Distribute Service Metric by BGP

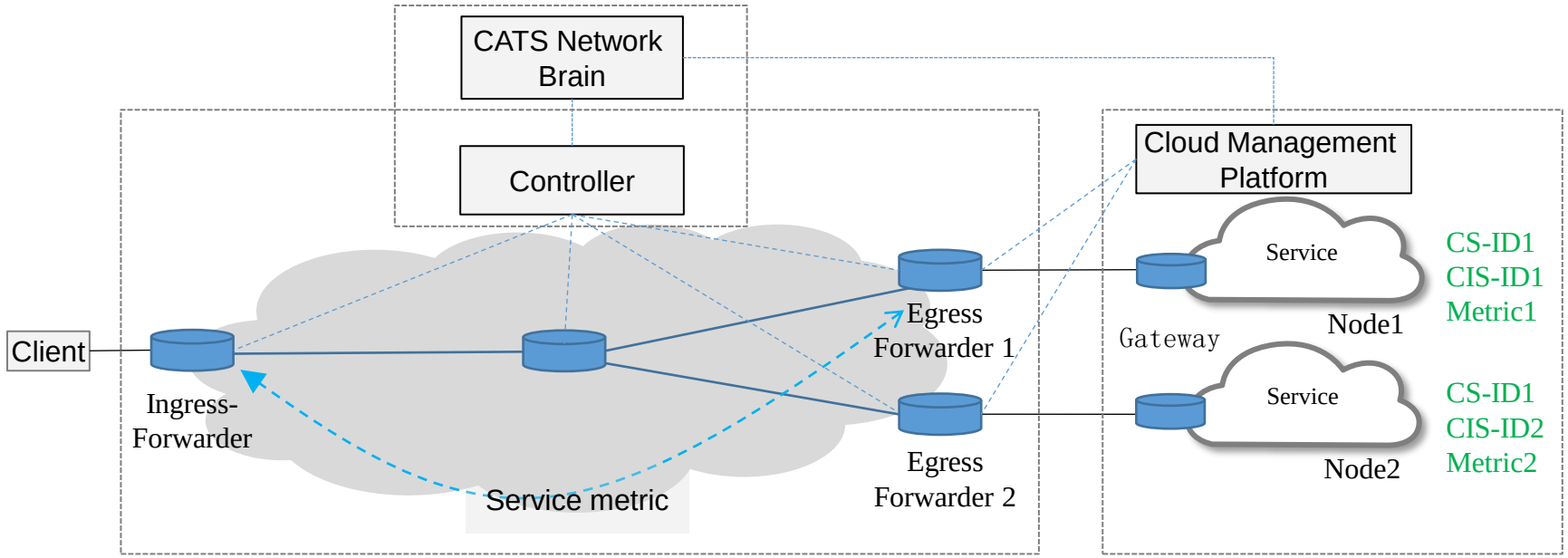
draft-lin-idr-distribute-service-metric-02

Changwang Lin(New H3C Technologies)

Huijuan Yao(China Mobile)

IETF 120

Background



- As described in draft-ietf-cats-framework-02, a C-SMA is required to collect service metrics in CATS network.
- C-SMA can be located in the CATS network brain or in the egress forwarder.
- Centralized model: When C-SMA can be located in the CATS network brain, The service metric is centrally controlled by the CATS network brain, and the generated CATS Route is sent to the device through the controller.
- Distributed model: When C-SMA can be located in the egress forwarder, CATS route with metrics is generated and published in the distributed network.
- CATS Routing at the ingress-forwarder should be based on both network and service metrics.

How to Distribute Service Metrics?

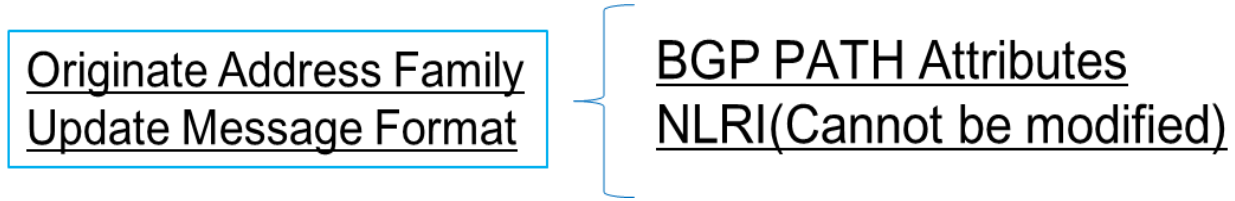
- API Interface
  - Transmitting service metric information through RESTful、gRPC、YANG protocols.
  - Suitable for centralized models, where interaction occurs between the controller and network devices.
- Service Discovery and Registration Center
  - When a service starts, it registers its compute service information to a central registry, and other services retrieve the latest information from this center, such as LISP.
  - Lacks mature applications.
- Routing Protocol
  - Using routing protocols to transmit compute service metric information. These protocols can be part of existing network routing protocols (e.g., BGP, OSPF) or custom proprietary protocols.
  - Extending BGP to include compute service metric information can be achieved by adding new BGP attributes or extensions to support the transmission of compute routing information.
  - BGP is a mature and reliable routing protocol with well-established RR (Route Reflector) mechanisms for information transmission.

Through the use of BGP extensions, it can effectively support the dissemination of service metric information.

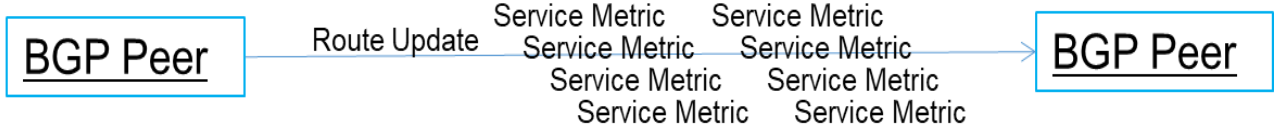
This document will discuss how to extend the BGP protocol to support the distribution of service metric information.

Why Do We Need a New BGP Address Family?

① Scalability
- CSID can only be an anycast address
- Service metrics can only be expanded through attributes



② Isolation
- Periodic updates affect regular route updates
- Cannot eliminate the impact on existing address family



③ Performance
- Differences in attributes result in smaller Update packets
- Unable to flexibly filter out unnecessary service metrics

BGP Update 1	BGP Update 2	BGP Update N
PATH:Service Metric 1	PATH:Service Metric 2	PATH:Service Metric n
NLRI : CSID 1	NLRI : CSID 2	NLRI : CSID n

④ Deployment Issues
- Requires enabling the add-path feature, where multiple paths need to be sent for a single prefix
- Needs to control the service metric propagation domain

Extending BGP with a New Service Metric Address Family

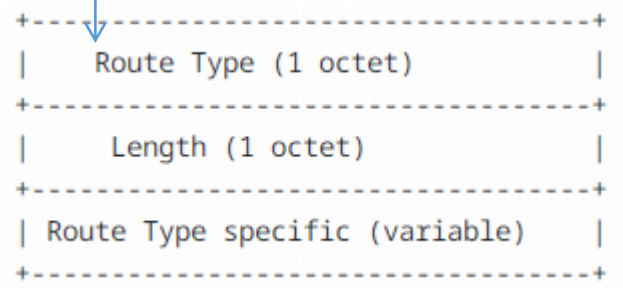
Design Goals

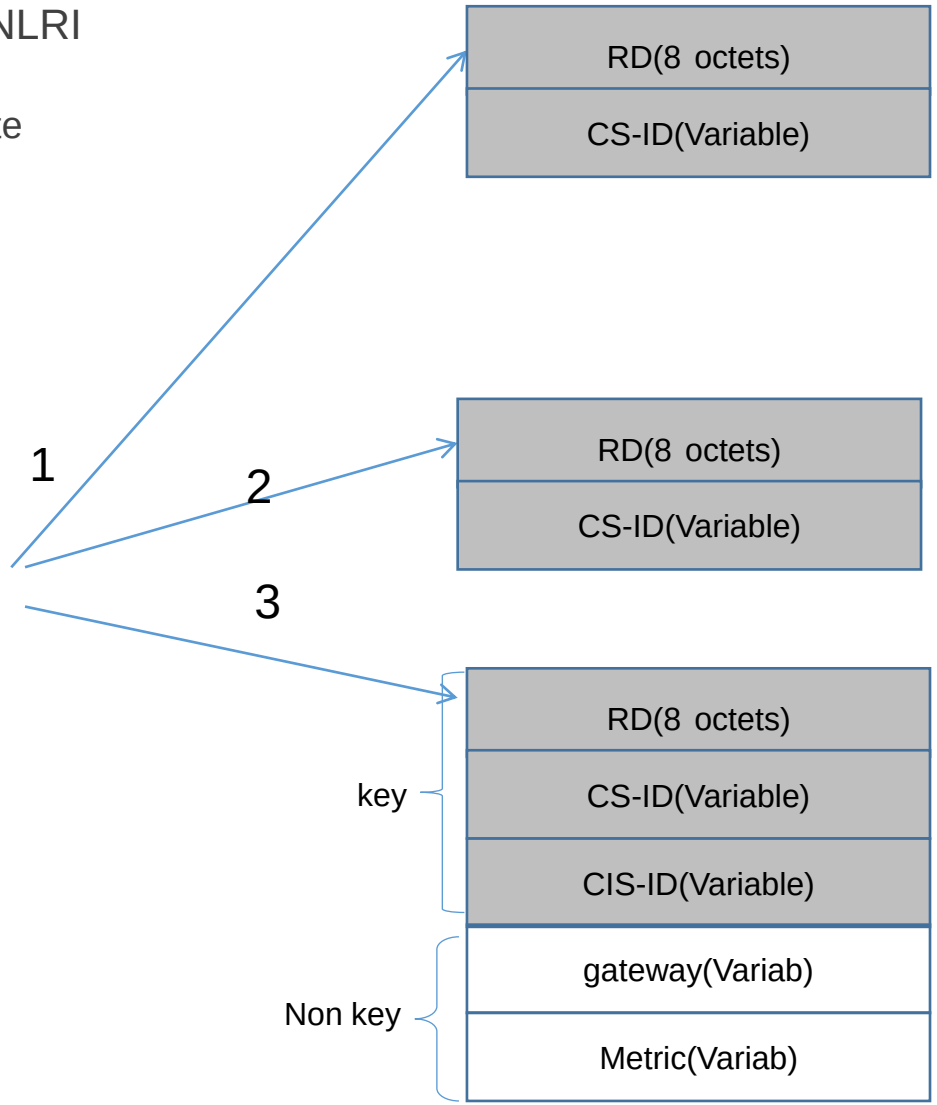
- Minimize the impact on existing route propagation.
- Service CS-ID identifiers are not limited to Anycast IP addresses.
- Improve the efficiency of sending periodic service metric updates.
- Send on-demand and control the propagation scope of service metrics.

	Originate Address Family	New Servcie metric Address Family
NLRI	Can not be modifed	Define flexible variable-length NLRIs
Path Attribute	Scalable	Scalable
Isolation	Lack of Isolation	Good isolation. The processing of different address spaces does not interfere with each other
Placement of Metric	Can only be in path attributes	It can be in Path attributes as well as in NLRI, offering good scalability.
Scalability	Lack of scalability	Strong scalability with flexible NLRI expansion, allowing NLRI to distinguish between key and no-key, facilitating the extension of processing mechanisms such as on-demand subscription mechanisms.
On-demand subscription mechanism	Not supported	Support. Implement on-demand subscription functionality to enable the on-demand transmission of service metrics.
Performance	Lack of Performance	Good performance.Set the next hop for service metrics routing to 0, and place the service-associated location address in the non-key part of NLRI, which also improve the packageability of routes.
Deployment	restrictions	Can be deployed independently

Extensible NLRI

- The format of the Service Metric NLRI
  - 1 - Service Metric Register route
  - 2 - Service Metric Subscribe route
  - 3 - Service Metric Update route





Flexible Service Metric Subscription

Register

- Egress send registration route to RR
- RR forwards the registration route to all Ingress nodes.

Subscribe

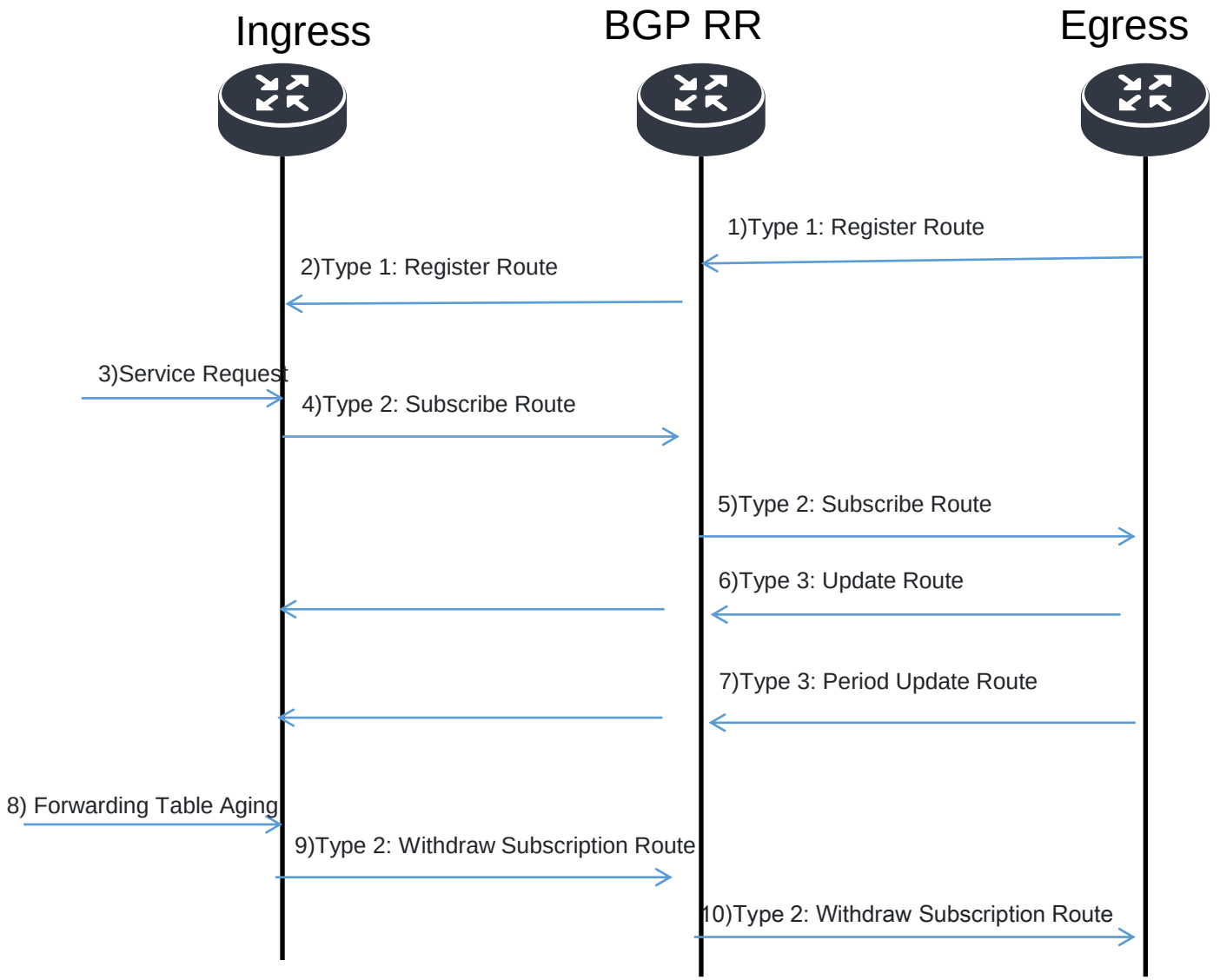
- Ingress receives a service request and sends a subscription route to the RR.
- The RR forwards the subscription route to the corresponding Egress nodes.
- Egress responds with an updated route to Ingress..

Period Update

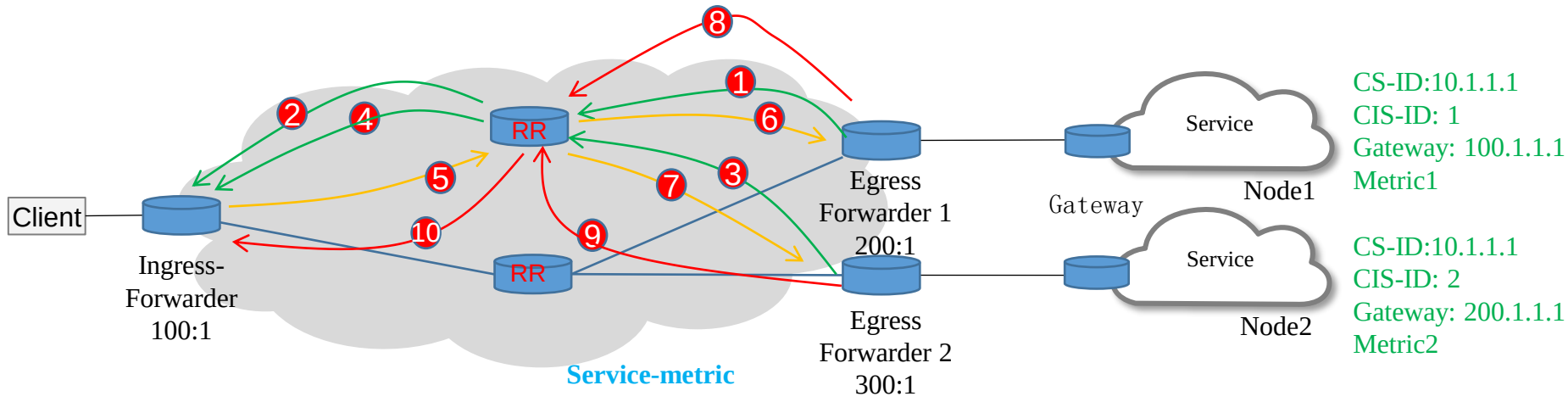
- Egress send periodic Update service metric route to subscribers.

UnSubscribe

- Upon forwarding table aging, send a withdraw subscription route to RR.
- RR notifies corresponding Egress of the route withdrawal.
- Egress no longer sends periodic route updates to the Ingress.



Use case



1 2 Type 1(Register) [RD: 200:1, CS-ID:10.1.1.1]

3 4 Type 1(Register) [RD: 300:1, CS-ID:10.1.1.1]

5 6 7 Type 2(Subscribe) [RD: 100:1, CS-ID:10.1.1.1]

8 Type 3(Update) [RD: 200:1, CS-ID:10.1.1.1, CIS-ID: 1,
Gateway: 100.1.1.1, Metric 1]

9 Type 3(Update) [RD: 300:1, CS-ID:10.1.1.1, CIS-ID: 2,
Gateway: 200.1.1.1, Metric 2]

10 Type 3(Update) [RD: 200:1, CS-ID:10.1.1.1, CIS-ID: 1,Gateway: 100.1.1.1, Metric 1]
[RD: 300:1, CS-ID:10.1.1.1, CIS-ID: 2, Gateway: 200.1.1.1, Metric 2]

Note: RR will bundle multiple prefixes (NLRIs) with the same attributes
into a single Update packet.

Next Steps

- Any questions or comments are Welcomed.

Thanks

9