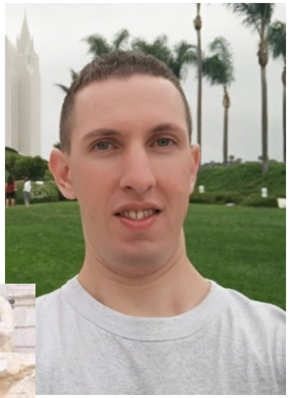
# ViSNext'21

First ACM CoNEXT Workshop on Design, Deployment, and Evaluation of Network-assisted Video Streaming

# Panel Session

# Panel Session

# Network-assisted Video Streaming and QoE Improvement: Challenges and Opportunities



## Panelists:

- Abdelhak Bentaleb – National University of Singapore, Singapore
- Lucia D'Acunto – TNO, The Netherlands
- Cornelius Hellge – Fraunhofer Heinrich Hertz Institute, Germany
- Lea Skorin-Kapov – University of Zagreb, Croatia

## Moderator:

- Hermann Hellwagner – University of Klagenfurt, Austria

# Initial Guiding Questions to Panelists

- **“Network assistance” required** to support video streaming **at all?**
  What type of network assistance, where, based on which information?
  Role of new networking paradigms (e.g., SDN, NFV)? Security/privacy concerns?
- Specific role, importance of, and (aggregation) level of **edge computing (MEC)?**
- If “network assistance” becomes important, how would **upcoming techniques (e.g., ML) and new protocols (e.g., QUIC)** be **adopted** in the networking industry?
- Would **significant new “entities”** appear, **similar to CDN providers** in today’s Internet for large-scale content distribution? If so, where?
- **Central challenge and central opportunity** of leveraging “network assistance” for video streaming and QoE improvement?

# ViSNext'21

First ACM CoNEXT Workshop on Design, Deployment, and Evaluation of Network-assisted Video Streaming

*Thank you, stay healthy, and farewell !!!*

# Panel
# Network-assisted video streaming and QoE improvement: challenges and opportunities

ACM CoNEXT-- ViSNext 2021

**Abdelhak Bentaleb, National University of Singapore**

# Why Network-assisted Video Streaming is Important ?

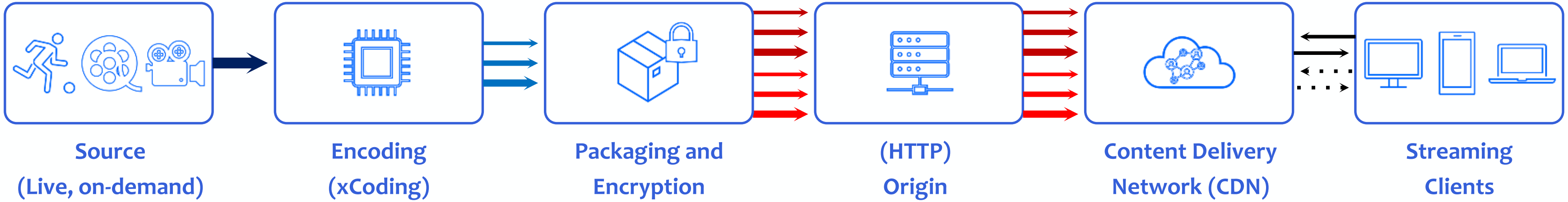
- State-of-the-art Approaches
  - In-network (e.g. network devices)
  - End-point (e.g client or server)

- Limitations
  - In-network:
    - limited by in-network devices
    - No  visibility to viewer QoE (QoS metrics only)
    - Hard to make substantial changes to the network core
  - End-point:
    - limited by individual endpoints' local visibility to network conditions
    - Single flow: Selfish algorithm that seek to optimize QoE without considering other flows in the network

# Why Network-assisted Video Streaming is Important ?

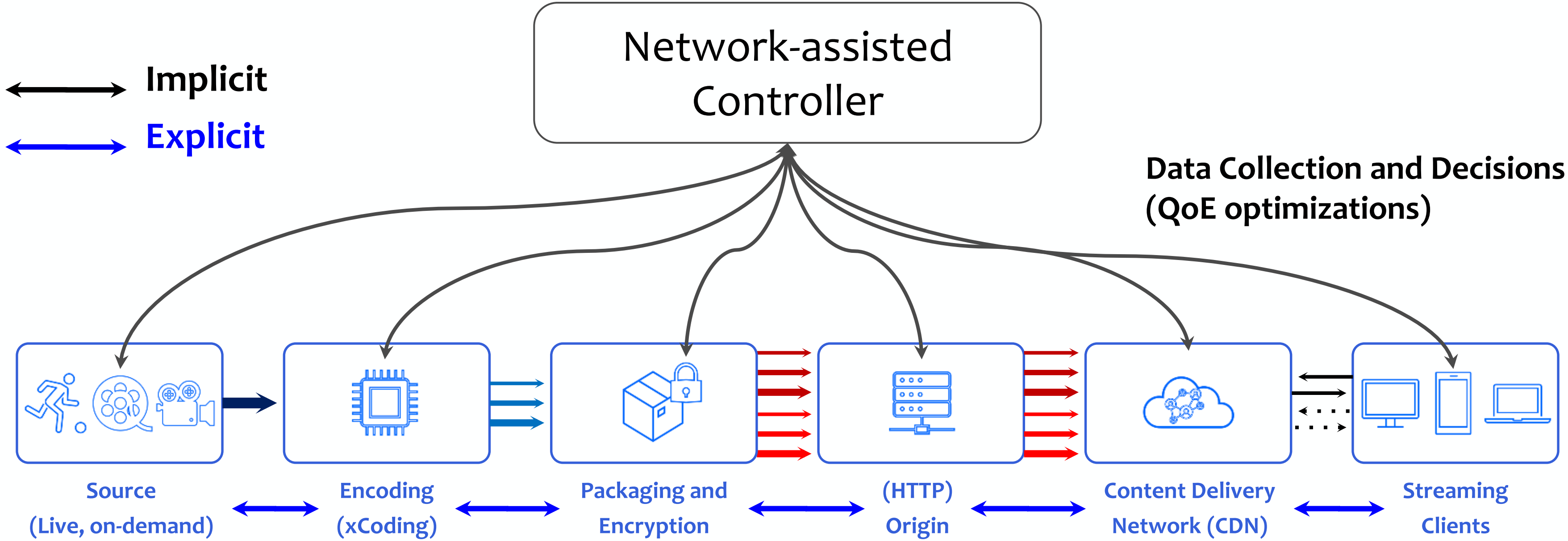
Network assisted approach can be **located** at **many entities** within delivery workflow, and has **full visibility to QoE impairment & network conditions.** The exchanged information can be collected in a centralized controller for making decisions that improve viewer QoE.

# End-to-End Workflow for HTTP Adaptive Streaming




- Why HTTP
  - Features well-understood naming/addressing and authentication/authorization infrastructure
  - Provides easy traversal for all kinds of middleboxes (e.g., NATs, firewalls)
  - Enables cloud access, leverages the existing (cheap) HTTP caching infrastructure
- Improved viewer experience
  - Reduces startup delay (upon zapping or seeking), frame skips and stalls
  - Provides adaptation capability based on network conditions and client status

# Network-assisted Video Streaming Architecture: General View




- Two types of network-assisted video streaming cooperation (depending on use case , network architecture support, Cost of deployment, and Scalability)
  - **Implicit Cooperation**: between network- assisted controller and different delivery entities
  - **Explicit Cooperation**: between entities of the delivery workflow. Define five modes:

    i) client <-> server (CDN, edge, origin), ii) client <-> client (P2P), iii) server <-> server, iv) client <-> network devices (gateway or router), and/or v) hybrid (i-iv)

# Network-assisted Video Streaming: Practical Use Cases

- Use Case 1:
  - Multiple clients computing for the available bandwidth at shared network environment (e.g. crowded mall)
  - Server, edge or gateway -based solution will solve this issue through bitrate guidance and bandwidth estimation/allocation
- Use Case 2
  - Learning based client-driven ABR
  - Server or edge -based (with more resource) solution to deploy and run the learning model for bitrate adaptation
- Use Case 3:
  - Flash crowd phenomena in live events (significant increase in number of clients in small time)
  - Virtual Reverse Proxy (VRP) solution deployed as VNF and located either at the edge or near to the origin server for request aggregation
- Use Case 4:
  - Rendering AR/VR content at the client side
  - Edge-based solution for VR content rendering

# Challenges and Opportunities of Network-assisted Video Streaming

- Multiple network-assisted solutions:
  - Based on standard (SAND: server and network assisted)
  - Commercial analytic solutions: Convivia, datazoom, etc

- Challenges:
  - Are these solutions scalable (support millions of concurrent requests)?
  - What is the cost of deployment? and What is about security/privacy concerns?
  - What information should be collected from each entity of the delivery workflow?
  - How this information should be collected and how to unify the manner of collection and metrics to be collected?

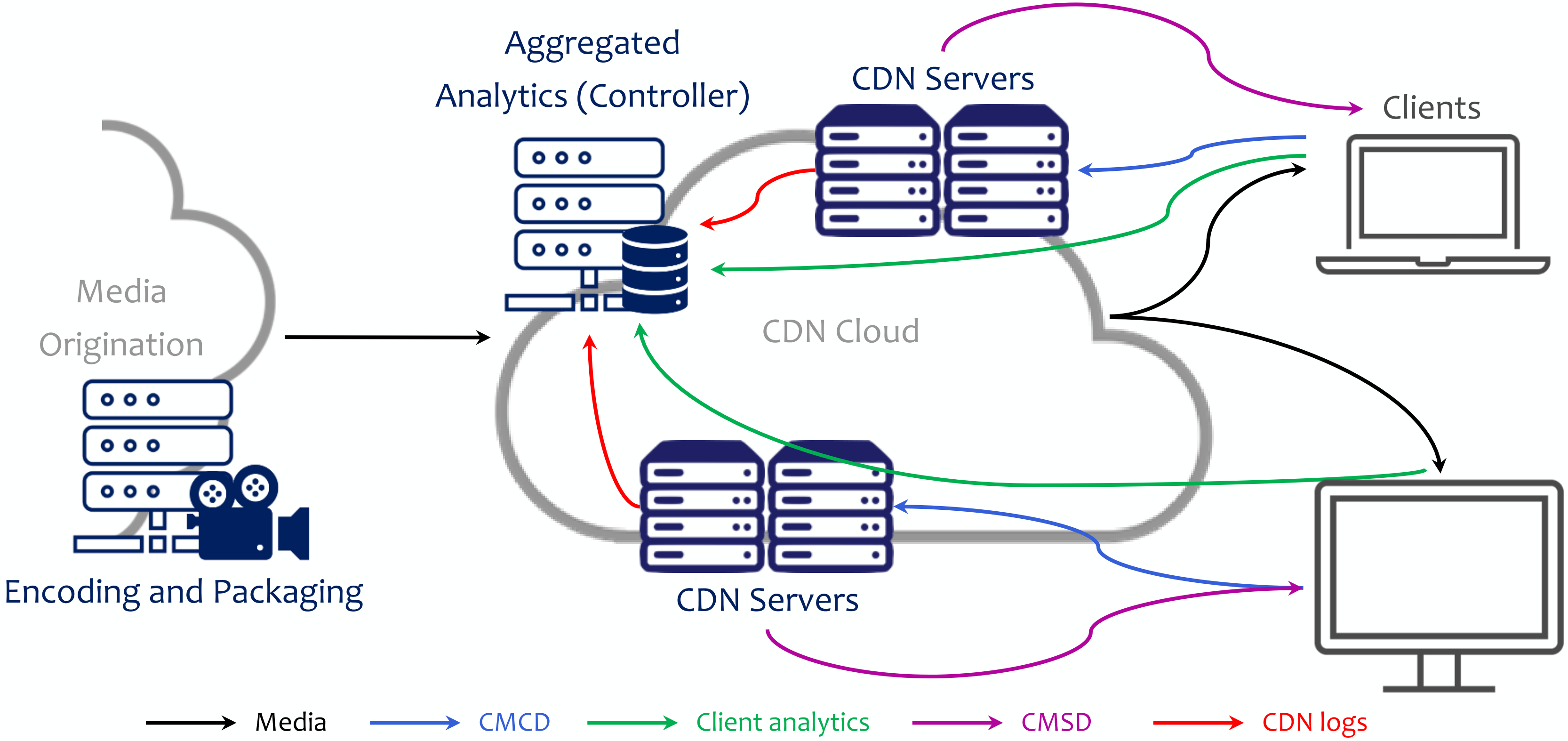
- Opportunities:
  - Emergence of new network paradigms (e.g., 5G and beyond) with their key enables (e.g., SDN, VNF, edge, MEC, etc).
  - New data unification/collection specification (CMCD: common-media-client-data and CMSD: common-media-server-data)
  - New protocols (e.g., QUIC/H3 and SRT).
  - AI-based techniques

# Security & Privacy of Network-assisted Video Streaming

- Possible Security and Privacy:
  - The malicious client reports wrong values to mislead the server or network entity in order to get more resource allocation for example.
  - A  pirate gets the exchanged information between different entities of the delivery workflow, he can infer useful data through fingerprinting, for example the title of the watched video which is one of the Infringement of privacy.
  - A malicious server or client may inject false data. This tactic may be part of replay, message insertion, or modification attacks. If the server-client communication is delivered over HTTP, then man-in-the-middle attacks are feasible.

- Recommendations:
  - The exchanged information should subject to web security model. HTTPS is recommended for transmission.
  - None of the critical information should be transmitted (like device identification, along with any persistent information across sessions).
  - All the exchanged information should be are optionally executed by the delivery entity, meaning that each entity can ignore them for security concerns.

# CMCD/CMSD: The Client-Server Cooperation



# CMCD/CMSD Parameter Values

## CMCD

| Parameter | Key | Type | Unit |
|---|---|---|---|
| **Encoded bitrate** | br | integer | kbps |
| **Buffer length** | bl | integer | ms |
| **…** | | | |
| **Measured throughput** | mtp | integer | kbps |
| **…** | | | |
| **Next object request** | nor | String | - |
| **…** | | | |

## CMSD

| Parameter | Key | Type | Unit |
|---|---|---|---|
| **Estimated throughput** | etp | integer | kbps |
| **Round trip time** | rtt | integer | ms |
| **…** | | | |
| **Max suggested bitrate** | mb | integer | kbps |
| **…** | | | |
| **Streaming format** | sf | Token [d,h,s,o] | - |
| **…** | | | |

# CMCD/CMCD: Mode of transmission

- CMCD: query arg with request or JSON based
    - Example CMCD query string

      **?CMCD=bl=4500,mtp=30000**

      which is encoded as

      **?CMCD=bl%3D4500%2Cmtp%3D30000**

- CMSD: header response with response to the client
    - Example CMSD response header (CMSD-Dynamic or CMSD-Static)

      **CMSD-Dynamic: etp=115;rtt=16;mb=5000**

## CTA-5004: Common Media Client Data (Published in Sept. 2020)

How could a client relay info about

- content ID and session ID
- current segment's type/duration/format
- device type, display size
- delivery deadline
- next segment (or byte range) to be requested
- current buffer length, latency, startup delay and playback rate
- stall stats

## CTA-5006: Common Media Server Data (Work started in May 2021)

How could a server relay info about

- server-side bandwidth estimates
- hints for the startup bitrate
- min/max limits for the playback bitrate
- redirection suggestions
- caching indications
- breadcrumb data
- server/network load signals

# Key Takeaways and Conclusions

Info exchange is useful when it is relevant, actionable and up-to-date

Running code is available to for testing and trialing

dash.js fastly THEOplayer Akamai

What information is relevant and actionable is the main question

Join the effort and contribute

WAVE WEB APPLICATION VIDEO ECOSYSTEM

# Thank you

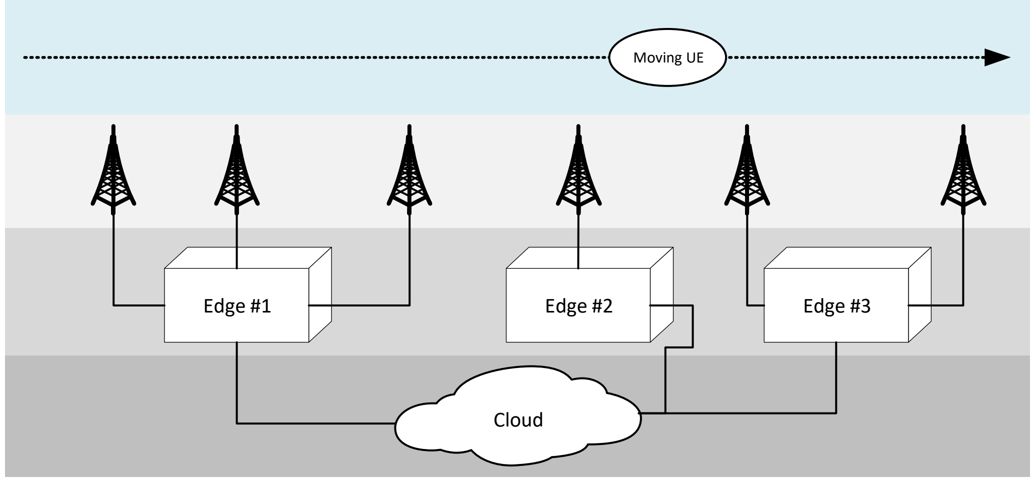
Bentaleb Abdelhak
bentaleb@comp.nus.edu.sg

TNO innovation for life

# NETWORK-ASSISTED VIDEO STREAMING AND QOE IMPROVEMENT: CHALLENGES AND OPPORTUNITIES

PANEL SESSION, VISNEXT 2021 | DR. IR. L. D'ACUNTO

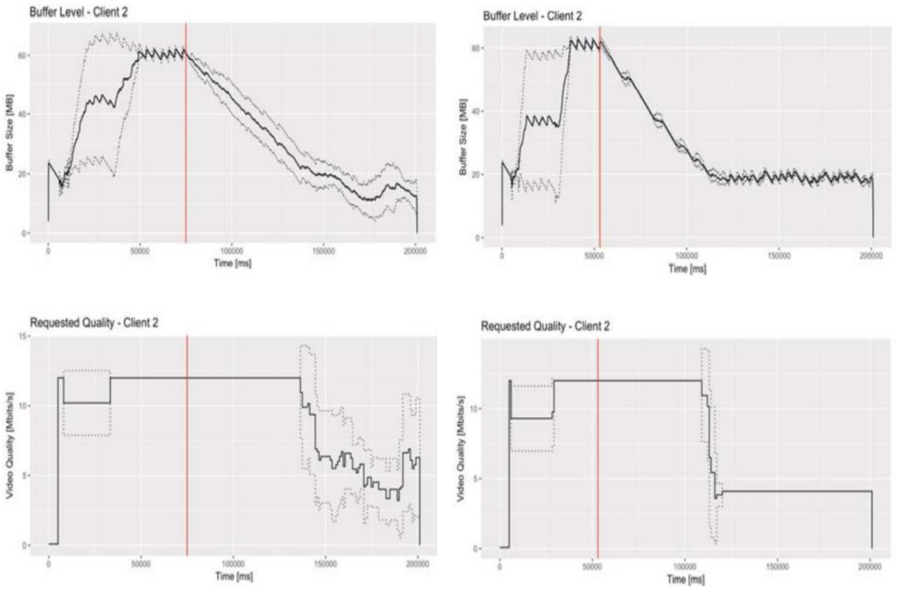
# NETWORK ASSISTANCE FOR VIDEO STREAMING GOES BEYOND CACHING AND BITRATE RECOMMENDATION




Support for handover without loss of service when using edge computing

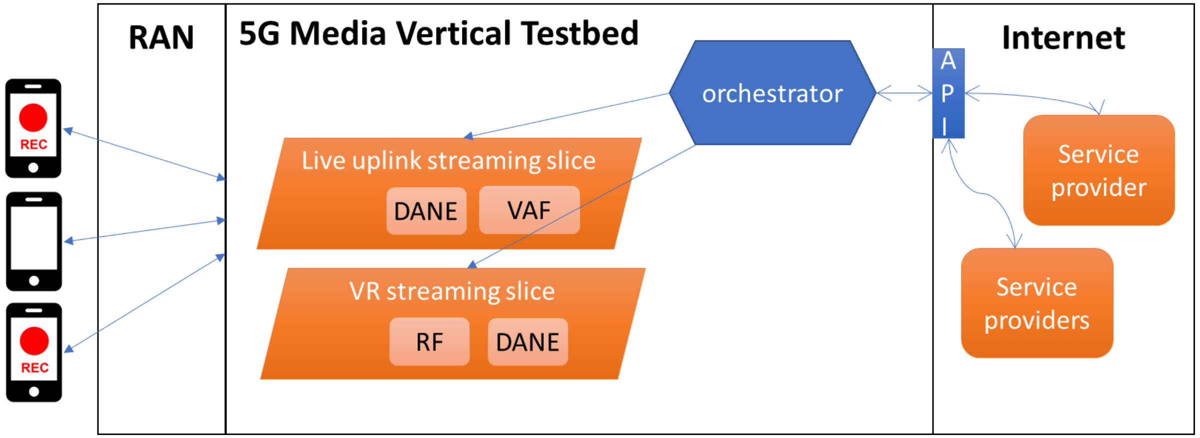
In-network processing (e.g. tiling / rendering for VR content)

# SDN/NFV AND NETWORK SLICING ARE ESSENTIAL TO THE EFFECTIVENESS OF NETWORK ASSISTANCE

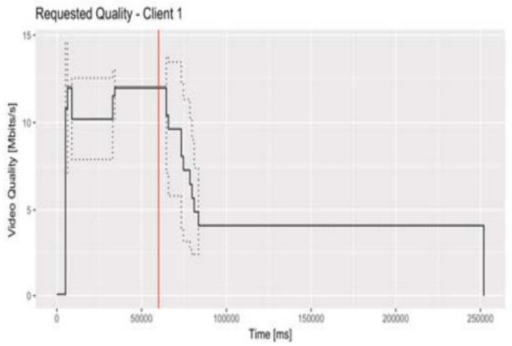


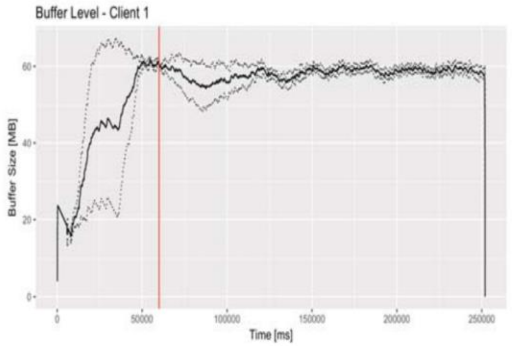

NO SDN vs SDN

Network slicing for personalisation and sandboxing



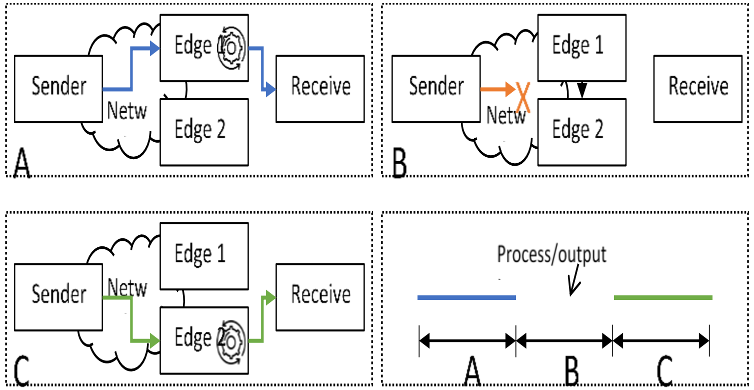
# NETWORK ASSISTANCE HAS THE POTENTIAL OF IMPROVING OPTIMIZATION OF RESOURCES AND INCREASING QOS

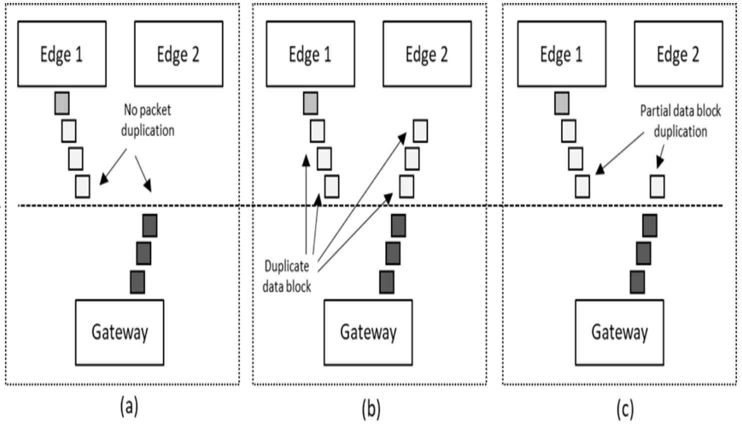



Optimize resource allocation + offer TV-quality services (no stalls)

Avoid stalling / rebuffering + optimize network and edge resources



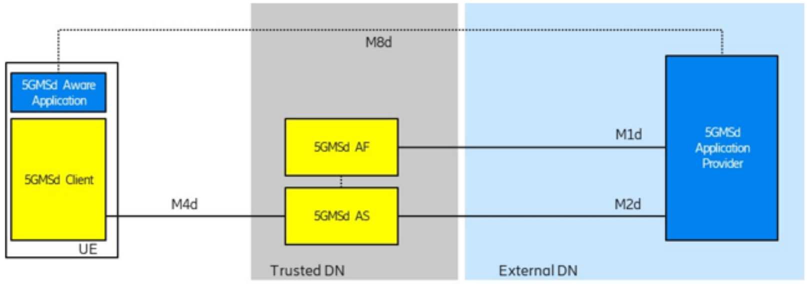



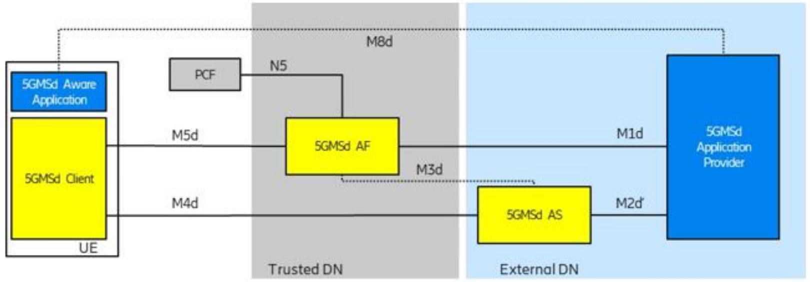
# COLLABORATION MODELS BETWEEN OPERATORS AND SERVICE PROVIDERS ARE A CHALLENGE

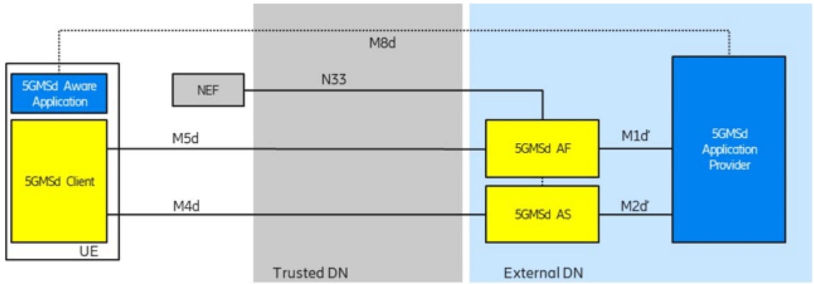
Where do the network assistance functionalities reside?

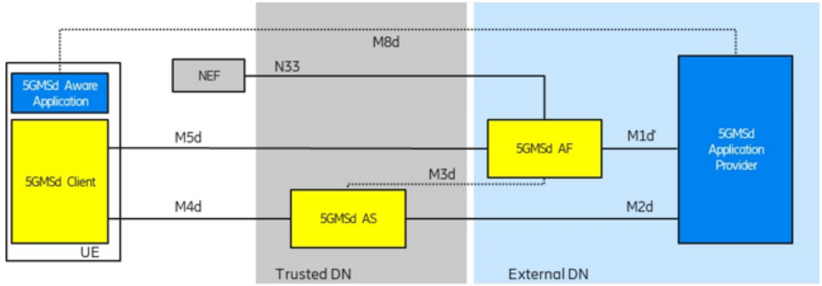
Which data (network, client, edge) do they use?

Who controls the network slice?



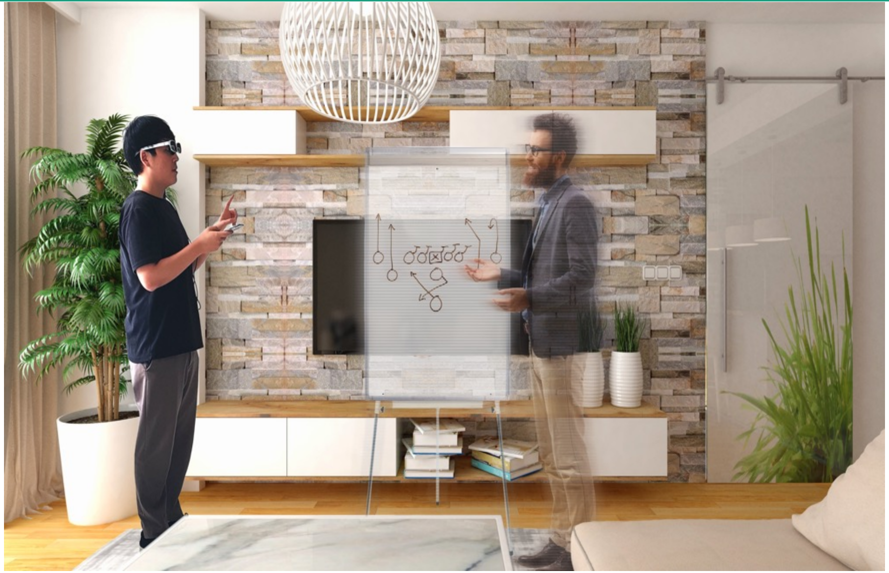
# THANK YOU FOR YOUR ATTENTION

TNO innovation for life

# VISNEXT'21

## PANEL: NETWORK-ASSISTED VIDEO STREAMING AND QOE IMPROVEMENTS: CHALLENGES AND OPPORTUNITIES

Dr. Cornelius Hellge





Fraunhofer HHI

# Some definitions

**Service/Latency Types/Protocols**

- Ultra-Low latency service (<100ms latency)
- Conversational services (<200ms latency)
- Low latency services (<=2s latency)
- Higher latency services (>2s latency)

WebRTC

HLS/DASH

**Service generation**

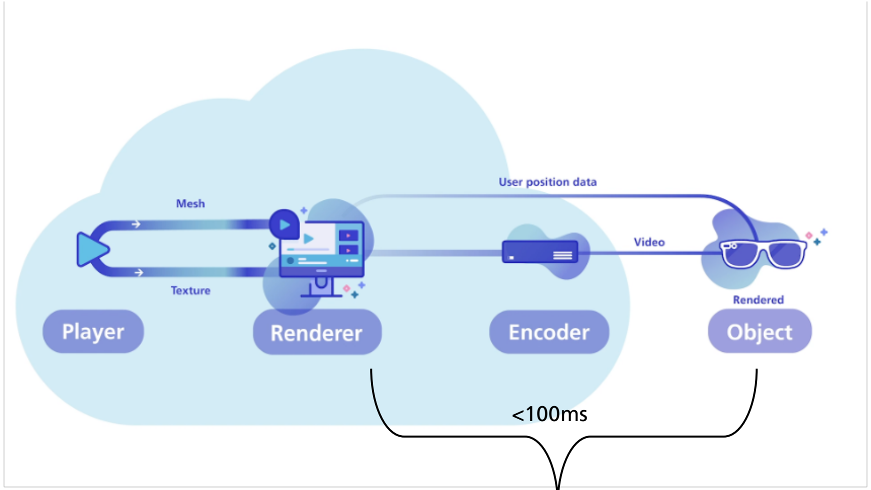
- Live (Captured)
- Live (Generated) (e.g. Cloud gaming)
- VoD

**Connection types**

- Fixed
- Wireless

# Example application (<100ms)

## MR/AR split-rendering



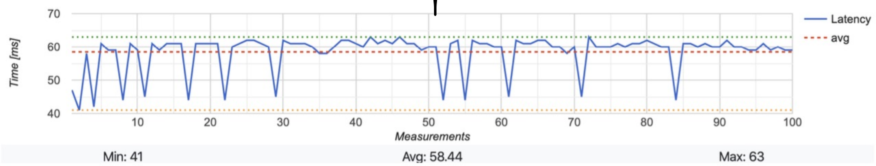
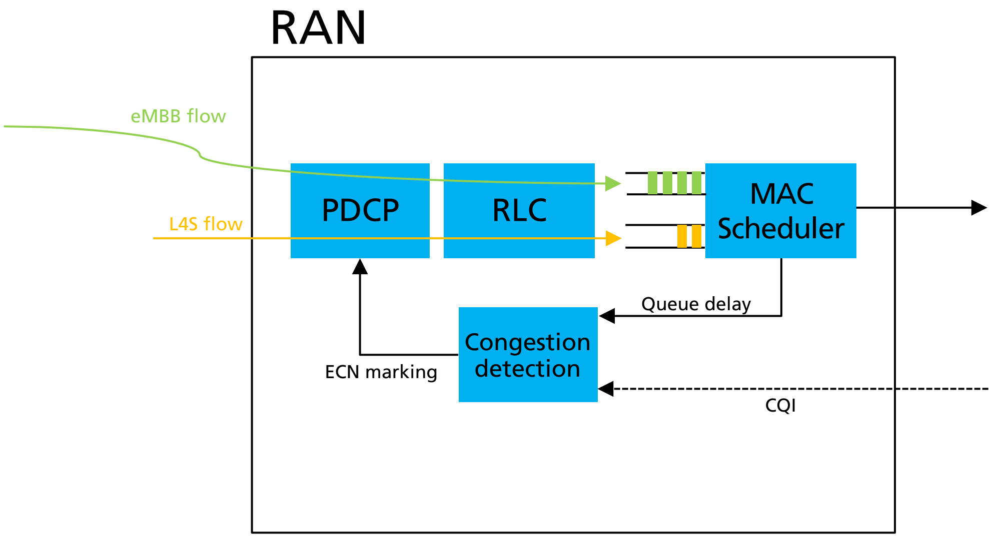




Figure 4 - Results of our motion-to-photon latency measurements.

www.hhi.fraunhofer.de/5GXR

# Network-assistance important for Ultra-low latency video services over mobile networks

- Ultra-low latency services (<100ms)
  - Cloud-gaming
  - Remote driving
  - XR/AR Split-rendering




- Direct feedback from RAN network
  - Prioritized L4S flows
  - AML with ECN based feedback
  - Immediate reaction to congestion

- **Does additional information to ECN feedback help?**
- **Benefit of ECN support for higher delay services?**
- **Does prediction of user behavior help to reduce service latency requirements?**

# Mobile Edge-Cloud (MEC)

- Mobile Edge-Cloud enables complex services on lightweight mobile devices
  - Complex rendering for VR/AR/XR services or cloud gaming
  - Real-time machine learning services (e.g. image recognition, translation)

- Position of MEC in network depends on latency requirements of service

- MEC in operator network
  - In operator network allows very low latencies (edge-cloud transparent to user)
  - Operator can enable network-support (AML)
  - Opens-up new business for operators

- MEC becomes the workhorse for mobile devices (e.g. glasses)
- Operators will become new entity for MEC provider

# General observation

- VoD and live services
  - Network support important for ultra-(very-)low latency live services
- Massive challenge for networks from new services
  - Video upload from Cars (e.g. remote driving, data collection for learning)
  - Immersive services
  - Video for machine-to-machine communication
- Potential research areas
  - Optimize network/resource planning depending on streaming bitrate characteristic (e.g. use mmWave Cells for fast upload/download)
  - Optimize all system components for low latency MEC services (encoding, ML for user/network prediction, rate-adaptation)
  - AI driven network slice optimization
- Standardization important for information exchange

# Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute, HHI

## WE PUT SCIENCE INTO ACTION.


Contact:

Dr. Cornelius Hellge
cornelius.hellge@hhi.fraunhofer.de


More information:

www.hhi.fraunhofer.de/5GXR

UNIVERSITY OF ZAGREB
Faculty of Electrical Engineering and Computing

MUEX lab
Multimedia Quality of Experience Research Laboratory

# Panel
# Network-assisted video streaming and QoE improvement: challenges and opportunities

Lea Skorin-Kapov
Univ. of Zagreb, Croatia

ViSNext’21: 1st ACM CoNEXT Workshop on Design, Deployment, and Evaluation of Network-assisted Video Streaming
Dec. 7, 2021

# Is “network assistance” required to support video streaming?

- issues with pure **client-based adaptation decisions** (poor bandwidth utilization, unstable quality, unfair resource allocation among users)
- **network assistance**: can be implemented at various stages of content delivery and at various point along the delivery path
  - players make improved adaptation decisions based on feedback from the network
  - the network performs quality adaptation based on global view (e.g., centralized solutions based on SDN, NFV)
  - exploitation of edge caching (reduced latency and network load)
- placement of network assistance entities: at the edge vs deeper in core network

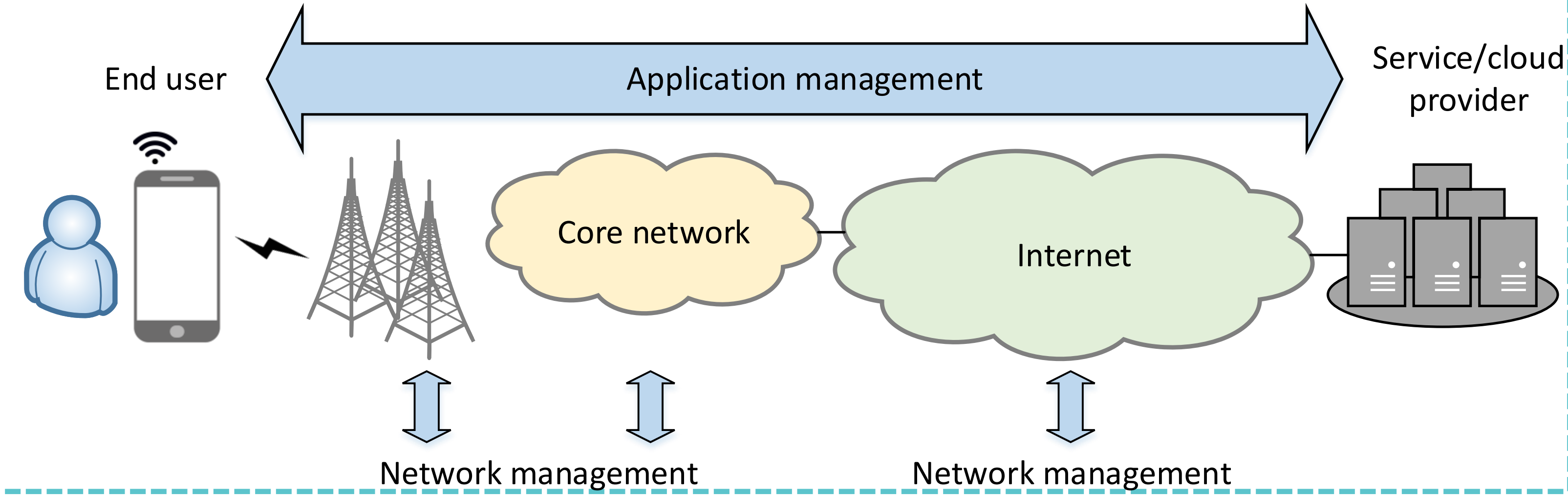
# Achieving QoE fairness

- QoS fairness ≠ QoE fairness
- Challenges: formulation of the QoE and fairness optimization problem
  - Which QoE metric to use?
  - How to address client heterogeneity?
  - Across what time frame (e.g., in fixed time slots, on the fly, etc)?
- To quantify and compare different approaches in terms of QoE fairness: **QoE Fairness Index** ($F = 1 - \sigma / \sigma_{max}$)

T. Hoßfeld, L. Skorin-Kapov , P. Heegaard, M. Varela, **Definition of QoE Fairness in Shared Systems**, *IEEE Comm. Letters*, 2017,
T. Hoßfeld, L. Skorin-Kapov , P. Heegaard, M. Varela, **A New QoE Fairness Index for QoE Management**, *Quality and User Experience*, 2018

# Challenges

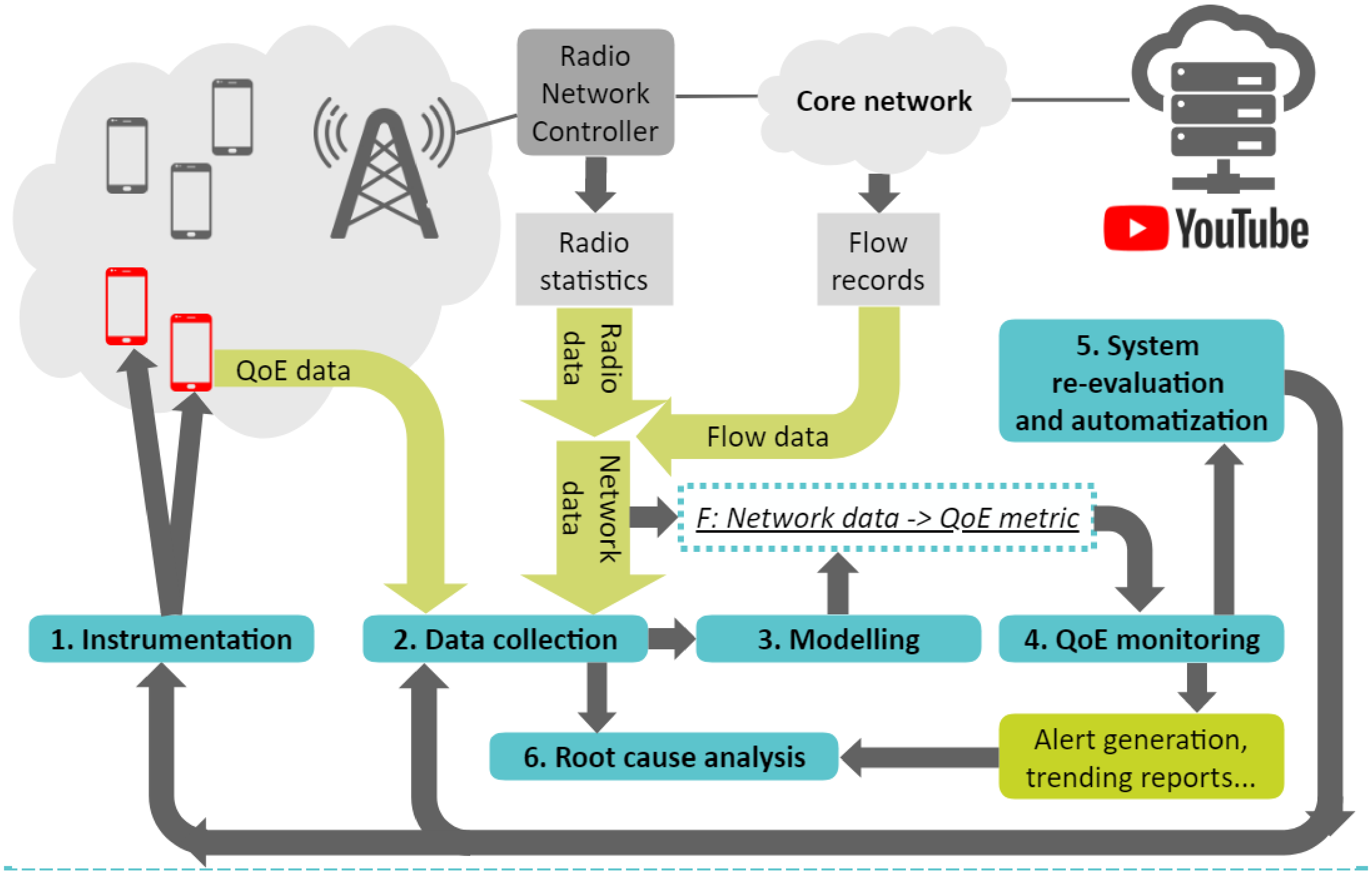
- Many (potentially new) actors involved in E2E service delivery
- Effectiveness of current QoE management mechanisms is potentially limited due to lack of information exchange and cooperation among actors
- Business relationships enabling cooperative (cross-layer) QoE management
  - incentives for information exchange
  - technical realization (e.g. MPEG SAND – real world implementation is still an issue)
  - ensuring end user privacy



# Issue: E2E traffic encryption

- E2E traffic encryption: operator network assistance and caching solutions that require access to content info may not be applicable
- Problem:
  - ISPs lack insight into OTT streaming performance and relevant app-level KPIs
  - How to perform in-network QoE monitoring, root cause analysis, dynamic resource allocation to help mitigate QoE impairments?

# In-network QoE monitoring leveraging ML





Orsolic, L. Skorin-Kapov, **A Framework for In-Network QoE Monitoring of Encrypted Video Streaming**, IEEE Access

I. Orsolic, D. Pevec, M. Sužnjević, L. Skorin-Kapov, **Machine Learning Approach to Classifying YouTube QoE Based on Encrypted Network Traffic**, MTAP, 2017

# Opportunities for leveraging network assistance

- Establishing **cross-layer cooperative QoE management schemes** to achieve both better end user QoE and more efficient resource utilization
- May become particularly relevant in the context of supporting emerging **immersive service scenarios**, which will impose BW requirements well beyond today’s VoD and live streaming services