addresses.
- For the App.net Server attached to a router, the router

has four Stub links to the same Server, L1, L2, L3, and L4 respectively. The cost to L1, L2, L3, and L4 is assigned differently for different egress routers. For example,

- o When attached to R1, the L1 has the lowest cost, say 10, when attached to R2, R3, and R4, the L1 can have a higher cost, say 30.
 - o ANYCAST L2 has the lowest cost when attached to R2, higher cost when attached to R1, R3, R4 respectively.
 - o ANYCAST L3 has the lowest cost when attached to R3, higher cost when attached to R1, R2, R4 respectively, and
 - o ANYCAST L4 has the lowest cost when attached to R4, higher cost when attached to R1, R2, R3 respectively
- When a UE queries for the "App.net" for the first time, the DNS reply has the location preferred ANYCAST address, say L1, based on where the query is initiated.
 - When the UE moves from one 5G site-A to Site-B, UE continues sending packets of the "App.net" to ANYCAST address L1. The routers will continue sending packets to R1 because the total cost for the App.net instance for ANYCAST L1 is lowest at R1. If any failure occurs making R1 not reachable, the packets of the "App.net" from UE-1 will be sent to R2, R3, or R4 (depending on the total cost to reach L1 attached to R2/R3/R4).

If the Application Server supports the HTTP redirect, more optimal forwarding can be achieved.

- When a UE queries for the "App.net" for the first time, the global DNS reply has the ANYCAST address G1, which has the same cost regardless of where the Application servers are attached.
- When the UE initiates the communication to G1, the packets from the UE will be sent to the Application Server that has the lowest cost, say the Server attached to R1. The Application server is instructed with HTTPs Redirect to reply with a location-specific URL, say App.net-Loc1. The client on the UE will query the DNS

for App.net-Loc1 and get the response of ANYCAST L1. The subsequent packets from the UE-1 for App.net are sent to L1.

9. Manageability Considerations

To be added.

10. Security Considerations

To be added.

11. IANA Considerations

Here are new Sub-TLV types requiring IANA registration:

Type = TBD1: Aggregated Load Measurement Index derived from the Weighted combination of bytes/packets sent to/received from the App server.

Type = TBD2: Raw measurements of packets/bytes to/from the App Server address.

Type = TBD3: Capacity value sub-TLV

Type = TBD4: Site preference value sub-TLV

12. References

12.1. 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>>.

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

12.2. Informative References

[3GPP-EdgeComputing] 3GPP TR 23.748, "3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Study on enhancement of support for Edge Computing in 5G Core network (5GC)", Release 17 work in progress, Aug 2020.

[5G-EC-Metrics] L. Dunbar, H. Song, J. Kaippallimalil, "IP Layer Metrics for 5G Edge Computing Service", [draft-dunbar-ippm-5g-edge-compute-ip-layer-metrics-00](#), work-in-progress, Oct 2020.

[5G-Edge-Sticky] L. Dunbar, J. Kaippallimalil, "IPv6 Solution for 5G Edge Computing Sticky Service", [draft-dunbar-6man-5g-ec-sticky-service-00](#), work-in-progress, Oct 2020.

[RFC5521] P. Mohapatra, E. Rosen, "The BGP Encapsulation Subsequent Address Family Identifier (SAFI) and the BGP Tunnel Encapsulation Attribute", April 2009.

[BGP-SDWAN-Port] L. Dunbar, H. Wang, W. Hao, "BGP Extension for SDWAN Overlay Networks", [draft-dunbar-idr-bgp-sdwan-overlay-ext-03](#), work-in-progress, Nov 2018.

[SDWAN-EDGE-Discovery] L. Dunbar, S. Hares, R. Raszuk, K. Majumdar, "BGP UPDATE for SDWAN Edge Discovery", [draft-dunbar-idr-sdwan-edge-discovery-00](#), work-in-progress, July 2020.

[Tunnel-Encap] E. Rosen, et al "The BGP Tunnel Encapsulation Attribute", [draft-ietf-idr-tunnel-encaps-10](#), Aug

2018.

13. Acknowledgments

Acknowledgements to Donald Eastlake for their review and contributions.

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

Authors' Addresses

Linda Dunbar
Futurewei
Email: ldunbar@futurewei.com

Kausik Majumdar
CommScope
350 W Java Drive, Sunnyvale, CA 94089
Email: kausik.majumdar@commscope.com

Haibo Wang

Huawei
Email: rainsword.wang@huawei.com

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