

A Use Case of Network Operation for Telecom Cloud

Cong Li

China Telecom

Yue Zhang

China Telecom

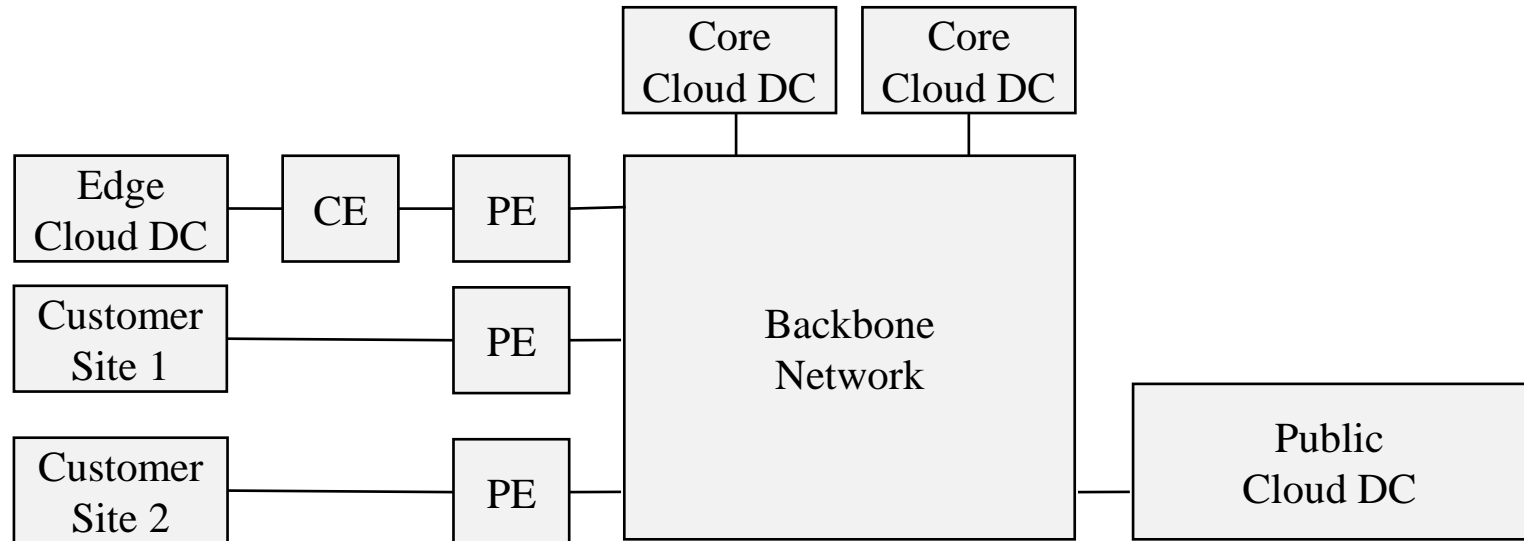
NeoTec@IETF 121, November 2024

Overview

- Telecom cloud refers to a cloud computing system owned and operated by Telecom Service Provider (TSP).
- Elasticity is a key characteristic of telecom cloud services, allowing for rapid deployment and flexible adjust of resources, including network resources.
- This document illustrates a typical use case of network for telecom cloud, including its architecture and major capabilities. A general model is also proposed to support further discussion.
- The purpose of this document is to show the challenges encountered and identify the works for collaboration particular within the IETF.

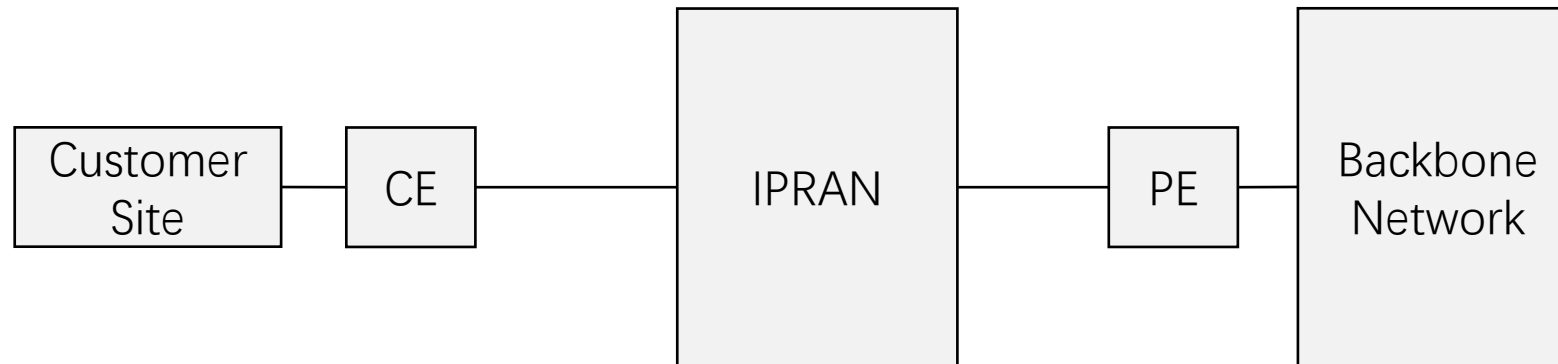
Introduction

- Telecom Cloud facilities include Core Cloud, edge cloud, telecom cloud and third-party public cloud may be interconnected through the network.
- A TSP deploys its backbone network using MPLS or SRv6+EVPN to connect cloud resource pools across a wide area, enabling customers to access both local and remote telecom clouds.
- The performance of the cloud service highly depends on the operation of the underlying network.



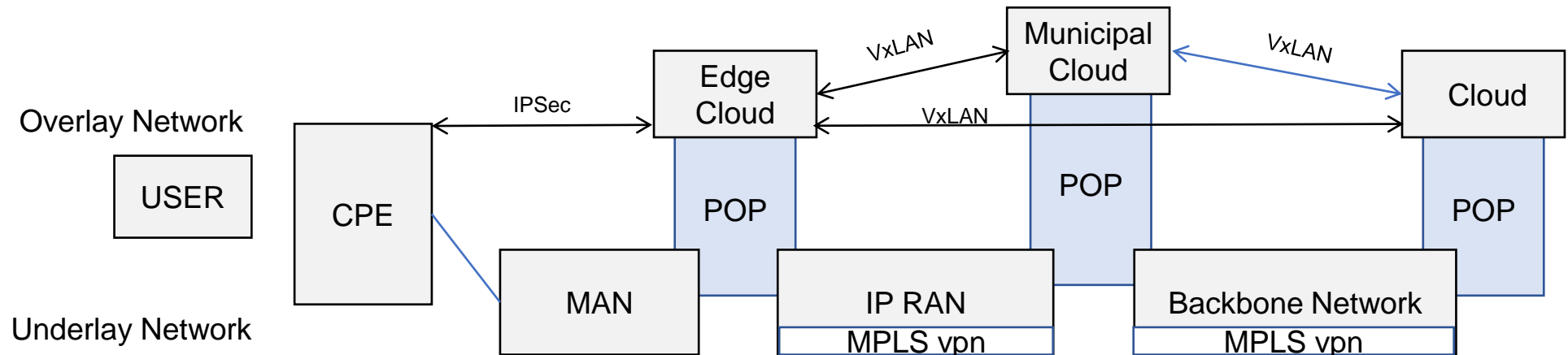
Cloud Access Service/Cloud leased Service

- Telecom cloud resource pool sites are usually deployed close to customers to reduce data transmission latency and provide a more localized service experience.
- Cloud access service enable customers to seamlessly connect their IT resources or data centers to the telecom cloud resource pools through the **Cloud Service Orchestrator**.
- Requirements to Cloud Access Service/Cloud Leased Service:
 - Agile service activation, low latency, high reliability of data transmission, etc.



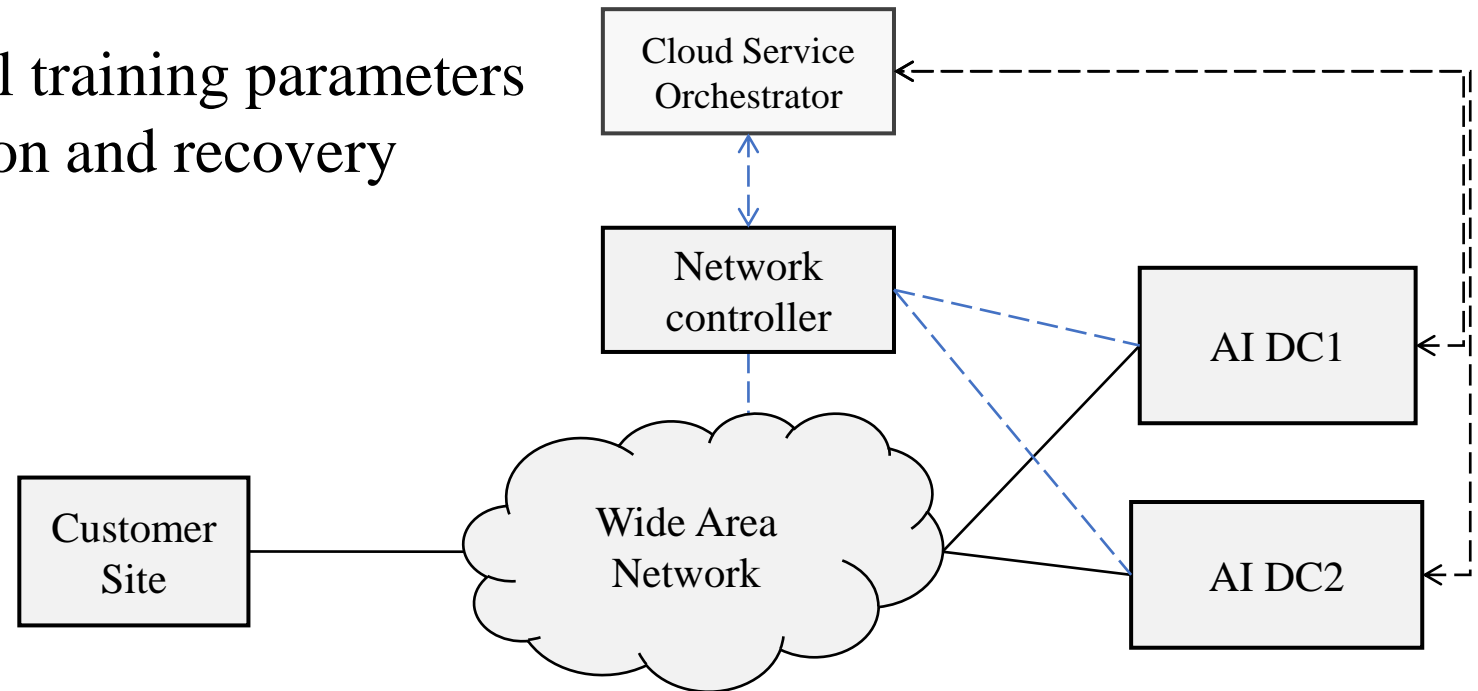
Inter-DC/Cloud Connection Service

- Inter-DC/Cloud Connection is an overlay network which needs configurations of tunnel endpoints, encapsulation protocols, load balance, ACL, etc.
- For large scale networks, service creation and delivery should be automatic,
 - Agile service activation
 - Exposure of network status to cloud, so workload may be deployed in an informed fashion
 - Dynamic adjustment of network resources to adapt to cloud scaling
 - High stability, reliability and secure



Network for AI Cloud

- Large Language Model(LLM) training puts more stringent requirements on the network, the network needs to support the massive scale (Tbits , 10Tbits) of data transmission and lossless connection between servers in cloud data centers.
 - Vacant bandwidth aggregation
 - Elephant flow path selection
 - Load balance based on model training parameters
 - Instant fault detection, location and recovery



Interface of Network Operation for Telecom Cloud

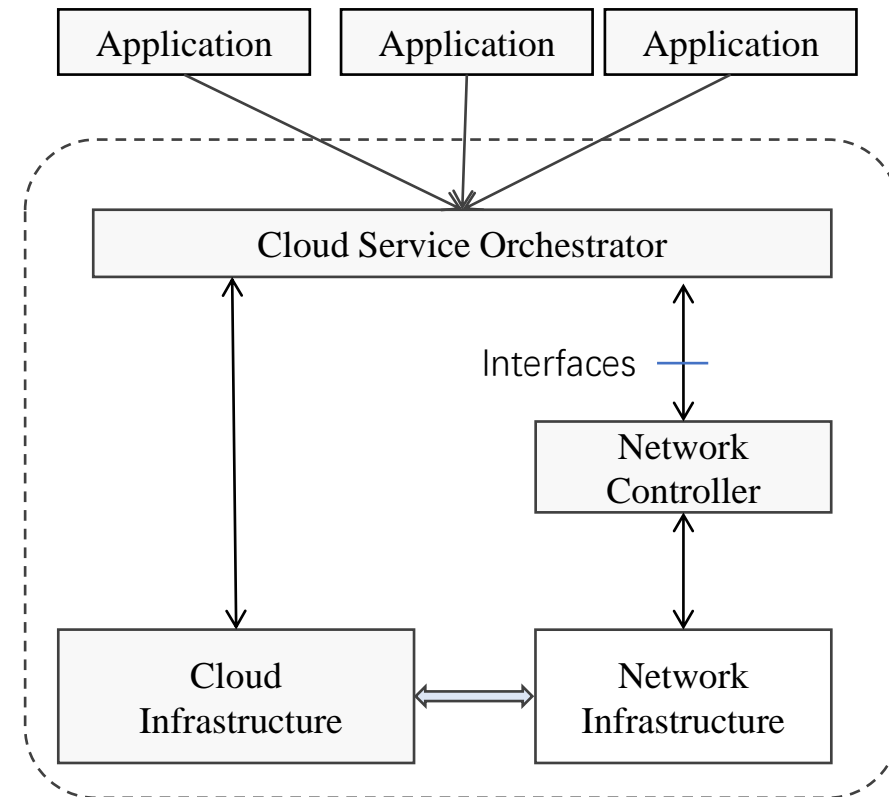
Interfaces need to be defined between the Cloud Service Orchestrator and the Network Controller, so as to achieve quick response to the requirement of cloud services.

(1) Interface from Cloud to Network, enforce connectivity service to provide guarantees

- SLO enforcement (latency, jitter, etc.)
- Network resource adjustment for cloud scaling
- Network traffic scheduling
- Security isolation requirement
- Load Balancer configuration
- AI-specific service information

(2) Interface from Network to Cloud, expose the network information to cloud

- Network performance data
- Data paths of a specific overlay network
- Usage of a specific node or link
- Network fault notification



Comments received and revisions

- Luis Miguel:
 - There is a need to enhance the capability of exposing network or cloud resources, such as network topology and cloud computing capabilities, and to provide a comprehensive and complete view.
 - It is necessary to establish an interface between the cloud management system(Orchestrator) and the network management system(Controller).
- This will be added in the next version.

Next Steps

- The interface requirement will be further refined based on the feedback from other WGs.
- Comments and suggestions are welcome.

Thank you!

Problem statement and gap analysis

Frankly, current TSP network can't yet perfectly meet the needs of the cloud, due to both managerial and technical reasons.

- Lack of definition for Key Interfaces, in particular, interface between controllers of network and cloud
- Lack of key data models, in particular, cloud resource model
- Lack of Unified Operation and Maintenance Management Monitoring (IOAM)
- Lack of Unified Cloud-Network Orchestration Mechanism