# EdgeBoost: Data-Driven Traffic Engineering and Configuration Verification for Cloud Service Assurance

PI: Prof. Mahesh K. Marina

## ABSTRACT

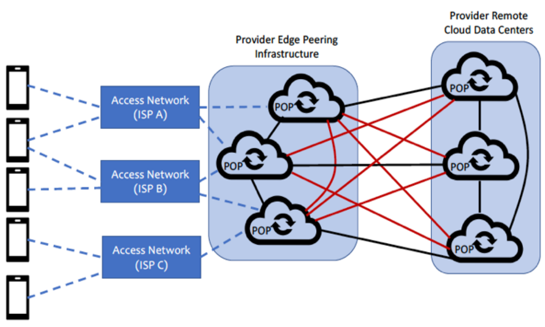
Most networked applications running on end devices (e.g., smartphones, IoT devices) are partially or wholly served by the remote processing and storage in the cloud. To better support such cloud dependent applications, service and content providers have been moving towards dedicated and custom network infrastructure that is not just limited to their remote geographically distributed data centers but extends its footprint deep into the edge. In recent times, this provider network infrastructure is being augmented with cloud compute and storage infrastructure at the edge nearer to devices so as to improve latency and user experience as well as provide bandwidth savings to services (e.g., video-oriented services, applications requiring real-time analytics capabilities).

The broad aim of this project is on the end-to-end service assurance in the setting outlined above. Here the term "service" can refer to a set/class/group of application flows that run on end-devices with similar requirements (e.g., maximum delay, minimum data rate, tolerable packet loss) as opposed to a particular application with particular requirements. Achieving service assurance in this setting is a significant challenge, principally because of the sharing of underlying network infrastructure among different services; public Internet making up part of the network infrastructure; and lack of visibility into certain parts of the infrastructure, especially between the device and edge. This project aims to address this challenge through machine learning driven techniques for predicting service demands and link qualities, intelligent and adaptive traffic engineering as well as for efficient troubleshooting. This project will also address the associated practical concerns of enforcing and verifying the control plane routing strategy decisions over the data plane.

## BACKGROUND

- Most networked applications on end devices partially or wholly served by remote processing and storage in the cloud
- To better support such cloud-dependent applications, service/content providers extending their infrastructure towards the edge
- Providers extend their footprint to the edge through their own private wide area network (WAN) while also peering with public Internet at the edge points of presence (POPs)
- Edge infrastructure not just for networking but also augmented with compute and storage, making it a cloud closer to end devices

## TARGET END-TO-END SCENARIO




- Provider's remote and edge POPs form a fully connected "overlay" mesh network
- Each overlay link can span multiple physical network links
- Black colored overlay links in the figure are solely through provider's private WAN
- Red colored ones include links from public Internet

## PROJECT OBJECTIVES

- Broad aim of this project is on *end-to-end service assurance* in the target scenario outlined above
- "Service" 🡐 a set/class/group of application flows that run on end-devices with similar requirements (e.g., maximum delay, minimum data rate, tolerable packet loss)
- Achieving service assurance implicitly requires performance isolation between different services sharing the underlying infrastructure and adaptation to dynamics over public Internet links
- Also need to track faults and performance bottlenecks on the path segment between edge PoPs and end devices

## TECHNICAL CHALLENGES

1. Accurate prediction of service demands and link performance characteristics over time
2. Adaptive traffic engineering at service level: path selection and bandwidth allocation for each service's traffic to meet its SLA
3. Enforcing and verifying control plane routing decisions over the data plane
4. Identifying faulty and anomalous path segments

## METHODOLOGY & WORK PLAN

1. Link performance and service demand matrix prediction
2. Intelligent and adaptive data-driven traffic engineering
3. Data plane centric policy enforcement, verification and feedback
4. Efficient, end-to-end network troubleshooting

*We will investigate machine learning driven techniques for (1), (2) and (4)*

## REFERENCES

1. K. Yap et al., "Taking the Edge off with Espresso: Scale, Reliability and Programmability for Global Internet Peering," in *ACM SIGCOMM*, 2017.
2. J. Jiang, V. Sekar, I. Stoica and H. Zhang, "Unleashing the Potential of Data-Driven Networking," in *COMSNETS*, 2017.
3. H. H. Liu et al., "Automatic Life Cycle Management of Network Configurations," in *ACM SIGCOMM Workshop on Self-Driving Networks (SelfDN)*, Aug 2018.
4. L. Ma, T. He, A. Swami, D. Towsley and K. Leung, "Network Capability in Localizing Node Failures via End-to-End Path Measurements," in *IEEE/ACM Transactions on Networking*, 25(1), Feb 2017.