**## Name**

Neotec (Network Operations in Telecom Cloud)

Note: After the BoF a different name may be chosen for the potential WG.

(Suggest to keep the name Neotec at least for the BoF, this can make good use of the Neotec mailing list and the influence of the past side meetings. If we find a better name for the WG, a new mailing list can be created and the participants will be redirected)

**## Description**

In the rapidly evolving landscape of cloud computing, the performance of cloud services delivery of Telecom Cloud Service Providers (TCSPs) heavily relies on the efficient management and coordination between the cloud infrastructures. In particular, there are challenges in coordinating the management of the various telecom edge clouds, where services are hosted, and the network infrastructures that interconnect them.

**Telecom edge cloud** delivers compute processing, storage, networks (e.g., fixed or mobile access) and other computing resources closer to the customer locations (e.g. customer premises or adjacent data centers (DCs)) to fulfill Service Level Objectives (SLOs) of user applications and connectivity. However, the telecom edge clouds could have different types of processing resources (CPU, GPU, FPGA, etc.), and be dimensioned with less resource capacity compared to a conventional DCs. Also, **cloud and network resources are typically managed independently**, which poses some challenges:

1. Without visibility into network SLOs status and available resources, cloud orchestrators struggle to ensure that the deployment of service function instances in edge DCs meets the customer network's strict SLO requirements, e.g., SD-WAN and SASE functions, and AI/ML algorithms applications.
2. When service function instances are scaled up or down, relocated, or when traffic matrices between service functions change, the network remains unaware and is unable to promptly adjust its resources to accommodate these changes.
3. The network is unaware of the traffic characteristics of telecom edge cloud applications, which might reduce service performance. Take the AI-cluster training case an example, without obtaining some flow characteristics (e.g., IP addresses, volume, or timeframe) from service orchestrators, the overall capability of the network is constrained by the slowest flow, which leads to uneven load distribution and low network throughput.

The Neotec WG focuses on introducing a **cloud-aware service orchestrator** and investigating the **interfaces** between the cloud-aware service orchestrators and existing network controllers, as well as cloud orchestrators. Specifically, these interfaces support the following functions:

* 1. Dynamical adjustment of load-balancing policies: Effective load-balancing policy is essential to guarantee the end-to-end performance and maximize the overall throughput for certain flows among edge clouds during a specific time span.
  2. Edge cloud resource abstraction and exposure: Exposing the cloud metrics to the network to assist the network in making decisions. For example, the computing metrics collected from cloud systems (e.g. Kubernetes or OpenStack) are transformed into ones that can be understood by the network.
  3. Inter-edge cloud network functions and status interfaces exposure: For example, expose the network topology and connectivity performance status to the cloud service orchestrators for service placement selection and deployment, such as the status of service functions associated with VPNs, as well as the status of SD-WAN and SASE.

With this interface, distinct approaches existing in both compute and network domains can be reconciled.

As TCSPs often rely on equipment from multiple vendors, standardized and interoperable solutions are essential to seamlessly integrate cloud and network resources.

The first use case is service function chaining (SFC) in the cloud. SFC, such as virtual firewall services, requires dynamic placement of service instances in the clouds, and there can be multiple instances of the same service function distributed across DCs at different locations. The performance of SFC services provided to a customer relies on both the availability of computing resources in the cloud, and the topology and congestion condition of networks among the DCs to build optimized service function paths.

One more use case is the cross-DC scheduling of computing and storage resources. Given the constraints of computing resources and storage costs, it iss crucial to facilitate storage migration, analysis, and processing of large-scale data across various DCs. By leveraging the real-time perception of the cloud-computing network resource states provided by the cloud, TCSP's network controller can dynamically allocate storage and computing resources via Metro Area Network (MAN). This capability allows for flexible cross-DC storage, data modeling, and AI training, ensuring optimal service quality and efficient resource utilization.

Another use case is that Machine Learning (ML) and Federated ML applications in 5G and beyond demand massive computing resources which may spread in multiple Cloud DCs, and it also relies on the network to provide on-demand connections with required bandwidth and latency. The performance and efficiency of such application can be improved by dynamic coordination between cloud-aware service orchestrator and network controller for optimized computing resources utilization and network throughput.

The Neotec BoF will discuss several use cases where better coordination between network controllers and service orchestrators is needed, enabling the exchange of resource, attributes, status, requirements and policy between these domains. It will analyze the gaps in existing IETF works for the coordination between network and clouds in management and operation, and hopefully identifies the potential work needed in IETF.

Neotec aims also to serve as a platform for the industry to exchange requirements, challenges, and experiences related to coordinated network operations for cloud-based services.

## Required Details

- Status: Non-WG Forming

- Responsible AD name: Mahesh Jethanandani

- BOF proponents: Chongfeng Xie (chongfeng.xie@foxmail.com), Luis Contreras ([luismiguel.contrerasmurillo@telefonica.com](mailto:luismiguel.contrerasmurillo@telefonica.com)), Gyan Mishra (hayabusagsm@gmail.com)

- Number of people expected to attend: 100

- Length of session (1 or 2 hours): 2 hours

- Conflicts (whole Areas and or WGs)

- Chair Conflicts TBD

- Technology Overlap: OPSAWG, NMOP, CATS, TEAS

- Key Participant Conflict: Chongfeng Xie, Luis Contreras, Linda Dunbar

## Information for IAB/IESG

To allow evaluation of your proposal, please include the following items

- Any protocols or practices that already exist in this space

OPSAWG has already defined connection interfaces such as Attachment Circuits (AC) between cloud gateways and network edge devices (draft-ietf-opsawg-teas-attachment-circuit), VPN service models between DCs (RFC8299 L3SM, RFC8466 L2SM), and Network Slice models between 5G use cases, as well as interfaces for network and VPN services topology and performance status (RFC 9375), however, none of these cover the traffic flow scheduling policy interface.

- Which (if any) modifications to existing protocols or practices are required

- Which (if any) entirely new protocols or practices are required

New YANG model(s) need be defined for the interface between the cloud-aware service orchestrators and network controllers to ensure that cloud metrics are provided to the network in an appropriate format to enforce connectivity services and offer guarantees.

Currently there are several WGs in IETF which work on topics related to network and cloud coordination:

- CATS WG focuses on the problem of how the network edge can steer traffic between clients of a compute service and sites offering the service. It works on a general framework for the distribution of compute and network metrics and transport of traffic from network edge to service instance, and identifies some common metrics, which will be used for traffic steering at the network edge node. It does not consider the coordination between network and cloud in management and operation.

- TEAS WG is responsible for defining traffic engineering architecture and identifying required related routing and path computation element functions. TEAS is also responsible for standardizing RSVP-TE signaling protocol mechanisms. It takes network capability and information into consideration for traffic engineering in network. It also delivers YANG models in support of traffic engineering. It does not consider the coordination between network and cloud in management and operation, either.

- OPSAWG deals with operational and management topics that are not in scope of an existing ops area working group and do not justify the formation of a new working group. Currently the WG is publishing ACaaS YANG model. This model can be used for the provisioning of ACs before or during computing service deployment to connect a cloud infrastructure to a network service provider network. Although it provides bearer services for communication between Cloud DCs, it does not address the exposure of network resources to the cloud and the real-time environmental status of Service Function instances, which are crucial for making dynamic network path decisions.

- Network Management Operations (NMOP) WG focuses on solving network management problems faced by operators. Currently it discusses operational issues faced by the deployment of existing network management technologies.

## Agenda

- BoF introduction and Administrivia [Chairs] 10 mins

- Use Cases and Problem Statements 48 mins

Case 1: China Telecom （Qiong） 12 mins

Case 2: Telefonica （Luis） 12 mins

Case 3: Alibaba/CU （Davey/Ran）12 mins

Case 4: Verizon（Gyan） 12 mins

- Gap Analysis and potential work in IETF(Bo Wu) 20 mins

- Open Discussion 30 mins

- Conclusion and Next steps [Chairs] 10 mins

- Speaker shuffling time 3 mins

## Links to the mailing list, draft charter if any, relevant Internet-Drafts, etc.

- Mailing List: [neotec@ietf.org](mailto:neotec@ietf.org), https://mailman3.ietf.org/mailman3/lists/neotec.ietf.org/

- Draft charter: TBD

- Relevant Internet-Drafts:

- Use Cases

<https://datatracker.ietf.org/doc/html/draft-zx-neotec-net4cloud-usecase-00>

<https://datatracker.ietf.org/doc/draft-dunbar-neotec-net-adjust-cloud-scaling-01>

<https://datatracker.ietf.org/doc/html/draft-gao-neotec-interface-cnc-00>

One use case draft from Luis

- YANG models:

<https://datatracker.ietf.org/doc/html/draft-llc-teas-dc-aware-topo-model-03>

<https://datatracker.ietf.org/doc/html/draft-dhody-teas-te-traffic-yang-05>

<https://datatracker.ietf.org/doc/draft-ietf-opsawg-teas-attachment-circuit-18>

<https://datatracker.ietf.org/doc/html/draft-ietf-opsawg-teas-common-ac-11>

<https://datatracker.ietf.org/doc/rfc8299/L3SM>

<https://datatracker.ietf.org/doc/rfc8466/L2SM>