Cognitive-driven Orchestration and Optimization of Cloud to Edge vCDN

Francesca Moscatelli, Nextworks S.r.l.
Challenges and Vision

• **Motivation:** Telcos, manufacturers and media content providers call for solutions to design and deploy custom functions for replication, distribution and adaptation of media contents
  - Fast-growing Media & Entertainment vertical industry
  - 5G as the enabler for high performance network services, high volumes, Any Device, Anytime, Anywhere, QoS

• **Goal:** Consolidate/build an Orchestration and DevOps platform for network media services and applications running on 5G networks

  - Hide the complexity of service development and deployment on the underlying 5G network and distributed cloud infrastructure
  - Orchestrate the deployment and scaling of media applications, with dynamic control of resource and ML-based cognitive optimization
5G-MEDIA High Level Architecture
Service Virtualization Platform & DevOps Tools
5G-MEDIA High Level Architecture

- **Service Virtualization Platform**
  - MANO (Service + Resource Orchestrator) based on ETSI OSM
  - Media Service MAPE
    - QoS/QoE monitoring used by Service/NFV orchestrator and VNFM
    - Cognitive Network Optimizer to dynamically optimize and re-configure service chains
  - VNF/NetApp Repository & Catalogue with V[N]F to be used across many M&E and network applications

- **Application Development SDK**
  - Tools for media applications DevOps (proof, package, emulate)
  - Serverless computing to focus on functions to code/execute instead of resource lifecycle mgmt (Faas)
  - Packaging of unikernels for lightweight atomic function VNFs
5G-MEDIA High Level Architecture (contd.)

- **Network Function Virtualization Infrastructures (NFVIs)**
  - To run the virtualization and abstraction layer on resources
  - VIM/NFVIs integration (OpenStack, OpenNebula, FaaS/OpenWhisk)

- **Core Network & Cloud**
  - For the deployment of legacy components and services esp. those instantiated on physical/specialized hardware

- **Micro/edge cloud**
  - To instantiate network and media functions closer to the consumer/user
UHD Media distribution over vCDN

Goal
Deliver new capabilities to media service providers by distributing UHD content (4K and 8K) with an optimal consumption of resources

Main Expected Benefits
Better experience for end users and new market opportunities in content delivery

Scenario - «My screen follows-me»
Users moving in the 5G network experience a seamless personal media experience from fixed video/audio device (e.g. at home) to personal devices (e.g. tablets, smartphones)
Scenario – service view

• **Media functions**
  - **Origin media server** based on Plex Media Server, including UHD media library
  - **Generic Audio-Video Client**

• **Caching functions** to implement a vCDN hierarchy
  - Mid and edge vCaches based on **Apache Traffic Server**

• **Generic network functions**
  - Load balancing (**vLoadBalancer**) for smart selection of vCaches, based on **HAProxy** tool
  - Name resolution for CDN (**vDNS**) based on **bind** tool
  - Security Front/Back-End (**vFirewall**) to protect users and service providers data, based on **VyOS** routing & **fw suite**
Scenario – control view

• **5G-MEDIA SVP** components
  • OpenStack Queens as NFVI/VIM for vCDN
  • MANO Service LCM via ETSI OSM R6
    • vnfAgents for cache hierarchy configuration
  • 5G Apps & Services Catalogue on top of the NFVO
  • vCDN Light SO to monitor and trigger manual scale-out for the service
  • (MAPE) Media App monitoring on vCaches
    • Telegraf + Logstash
    • Kafka + Influxdb + Grafana
    • Plex Tautulli supervisor
  • Automated CNO-driven fine-grained media service scaling and re-configuration driven by anomaly prediction
Scenario – CNO Algorithm

- **Machine Learning for vCDN anomaly detection**

- **Problem**: Identify vCDN anomalies caused by flash crowds

- **Solution**: Supervised Deep Learning algorithms based on Neural Networks
  - Option#1: Anomaly prediction based on vCache traffic load metrics
  - Option#2: Anomaly prediction based on several measured media app and network metrics
    - **virtual infrastructure**: packet rates, cpu usage, memory usage
    - **vCache**: cache hits, # active clients, http connections, byte rates
    - **Origin Server**: # active streams, expected bandwidth

- **Output CNO actions**
  - vCache VNF scale out
  - vCache VNF scale in
  - No operation
UHD Media over vCDN

“My screen follows-me” - Pilot
Pilot – Network Service Topology
@NXW portable testbed:

• Compute resources:
  • Control Node
    • Lenovo Laptop
    • CPU: 8
    • RAM: 32 GB
  • Core Server
    • Intel NUC i7
    • CPU: 8
    • RAM: 32 GB
  • Edge Server
    • Intel NUC i7
    • CPU: 8
    • RAM: 32 GB

• Networks:
  • VL-MGMT
  • VL-ORIGIN
  • VL-CACHE
  • VL-USERS
  • OS-CTRL
Pilot - Workflow: step #1 -> no CNO in place

Step 1.1
vCDN not congested -> two users can stream without perceived video degradation
Pilot - Workflow: step #1 -> no CNO in place

Step 1.2 vCDN is congested at edge vCache
user #1 perceives video degradation
user #2 connects to the vCDN and experience degradation
Step 2.1 vCDN is congested at edge vCache
CNO predicts the vCache anomaly and automatically trigger a vCDN scale
Step 2.2
user #2 connects to the vCDN and can stream without perceived video degradation

© ETSI 2019
OSM 5G Day, Lucca (IT) - Nextworks S.r.l
Future Work

- **Consolidation of CNO** algorithms and optimization scenarios
- Integration of **vTranscoder VNF**
  - as alternative to native Plex transcoding features
- Integration of the **vLoadBalancer** into the NS topology
  - more agile per-vCDN service vCache load balancing features
- Implementation and integration of **FaaS vTranscoder**
  - for deployment and operation of multi-VIM (FaaS + non-FaaS) vCDNs
- Integration of **QoE probing** functions
- Automated configuration of the vCaches hierarchy via proxy charms
- **Keep cooperating with TATA Elxsi for a vCDN PoC with TEOSM**
Thanks!

Francesca Moscatelli – f.moscatelli@nextworks.it