Computing-Aware Networking(CAN)

draft-liu-dyncast-ps-usecases-04
draft-liu-dyncast-gap-reqs-00
draft-li-dyncast-architecture-04
draft-liu-can-computing-resource-modeling-00

P. Liu, China Mobile

Motivations – Why CAN?

Computing-Aware Networking (CAN) aims at computing and network resource optimization by steering traffic to appropriate computing resources, considering not only networking metric but also computing resource metric and service affiliation.

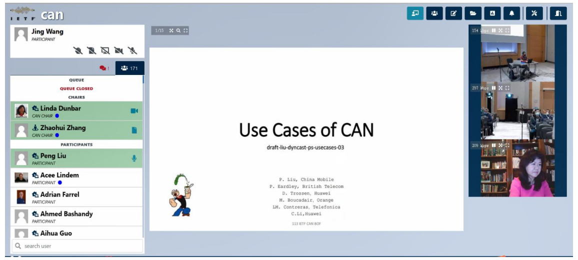
CAN is to solve the problem of “ the ‘closest’ is not the ‘best’”, providing the best user experience of low latency, high reliability and stable services experience when moving across different areas, based on the Increasing development of integrated computing and networking infrastructure, including CDN and edge computing.

CAN BoF at IETF 113

A non-wg forming BoF with178 people participated in the discussion

Use cases:

- The deployment of edge computing and user mobility show the necessity of “steering traffic among different edge sites” .
- Majority of modern applications have multiple instances instantiated at many locations and have short flows
- Many apps, such as AR/VR typically requiring 20ms E2E delay, need "considering more factors when steering traffic ”



Linda Dunbar CAN CHAIR
Zhaohui Zhang CAN CHAIR
Peng Liu PARTICIPANT
Acee Lindem PARTICIPANT
Adrian Farrel PARTICIPANT
Ahmed Bashandy PARTICIPANT
Aihua Guo

Use Cases of CAN

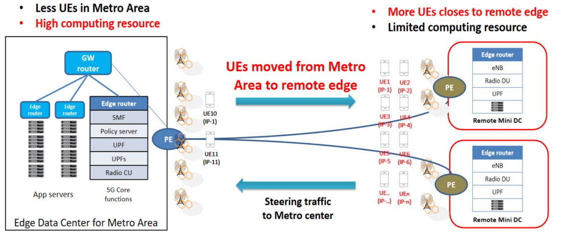
draft-liu-dyncast-ps-usecases-03

P. Liu, China Mobile
P. Eardley, British Telecom
D. Trossen, Huawei
M. Boucadair, Orange
LM. Contreras, Telefonica
C.Li,Huawei

113 IETF CAN BOF

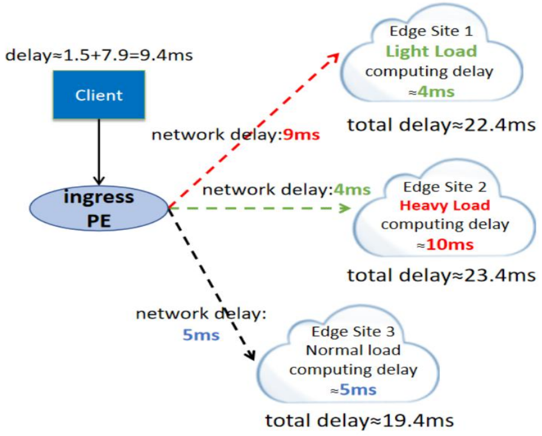
Gap analysis:

- DNS was not designed for dynamic scenarios with fast changes of serving service instance
  - •DNS caching and learning computing status are additional problems
- Load balancer should learn about the network status, while as an off-path approach, it also adds latency and bottleneck



Potential Solutions:

- **Dynamic anycast:** edge routers get the computing metrics and select the service instance
- **On-path load balancer:** centralized/decentralized LBs coordinates, also learns the network status



Conclusion:

- The BoF was successful in agreeing that the **use cases are important**.
- **Study existing solutions**, e.g. ALTO. Please continue pay attention to reusing the existing work.
- Load Balancer/D-router or Dyncast have own characteristics and could **be further investigated**.
- The next hop of this work (WG to land in) is **still not clear**, please continue working.

Issues and Responses

36 issues raised from the BoF discussion; Lots of discussions on dyncast@ietf.org and rtgwg@ietf.org

Summary of key issues(https://github.com/CAN-IETF/CAN-BoF-ietf113/issues):

What is the relationship between CAN and service deployment, discovery and the upper layer protocol? #19 #21 #23

The whole process includes service deployment, service discovery and service traffic steering. CAN assumes the services have been deployed and discovered in multi-edge sites, aiming at the traffic steering. One the one hand, CAN should coordinate with upper layer for service deployment and discovery. On the other hand, the computing metrics collection in CAN will also benefit service deployment and discovery.

What is the relation between CAN and ALTO? #5

ALTO solves the problem of service instance selection as an off-path solution, which can be seen as an alternative way of addressing the problem space of CAN at the Application Layer. So in that respect, even targeting a common problem, both provide different approaches, then imposing different needs but also taking different assumptions on how applications and networks interact. Off-path systems, e.g., ALTO, replies for applications/services before traffic delivery might not be optimal or valid after the handover. So, more details are needed of ALTO including some extension to support multi-deployment, quick interaction, integrate more performance metric information.

Why compute resource is different? Is it measurable? Per app per metrics? #8 #10 #14 #27 #34

From the network-based perspective, it could be found that compute resources are also often measured in terms similar to network resources (e.g., request throughput, request completion times). However, compute resource are heterogeneous to have different kind of chips, hardwares, components, etc. and the usage of compute resources are often more service-specific than the sharing of network resources over a possibly large number of services.Hence, decisions on which compute resources to use out of possibly many needs to take service-specific compute resource metrics into account, in addition to the network resources being used.

How to use the measurement would be solution related. See IFIP Networking 2022 paper on how to simply expose "computing capability" and achieve better steering with such simple measure.

Will CAN dyncast impact each and every router? #29 #30 #28

No. CAN doesn't impact every router.

Will CAN compute paths? #32

CAN does not compute path directly, it selects computing endpoints as well as the network path. For selecting path, CAN is aware of the network path status and coordinates with the existing routing/path solutions.

New version of Drafts

New versions of *problem statement*, *use cases*, *gap analysis* and *requirements* were posted, following the BoF discussion and addressing some of the issues:

•**draft-liu-dyncast-ps-usecases-04**

- Updated the problem statement and use cases (https://datatracker.ietf.org/doc/draft-liu-dyncast-ps-usecases/)

•**draft-liu-dyncast-gap-reqs-00**

- Replacing previous version by adding the gap analysis and updating requirements (https://datatracker.ietf.org/doc/draft-liu-dyncast-gap-reqs/)

•**draft-li-dyncast-architecture-04**

- Addressed some issues in the BoF (https://datatracker.ietf.org/doc/draft-li-dyncast-architecture/)

New draft of computing resource modeling,

•**draft-liu-can-computing-resource-modeling-00**

- describes possible modeling of compute resource and usage in CAN (https://datatracker.ietf.org/doc/draft-liu-can-computing-resource-modeling/)

Next Steps

- CAN will focus on routing area related work.
- Rename the drafts to ‘CAN’？
- Request a CAN mailing list?
- Refine the drafts.
- Apply for another CAN BoF.

- Welcome more comments and suggestions.

Thank you!