FACULTY OF ENGINEERING AND ARCHITECTURE

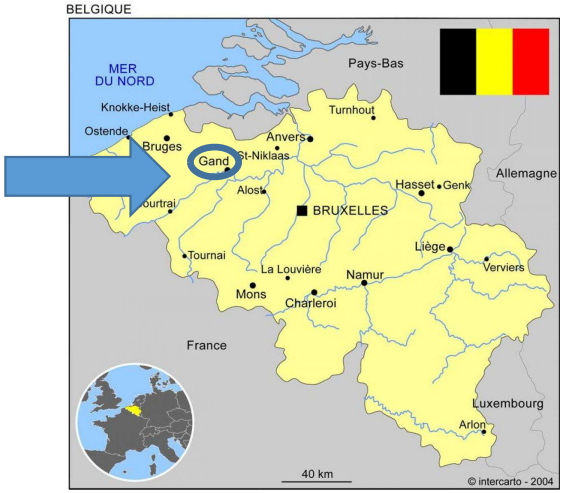
DEPARTMENT OF INFORMATION TECHNOLOGY
IDLAB-IMEC

ACHIEVING TRULY IMMERSIVE HOLOGRAPHIC-TYPE COMMUNICATION

Filip De Turck

GHENT UNIVERSITY
IDLab
INTERNET & DATA LAB
IEEE ComSoc
IEEE Communications Society
IEEE
imec

WELCOME - AFFILIATION



GHENT UNIVERSITY

Comprehensive university
Founded in 1817, ranked 74 (Shanghai ranking)
43,281 students

imec

R&D, nano electronics and digital technologies
Founded in 1984
3,500 researchers

IDLab
INTERNET & DATA LAB

Internet Technology and Data Science Lab
300 internet experts and data scientists
500+ collaborations with innovative industry partners

NOKIA Bell Labs · televic · BARCO · smappee · IBM · Adobe · orange · T · Telefónica · proximus

https://idlab.technology/

PROF. FILIP DE TURCK

- Professor at Ghent University-imec, Belgium
- Expertise in softwarized network management, network automation and service delivery,
- Many research projects in collaboration with industry
- Former Editor-in-Chief of IEEE Transactions on Network and Service Management (TNSM)
- Past Chair of IEEE Technical Committee on Network Operations and Management (CNOM)

GHENT UNIVERSITY

imec

IEEE 25 YEARS MEMBERSHIP MILESTONE

GHENT UNIVERSITY IDLab INTERNET & DATA LAB

imec


IEEE ComSoc Technical Committee (TC) on Network Operations and Management (IEEE CNOM TC)
IEEE ComSoc
IEEE Communications Society
IEEE
Advancing Technology for Humanity

ComSoc Technical Committees

Big Data

Cognitive Networks

Communication Theory

Communications & Information Security

Communications Quality and Reliability

Communications Software

Communications Switching & Routing

Communications Systems Integration & Modeling

Computer Communications

Data Storage

e-Health

Green Communications & Computing

Information Infrastructure & Networking

Internet

Internet of Things, Ad Hoc & Sensor Networks

Molecular, Biological and Multi-Scale Communications

Multimedia Communications

Network Operations & Management

Optical Networking

Power Line Communications

Radio Communications

Satellite & Space Communications

Signal Processing and Computing for Communications

Smart Grid Communications

Social Networks

Tactile Internet

Transmission, Access, & Optical Systems

Wireless Communications

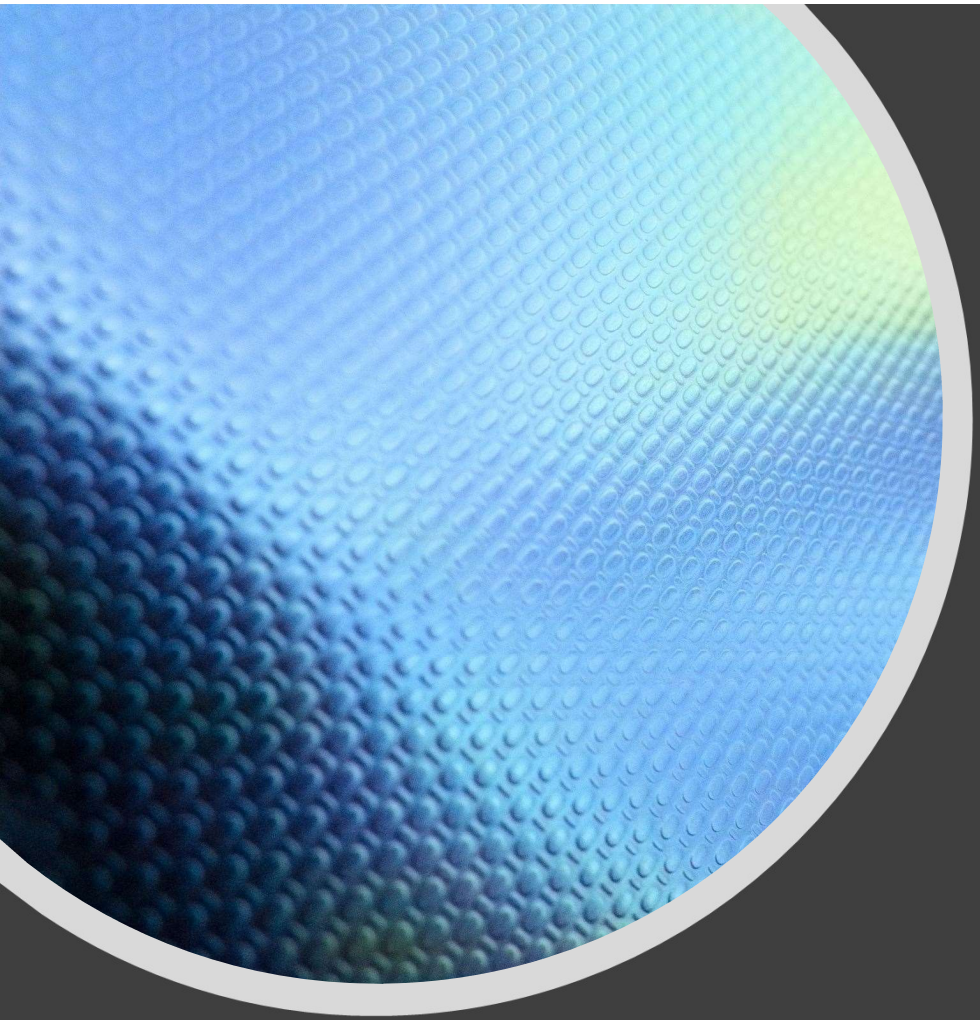
IEEE ComSoc
IEEE Communications Society

IEEE
Advancing Technology for Humanity

Network Resource Management

- Is of prime importance, as it allows to
  - make efficient use of the available resources,
  - offer service guarantees,
  - and make sure that services can be delivered with high quality of experience to end users.

- Need for efficient network resource management algorithms and methodologies.



Softwarized Networks

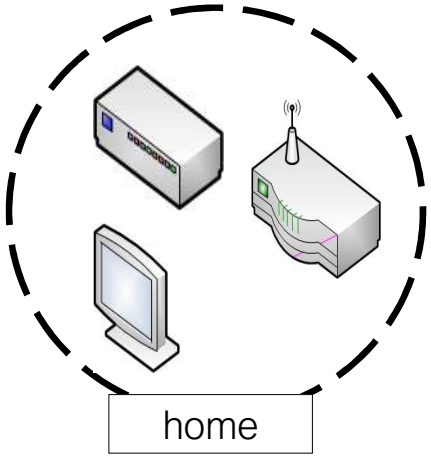
Softwarized networks bring virtualization concepts to the network

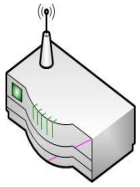
Have proven to be particularly important for the industry

Papers on efficient placement of virtualized network functions


What is network function virtualisation?

USE CASE: RESIDENTIAL EQUIPMENT



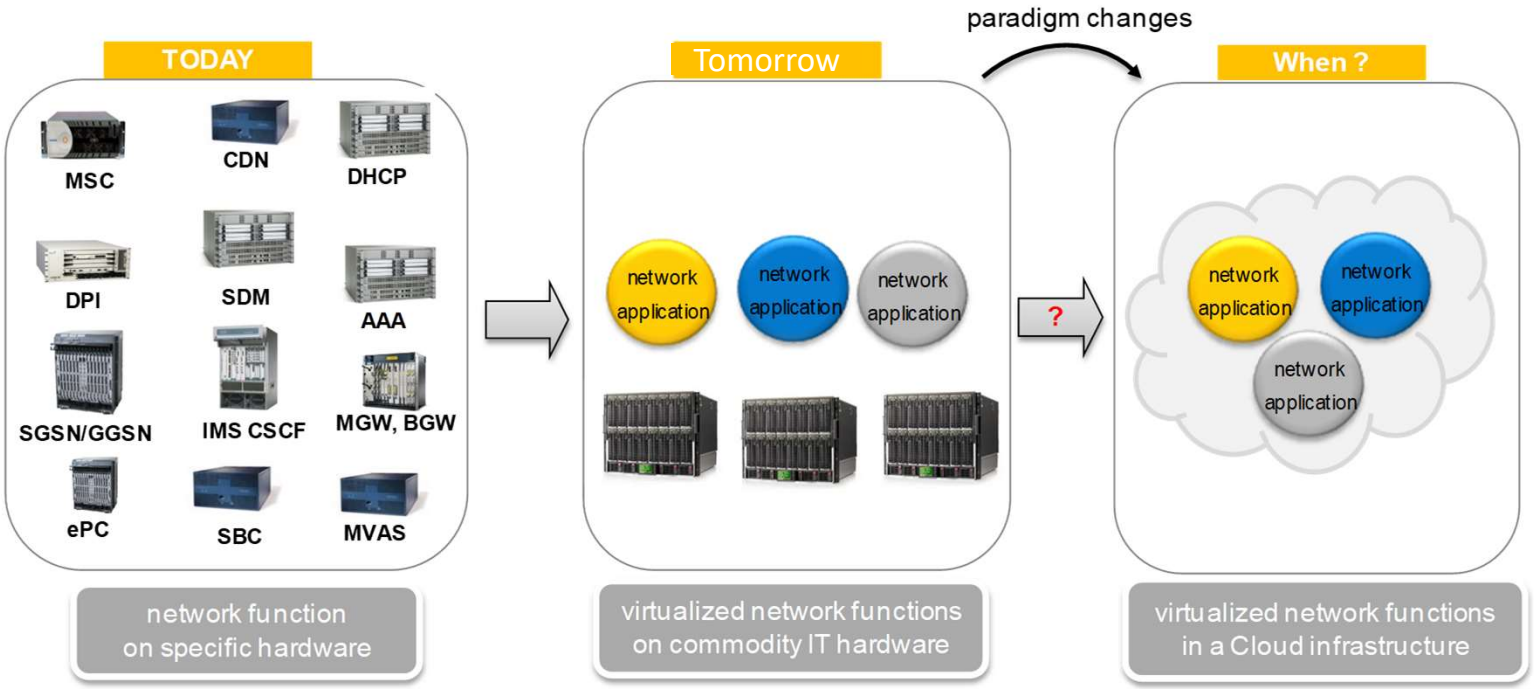


Home gateway (HGW):
- DHCP server
- NAT router
- Firewall
- QoE monitoring

Set top box (STB):
- Decoder
- Storage
- Middleware
- Streaming client

= network functions

NETWORK FUNCTION VIRTUALIZATION



GHENT UNIVERSITY IDLab INTERNET & DATA LAB imec

Advantages

for Network Operator:

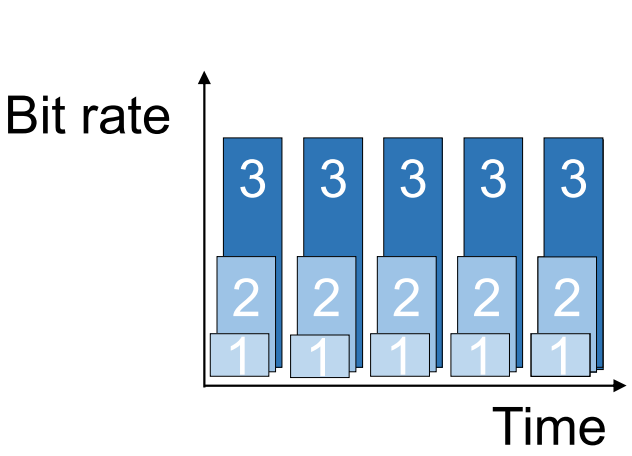
- Maximize resource utilization and optimize energy usage
- Faster and easier deployment, configuration, and updating of network functions
- Support for the Network-as-a-Service business model

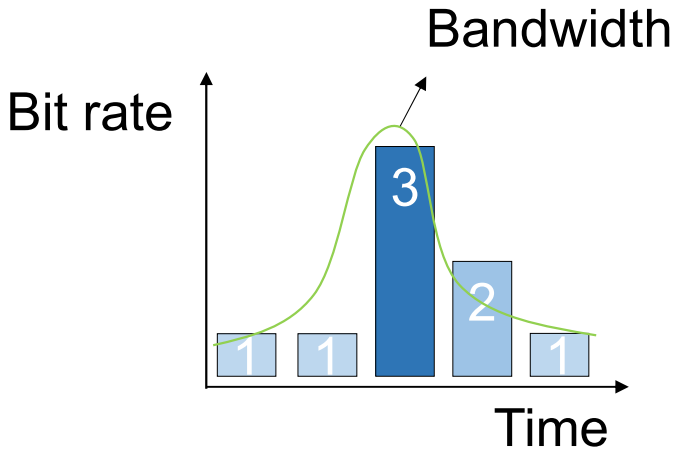
Advantages

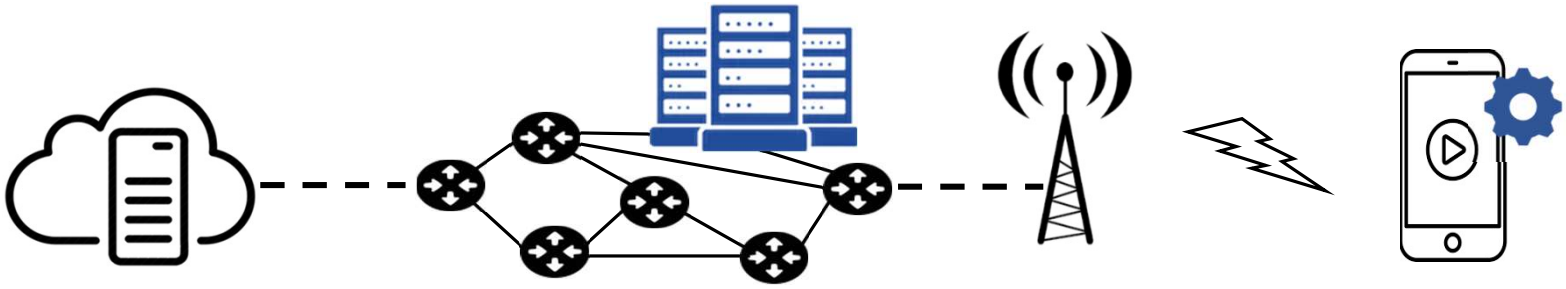
for Service Provider:

- Dynamically scale network, computing and storage resources based on service requirements
- Reduced time to market for services

ADAPTIVE SERVICE DELIVERY

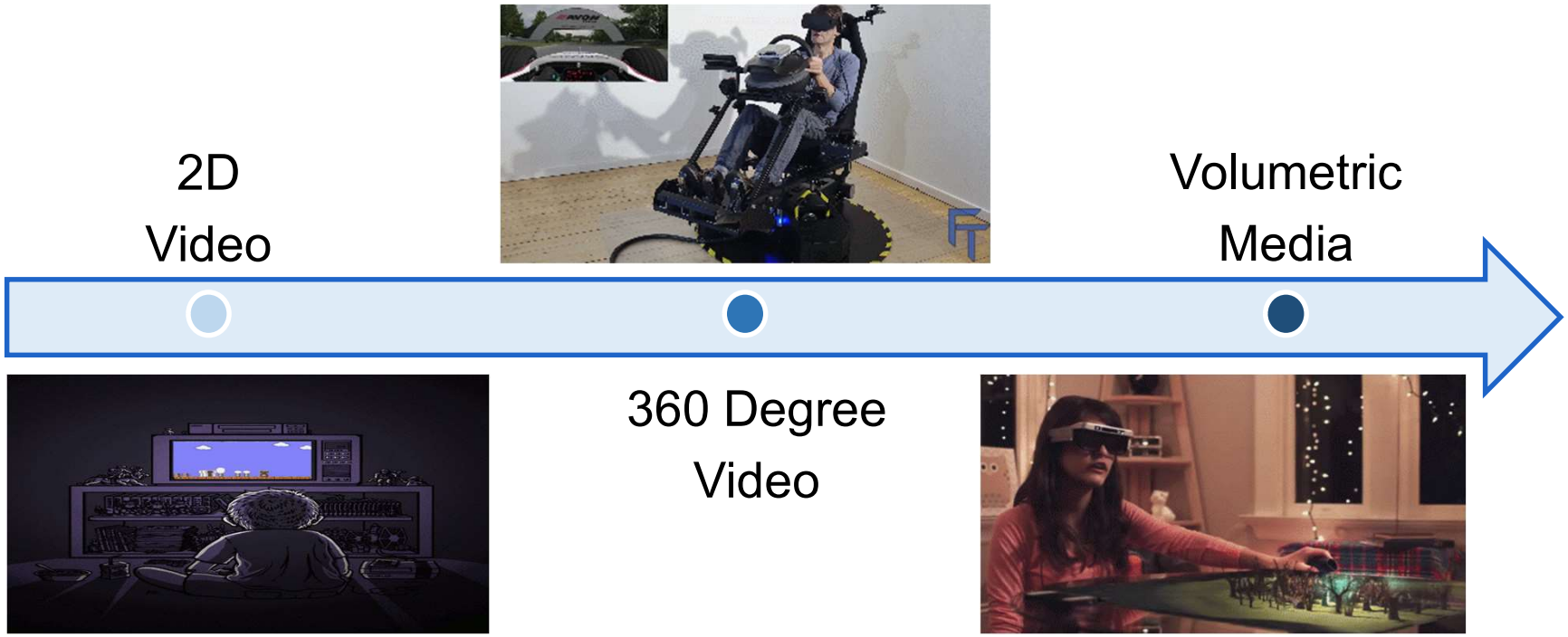




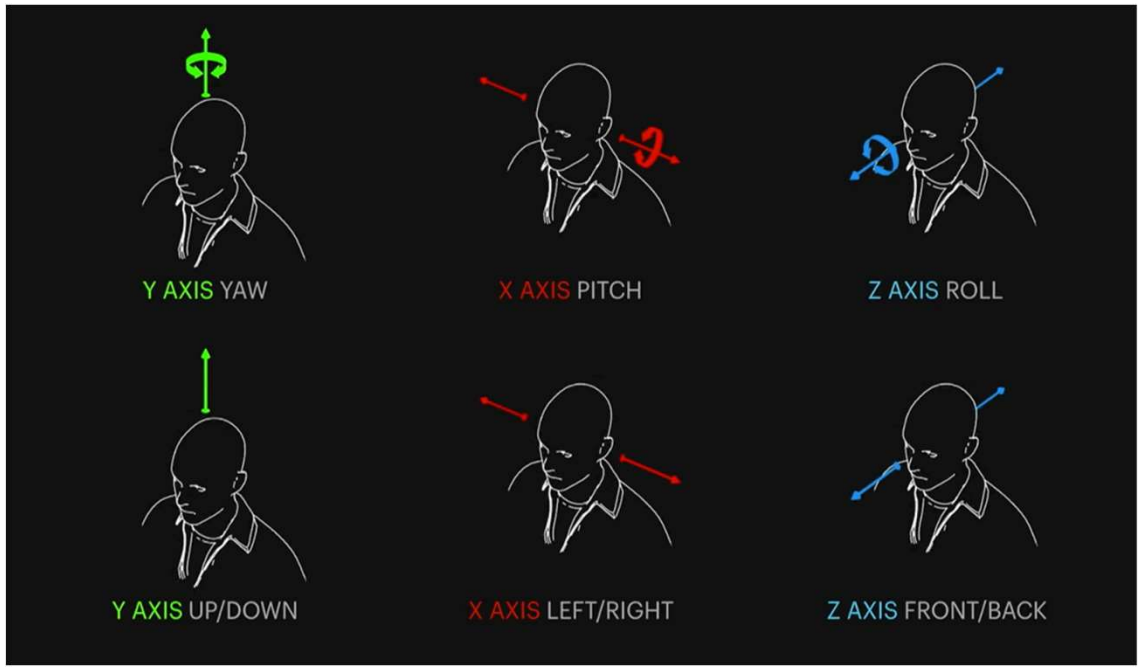


VOLUMETRIC MEDIA

FROM 2D TO IMMERSIVE VIDEO



WHAT DOES VOLUMETRIC MEDIA PROVIDE?



Adds parallax → multiple views

Six Degrees of Freedom

USE CASES

Holographic Collaboration & Conferencing [1]

Tele-Surgery & Remote Patient Monitoring [2]

Remote Industrial Monitoring & Management [3]

[1] S. Gunkel, H. M. Stokking, M. J. Prins, N. van der Stap, F. B. ter Haar, and O. A. Niamut, Virtual reality conferencing: multi-user immersive VR experiences on the web. In Proceedings of the 9th ACM Multimedia Systems Conference (MMSys '18). Association for Computing Machinery, New York, NY, USA, 498–501.

[2] J. Heyse, M. Torres Vega, T. De Jonge, F. De Backere, and F. De Turck, A Personalised Emotion-Based Model for Relaxation in Virtual Reality. *Appl. Sci.* **2020**, *10*, 6124.

[3] M. Torres Vega, et al., Immersive Interconnected Virtual and Augmented Reality: A 5G and IoT Perspective. J Netw Syst Manage 28, 796–826 (2020).

Holograms are not science-fiction anymore

Star Wars: Episode IV – A New Hope (1977)

- Local application
- Bandwidth requirements (Gb/s) too high for remote access
  → Compression

newsround

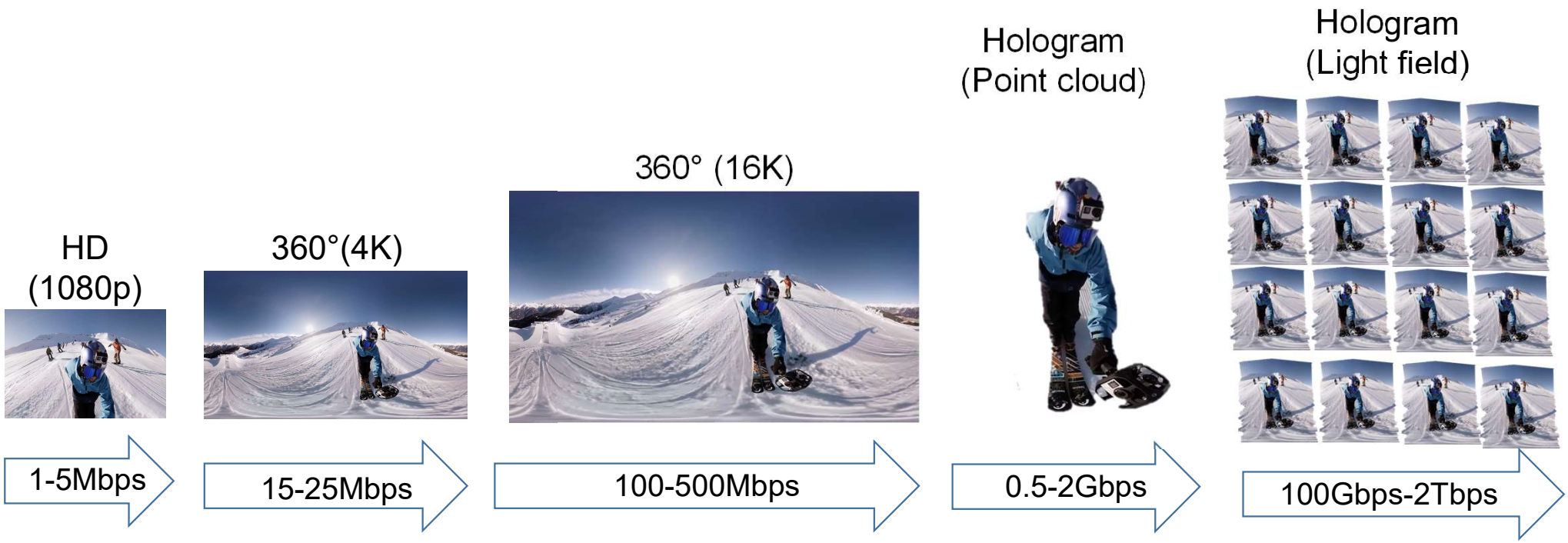
A German circus uses holograms instead of animals and it looks amazing!

6 Jun 2019 Last updated at 19:50

A German circus has become the first in the world to use holograms instead of real animals in its acts.

https://www.bbc.co.uk/newsround/48543263

ULTRA-HIGH BANDWIDTH REQUIREMENT [1]



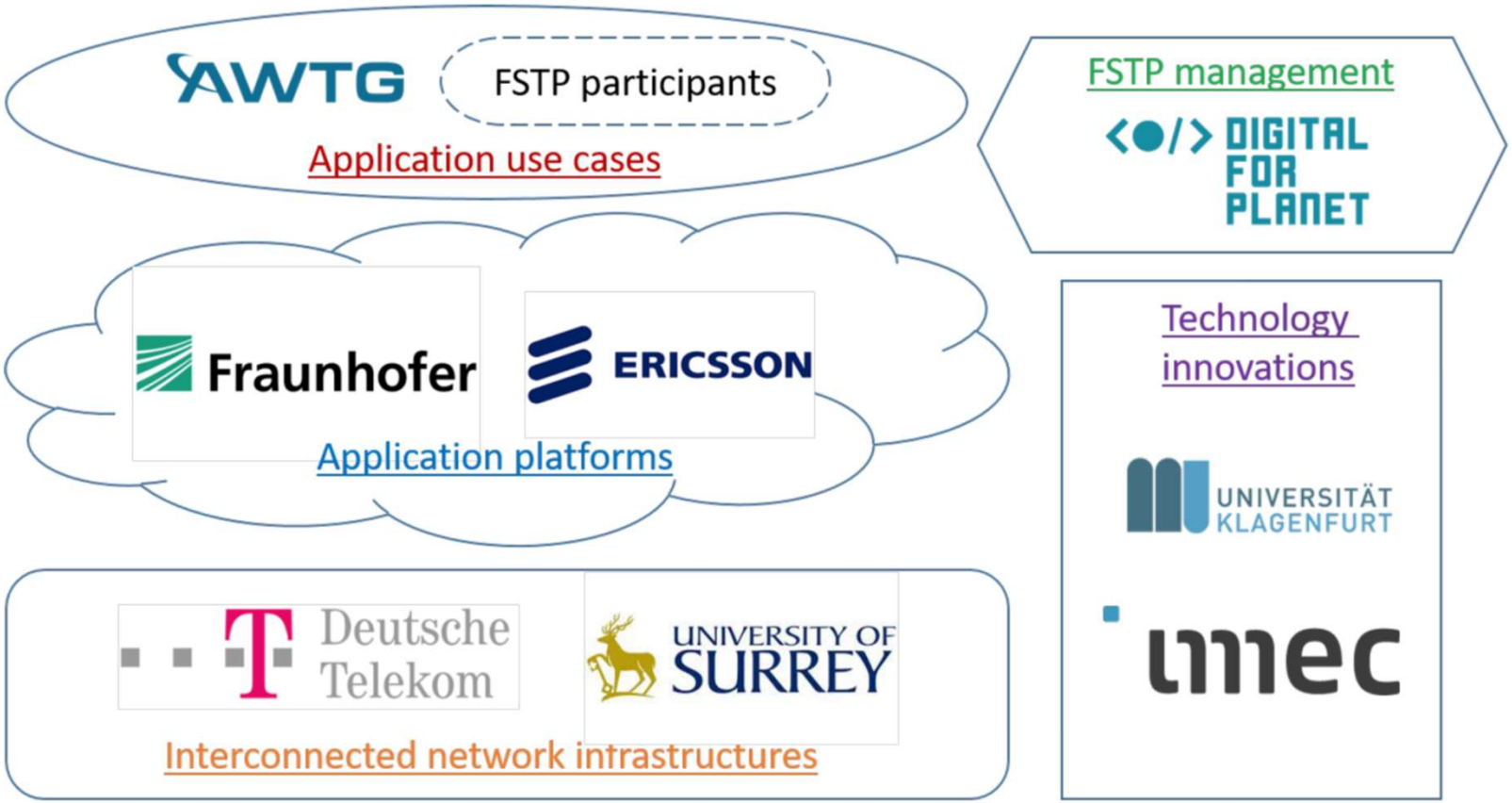
[1] A. Clemm, M. Torres Vega, H. K. Ravuri, T. Wauters and F. De Turck, "Towards Truly Immersive Holographic-Type Communication: Challenges and Solutions," in *IEEE Communications Magazine*, vol. 58, no. 1, pp. 93-99, January 2020.

EU SPIRIT PROJECT

2022-2025

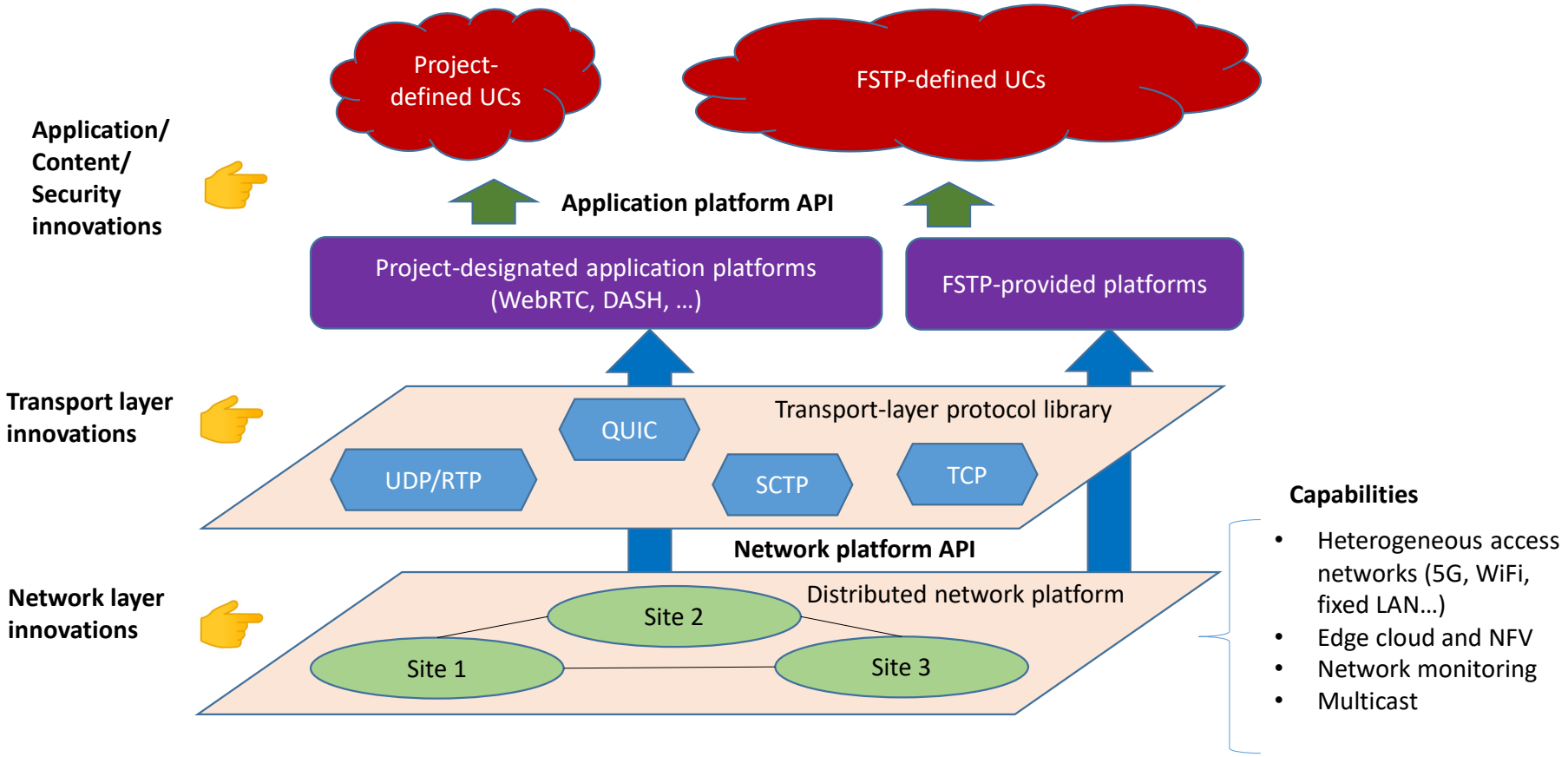
https://www.spirit-project.eu/

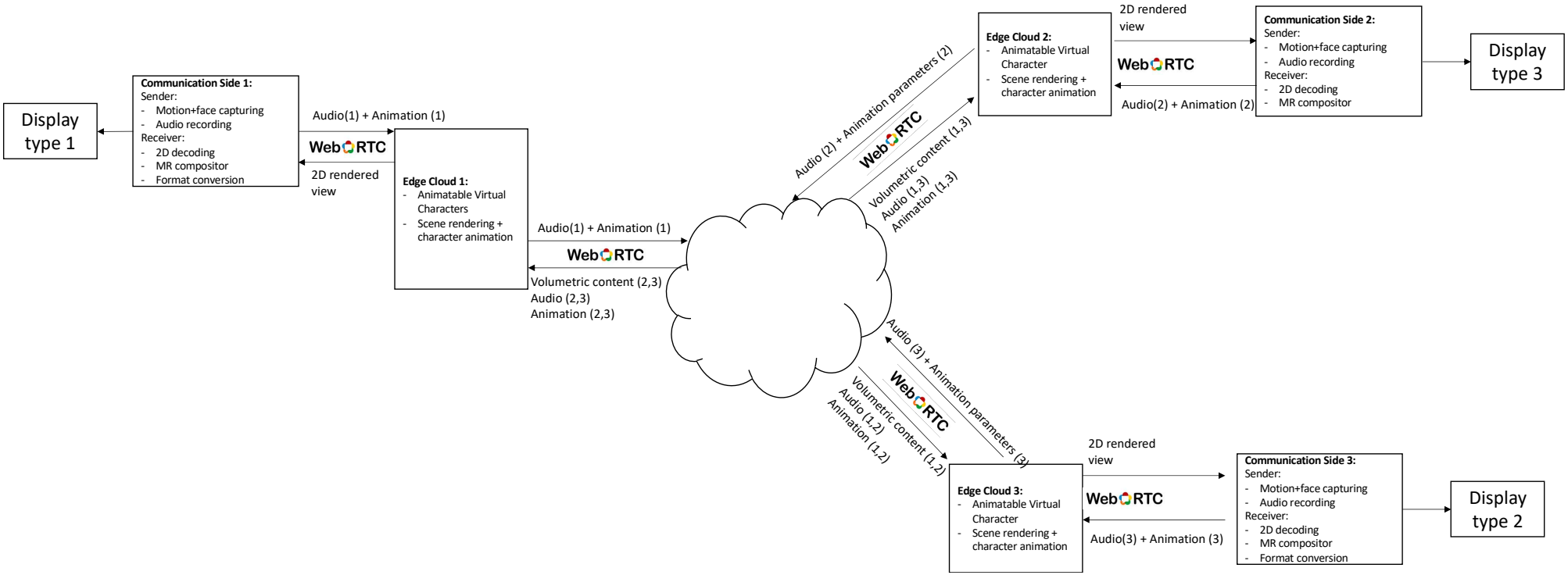
SPIRIT CONSORTIUM

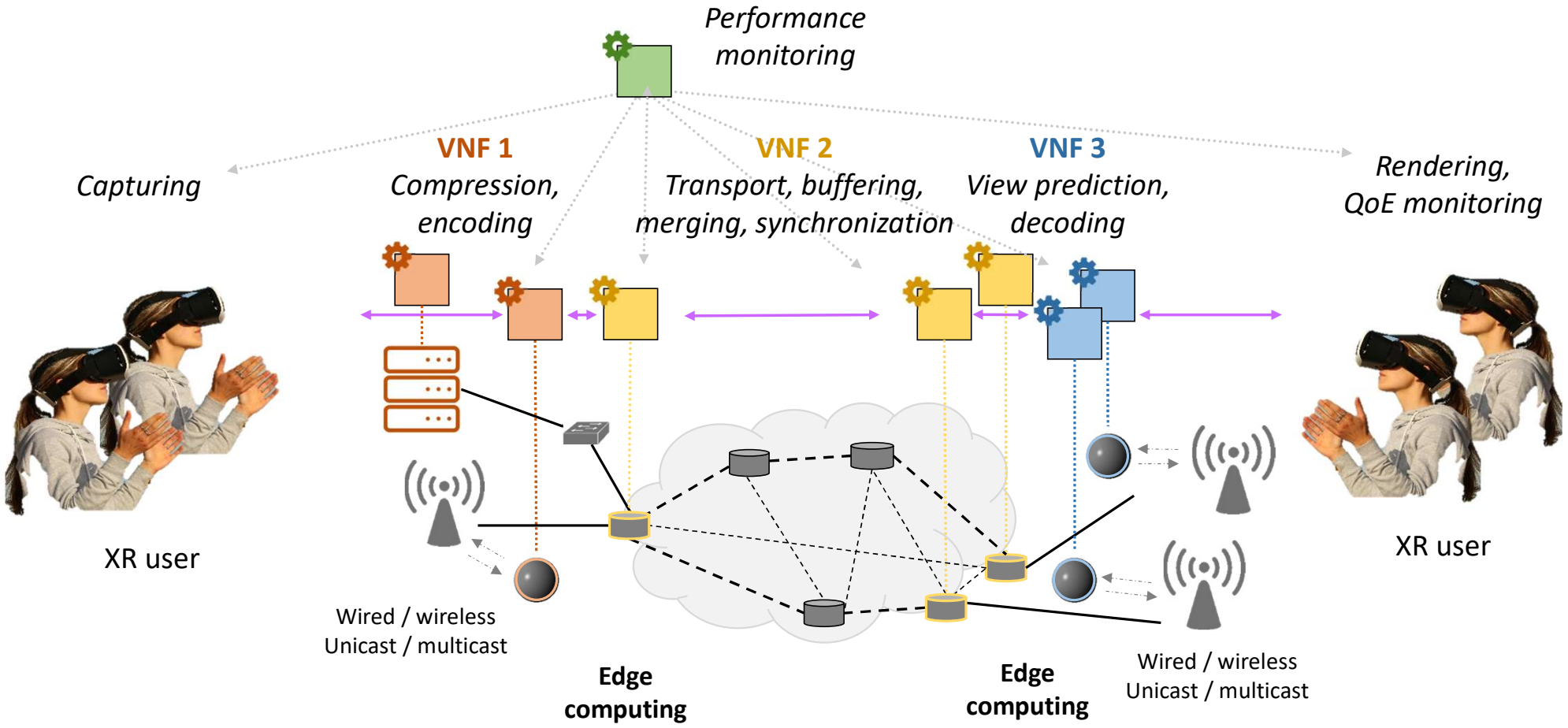


https://www.spirit-project.eu/

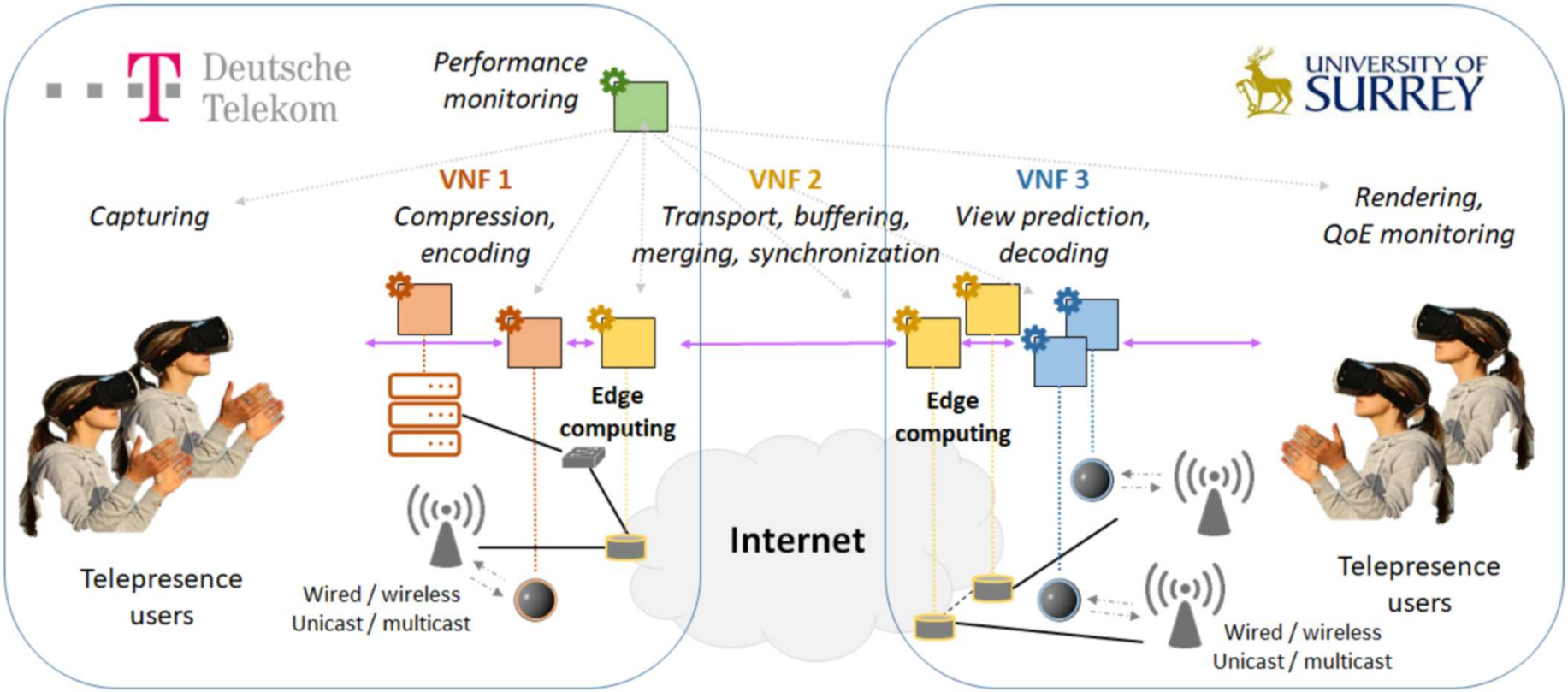
SPIRIT: SCALABLE PLATFORM FOR INNOVATIONS ON REAL-TIME IMMERSIVE TELEPRESENCE




Display type 1
Communication Side 1:
Sender:
- Motion+face capturing
- Audio recording
Receiver:
- 2D decoding
- MR compositor
- Format conversion
Audio(1) + Animation (1)
WebRTC
2D rendered view
Edge Cloud 1:
- Animatable Virtual Characters
- Scene rendering + character animation
Audio(1) + Animation (1)
WebRTC
Volumetric content (2,3)
Audio (2,3)
Animation (2,3)
Audio (2) + Animation parameters (2)
WebRTC
Volumetric content (1,3)
Audio (1,3)
Animation (1,3)
Edge Cloud 2:
- Animatable Virtual Character
- Scene rendering + character animation
2D rendered view
WebRTC
Audio(2) + Animation (2)
Communication Side 2:
Sender:
- Motion+face capturing
- Audio recording
Receiver:
- 2D decoding
- MR compositor
Display type 3
Audio (3) + Animation parameters (3)
WebRTC
Volumetric content (1,2)
Audio (1,2)
Animation (1,2)
Edge Cloud 3:
- Animatable Virtual Character
- Scene rendering + character animation
2D rendered view
WebRTC
Audio(3) + Animation (3)
Communication Side 3:
Sender:
- Motion+face capturing
- Audio recording
Receiver:
- 2D decoding
- MR compositor
- Format conversion
Display type 2


Performance monitoring
Capturing
VNF 1
Compression, encoding
VNF 2
Transport, buffering, merging, synchronization
VNF 3
View prediction, decoding
Rendering, QoE monitoring
XR user
XR user
Wired / wireless
Unicast / multicast
Edge computing
Edge computing
Wired / wireless
Unicast / multicast

SPIRIT PLATFORM



https://www.spirit-project.eu/

SPIRIT INTERNAL USE CASES

- Real-time human-to-human interactions
  - Holographic conversations

A dad spending time with his child

An athlete getting help from their doctor with an injury

A customer seeing her new car with a salesperson explaining the features

- Real-time human-machine interactions
  - Human-initiated
    - multi-site machine supervision, supported by autonomous mobile robots for intralogistics
  - Machine-initiated
    - autonomous mobile robots alert humans to solve contextual problems

SPIRIT FINANCIAL SUPPORT TO THIRD PARTIES (FSTP)

- Two Open Calls
  - 3.5 M euro
  - OC1 (2nd project year)
    - 10 projects
    - <= 200K euro funding each
  - OC2 (3rd project year)
    - 15 projects
    - <= 100K euro funding each
  - Target application domains
    - healthcare, retail, education, training, entertainment, manufacturing, tourism

https://www.spirit-project.eu/

CAN CURRENT TRANSPORT AND APPLICATION INFRASTRUCTURES DEAL WITH VOLUMETRIC MEDIA DELIVERY?

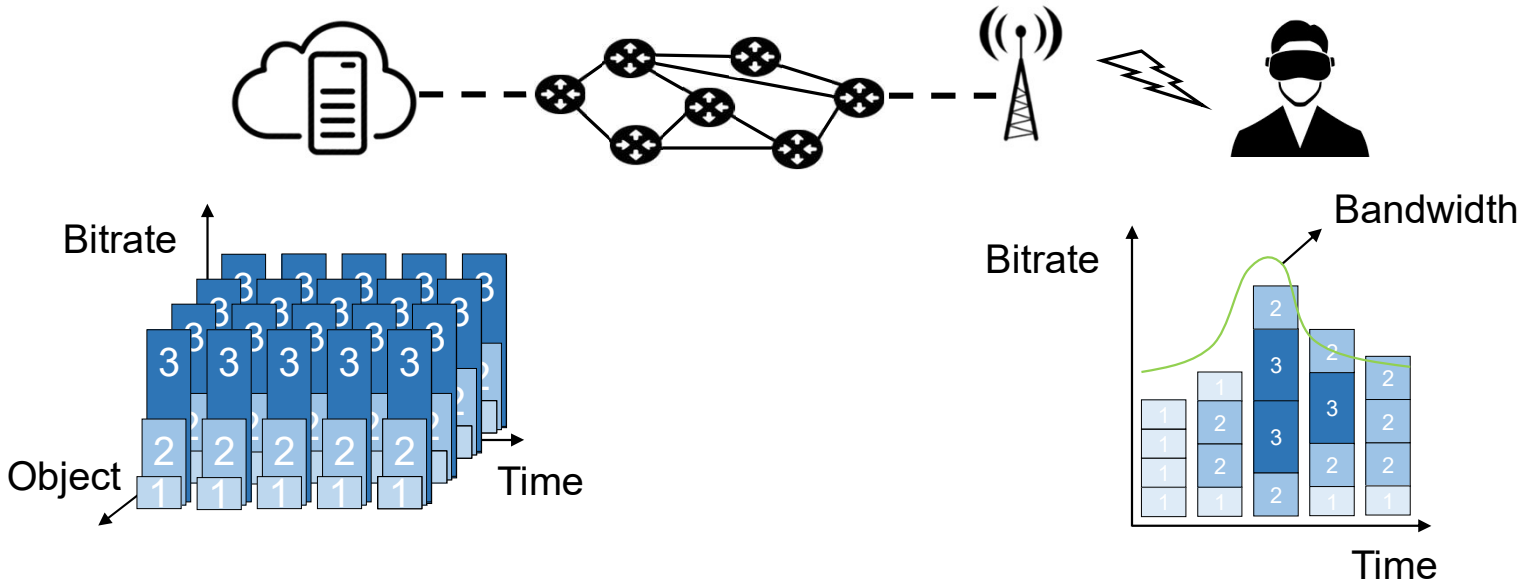
DYNAMIC POINT-CLOUD SCENES REQUIRE A SIGNIFICANT AMOUNT OF DATA

4.1 Gb/s

5.6 Gb/s

Streaming a scene with 4 similar objects would require 19.2 Gb/s!

WE CAN STREAM IT USING ADAPTIVE STREAMING

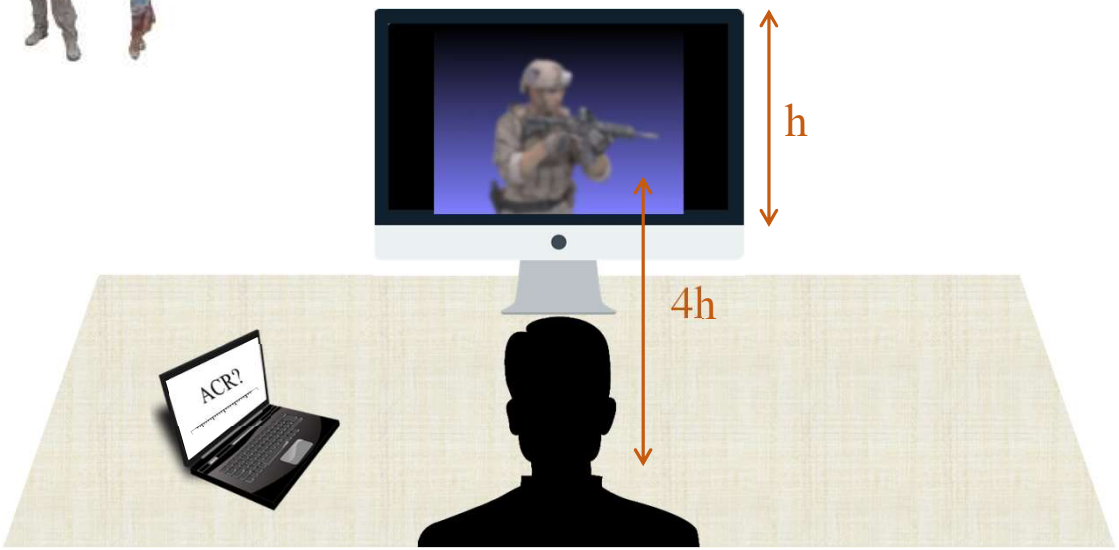


WE CAN EVALUATE THE QUALITY BY MEANS OF SUBJECTIVE STUDIES…

- Four Point Clouds
  (Raw data rate 19Gb/s)

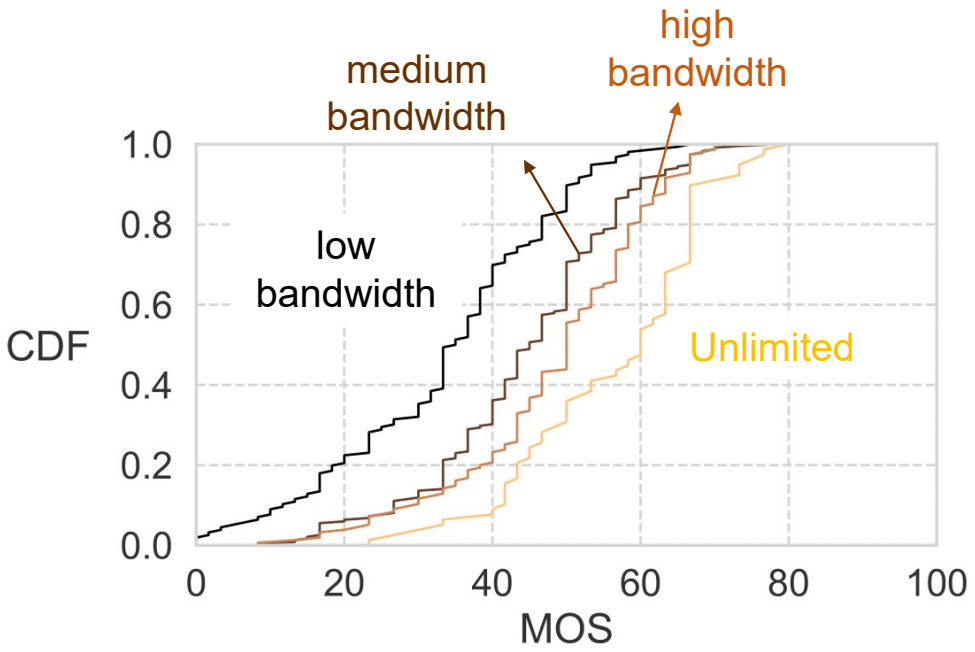
- Three different video sequences

- 8 configurations/video
  (bandwidth, allocation, prediction)



Single Stimulus (30 subjects)

EVALUATION OF SUBJECTIVE QOE OUR OBSERVATIONS [1]



- Subjects can distinguish between different bitrates
- However, the difference in QoE is not significant
- People do not perceive delivered data as good quality: MOS < 80% (4) in all cases and close to 60% in average

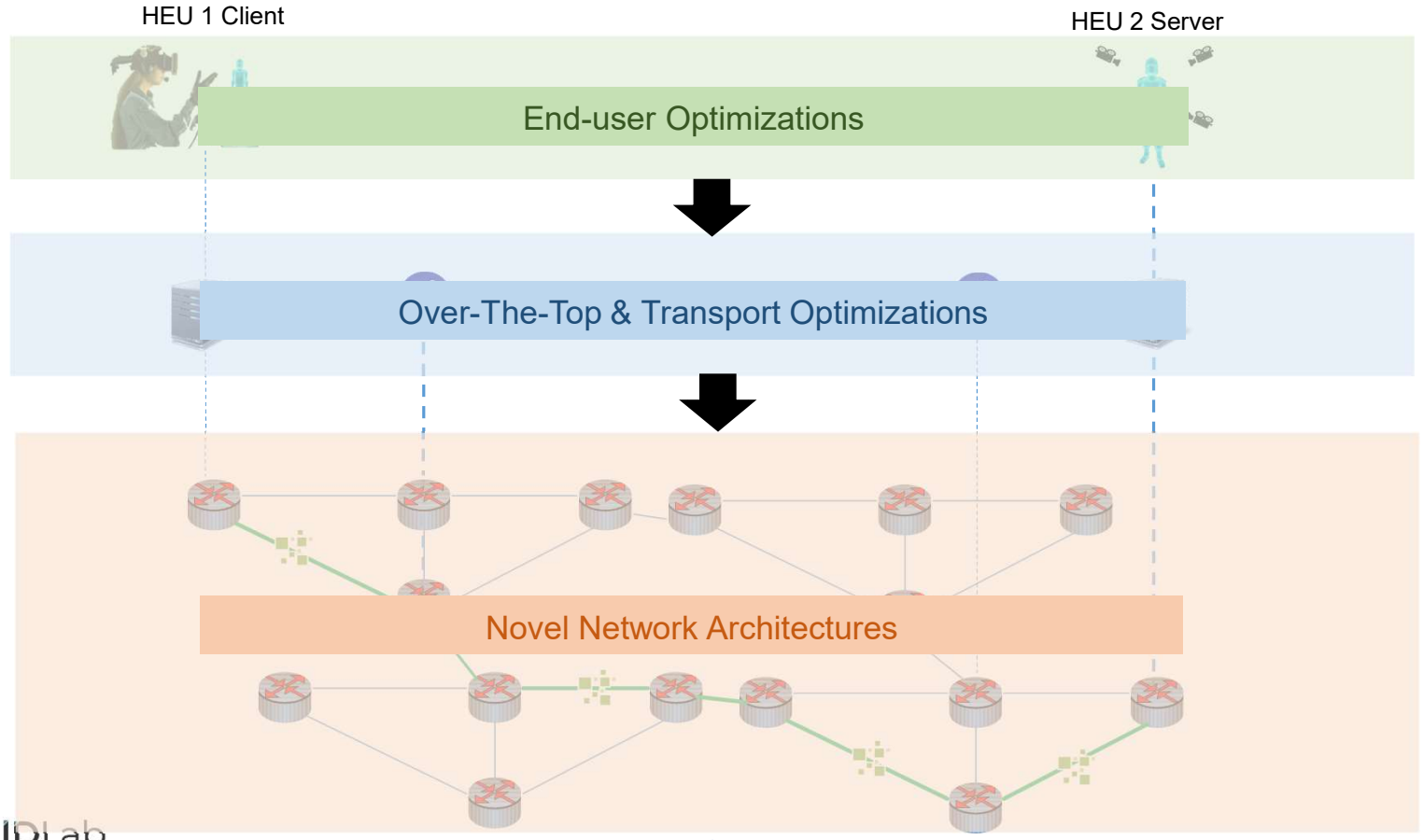
[1] J. van der Hooft, M. Torres Vega, C. Timmerer, A. C. Begen, F. De Turck and R. Schatz, "Objective and Subjective QoE Evaluation for Adaptive Point Cloud Streaming," 2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX), 2020.

WHAT DOES IT MEAN TO THE NETWORK?

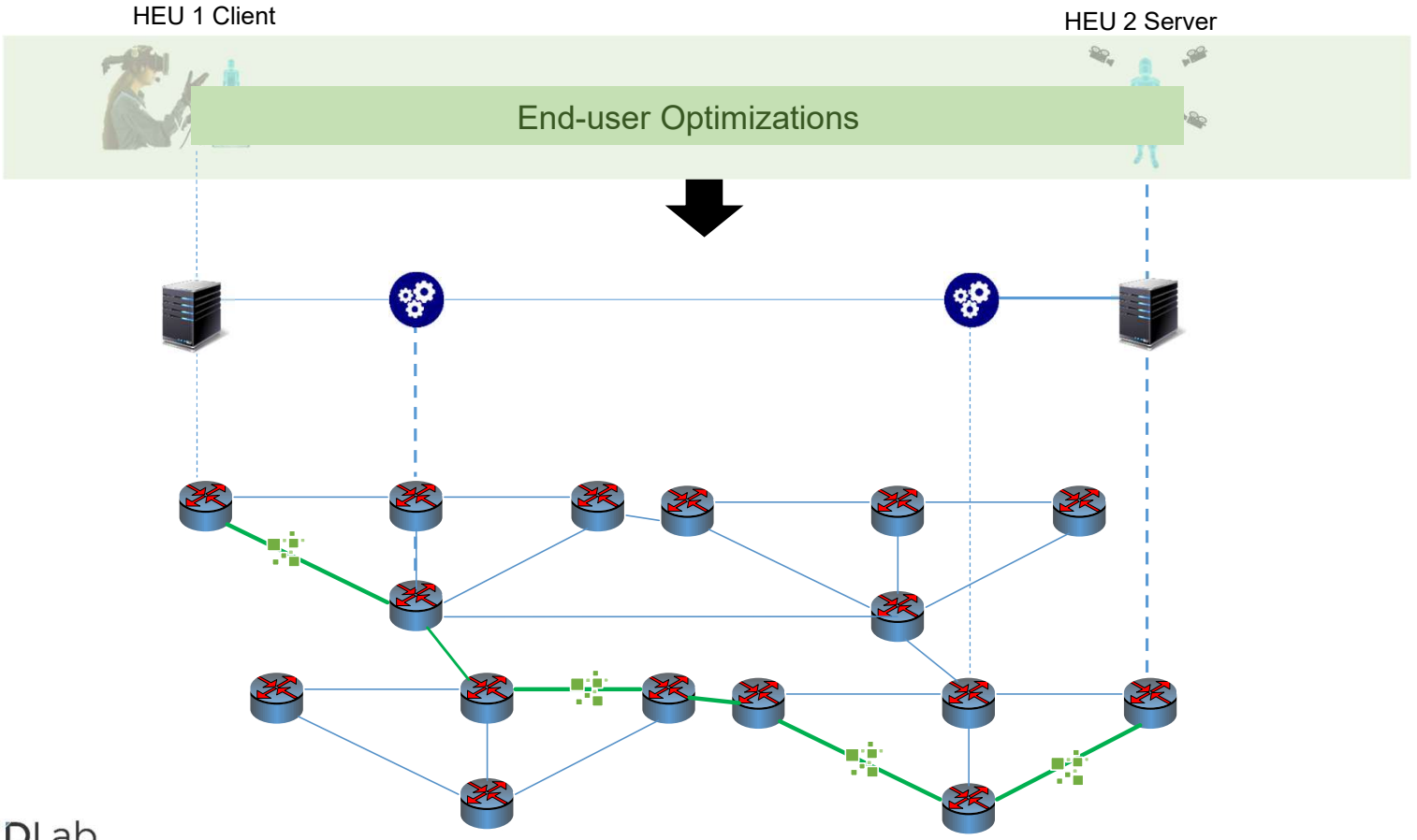
- End-user based or over the top optimizations are not sufficient to satisfy the user (MOS < 60%)
- These techniques do not cater to the latency requirement
- Network layer needs to complement the application layer approaches
- Cross-layer based end-to-end architecture for volumetric media delivery

HOW TO ACHIEVE TRULY IMMERSIVE VOLUMETRIC DELIVERY? A CROSS-LAYER APPROACH

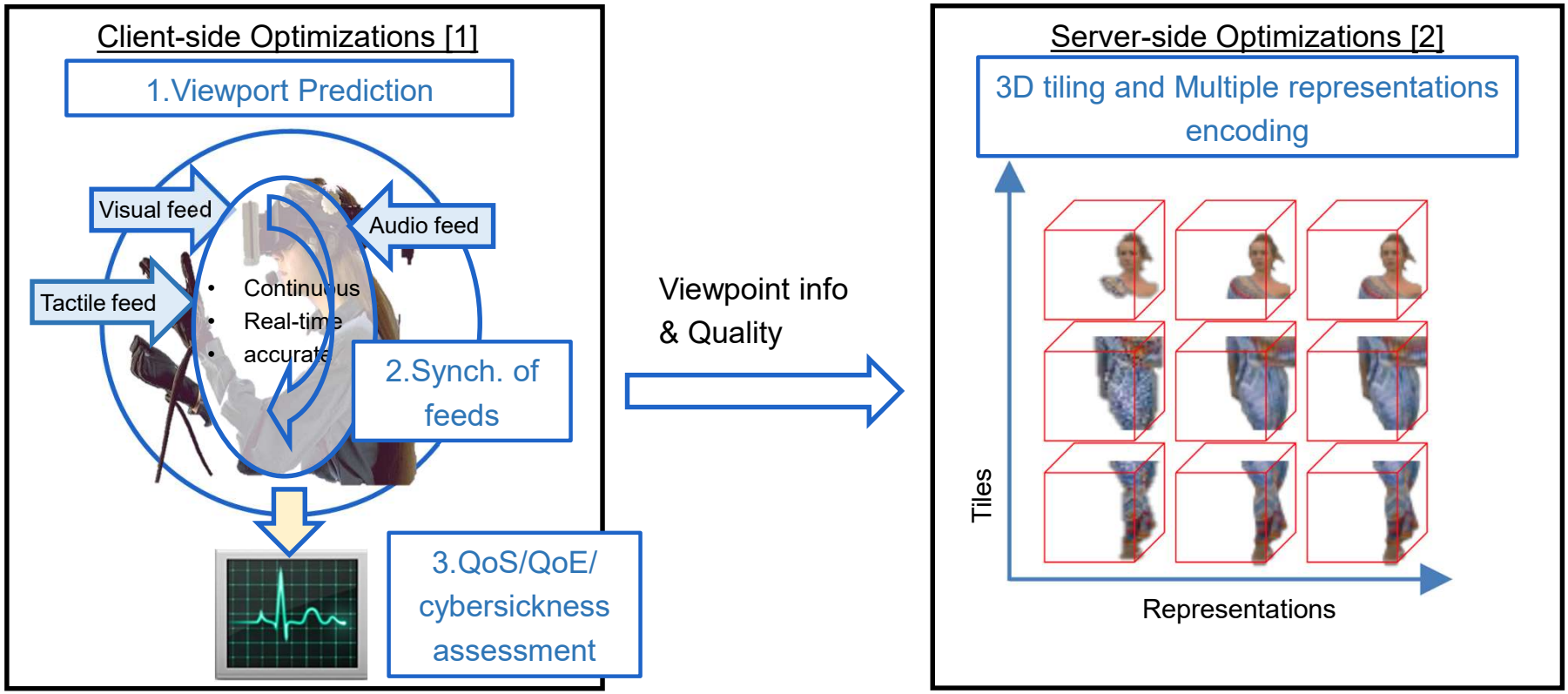
TRULY IMMERSIVE HTC: A CROSS-LAYER APPROACH



A CROSS-LAYER APPROACH: END-USER



END-USER OPTIMIZATIONS

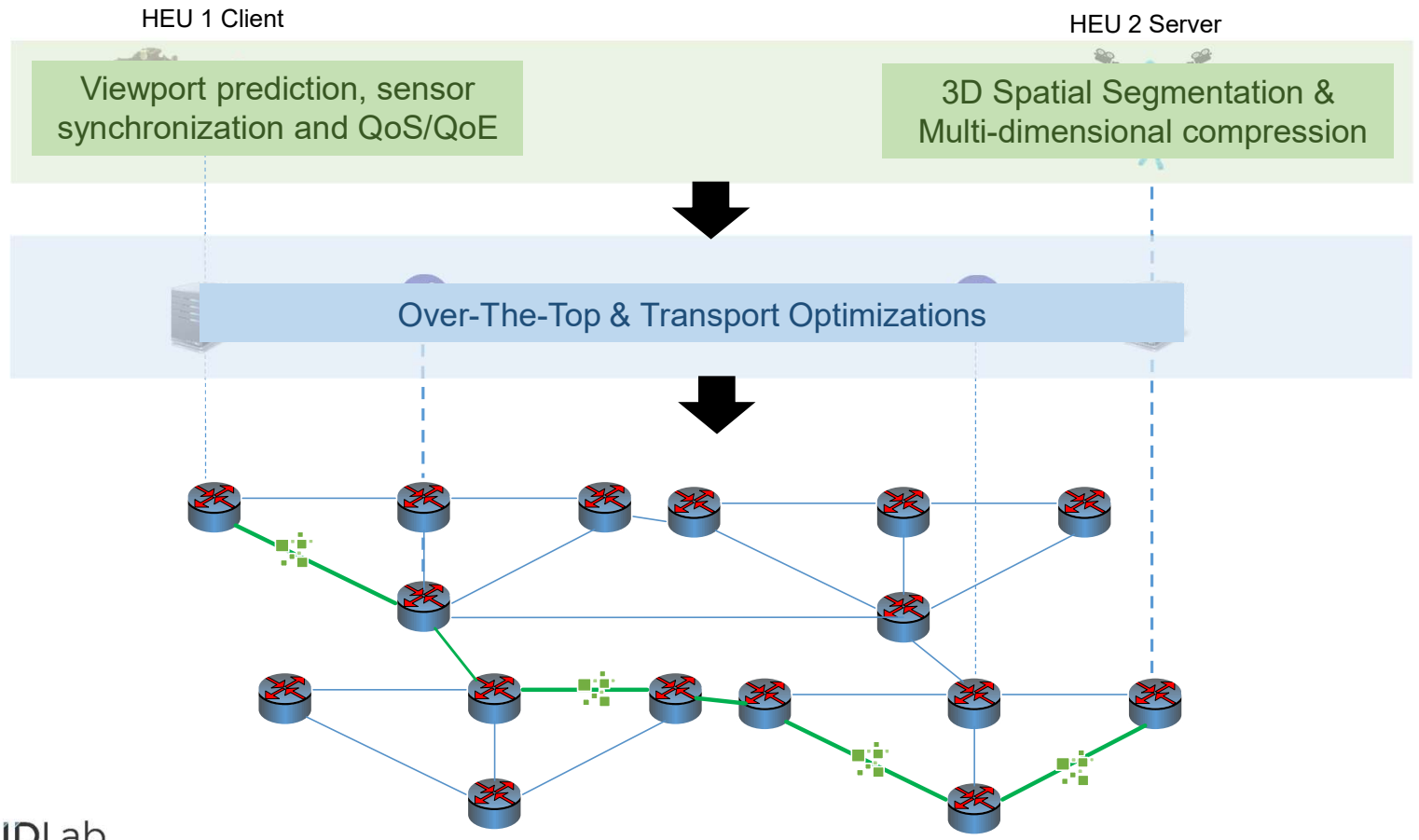


[1] S. Van Damme, M. Torres Vega and F. De Turck, "Human-centric Quality Management of Immersive Multimedia Applications," *2020 6th IEEE Conference on Network Softwarization (NetSoft)*, 2020, pp. 57-64

[2] J. Park, P. A. Chou, and J.-N. Hwang, "Rate-Utility Optimized Streaming of Volumetric Media for Augmented Reality," IEEE JETCAS, vol. 9, 2018, pp. 149–62

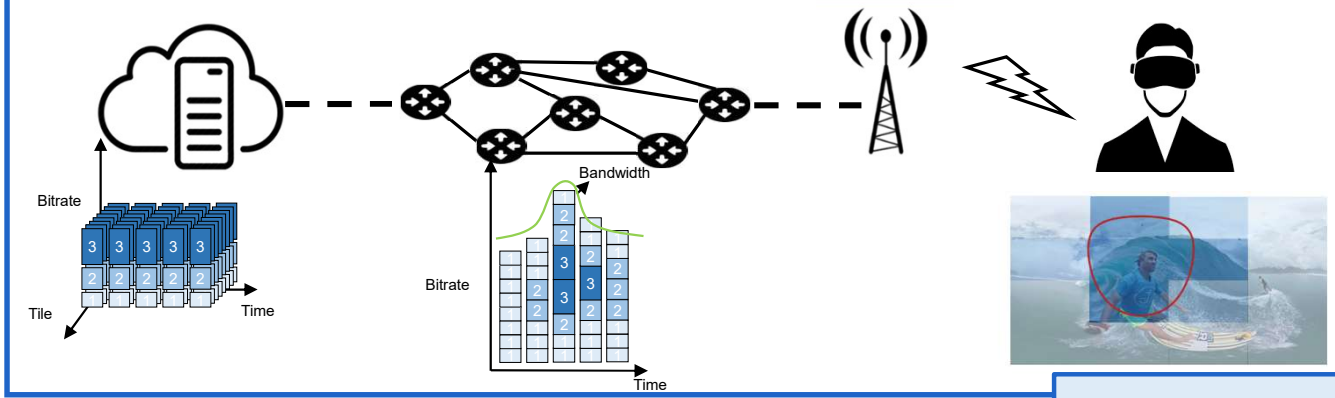
GHENT UNIVERSITY

A CROSS-LAYER APPROACH: TRANSPORT



CURRENT VIDEO TRANSMISSION: QUALITY VS DELAY

HTTP ADAPTIVE STREAMING: QUALITY OPTIMIZATION



☺ Quality and Bandwidth optimization

😐 Segmentation: no life

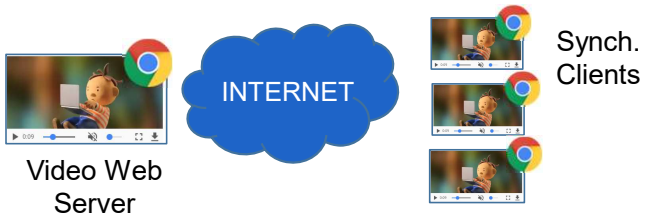
☹ Processing, buffering and protocol overhead: no real-time

Is it possible to get the best of both worlds?

RTP/UDP STREAMING: LATENCY OPTIMIZATION

RTP/UDP streaming

WebRTC & QUIC: Browse-based real time streaming

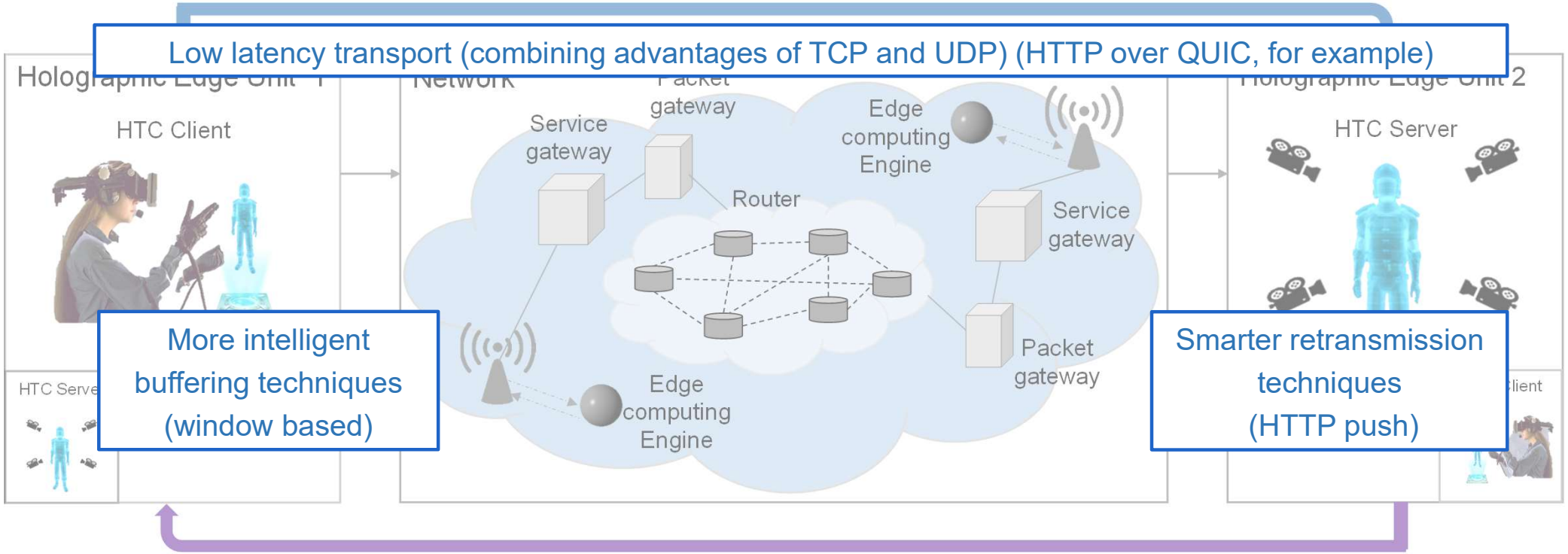


☺ Latency optimization

😐 Very limited quality control -> problem for HTC

☹ Very low resilience (packet loss prone)

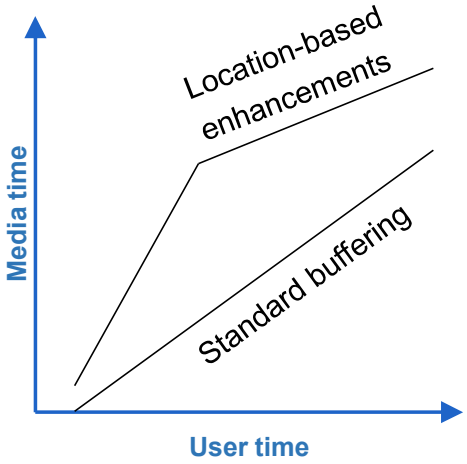
OVER THE TOP & TRANSPORT OPTIMIZATIONS



GHENT UNIVERSITY IDLab INTERNET & DATA LAB imec

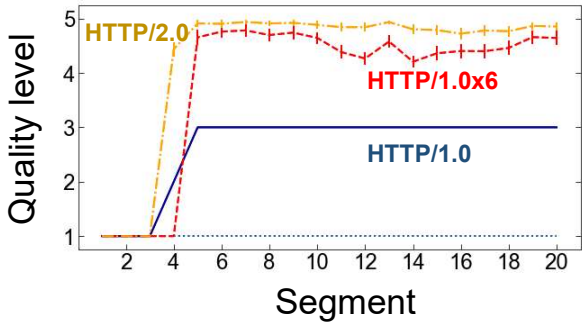
TRANSPORT AND OTT OPTIMIZATIONS

More intelligent buffering techniques: Window-based buffering [1]



- Buffer= interval or window
- Allows to access not only the end of the buffer.
- Able to respond with low latency to unexpected user interactions.

Smarter retransmission techniques: HTTP2 push for 360-degree video [2]

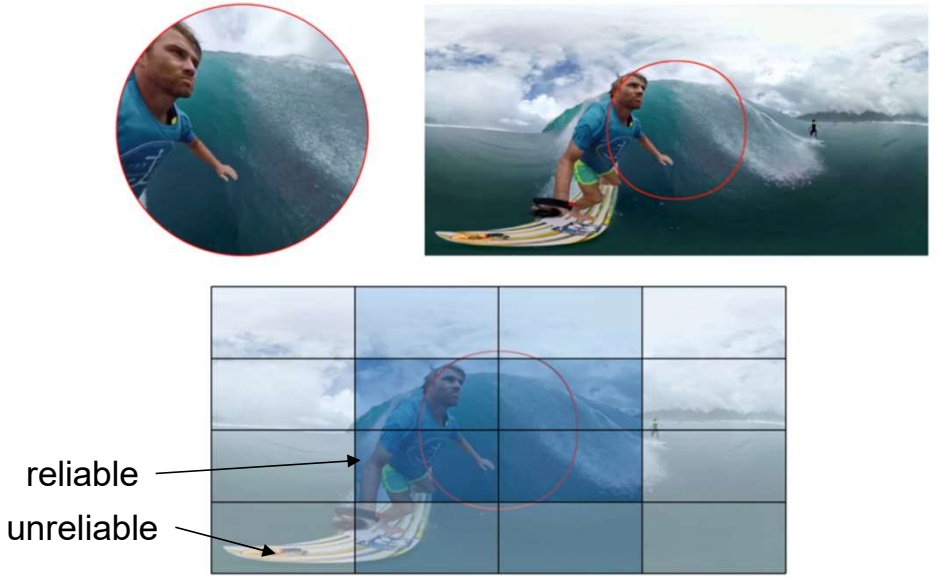


- Thanks to push, HTTP2 (yellow) can acknowledge multiple packets simultaneously
- It allows for highest and more stable quality delivery

[1] J. Park, P. A. Chou, and J.-N. Hwang, "Rate-Utility Optimized Streaming of Volumetric Media for Augmented Reality," IEEE JETCAS, vol. 9, 2018, pp. 149–62

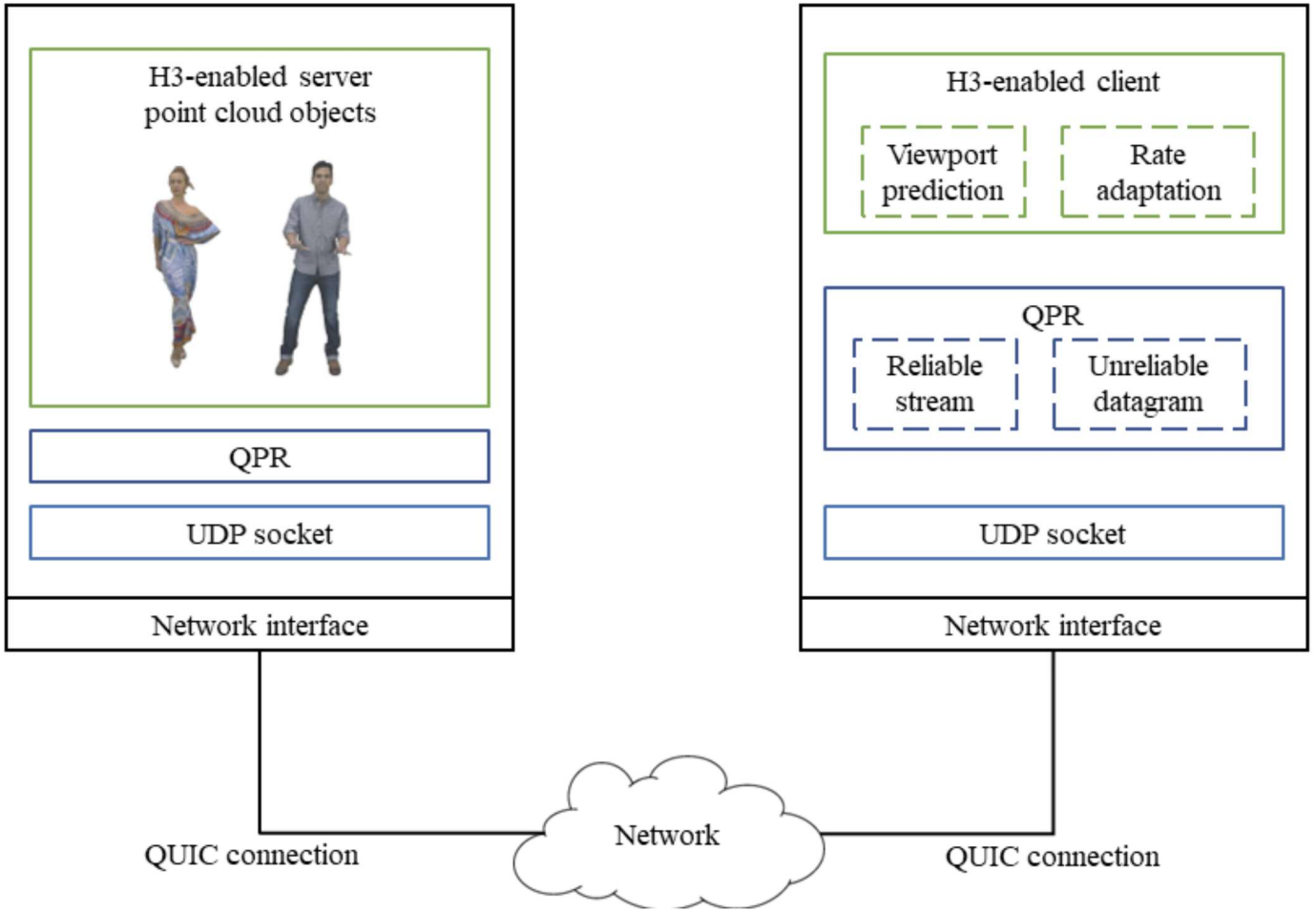
[2] J. van der Hooft, M. Torres Vega, S. Petrangeli, T. Wauters, and F. De Turck, "Tile-based Adaptive Streaming for Virtual Reality Video", ACM Trans. Multimedia Comput. Commun. Appl. 15, 4, Article 110 (January 2020), 24 pages.

TRANSPORT LAYER AND CONTROL LAYER INNOVATIONS



partially reliable QUIC
(QUIC-PR or QPR)

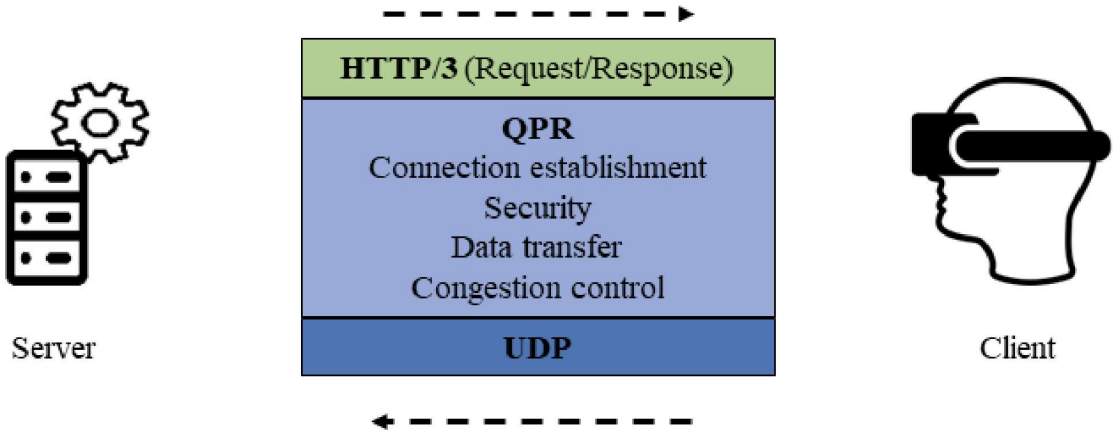
ADAPTIVE PARTIALLY RELIABLE DELIVERY OF IMMERSIVE MEDIA OVER QUIC-HTTP/3



GHENT UNIVERSITY

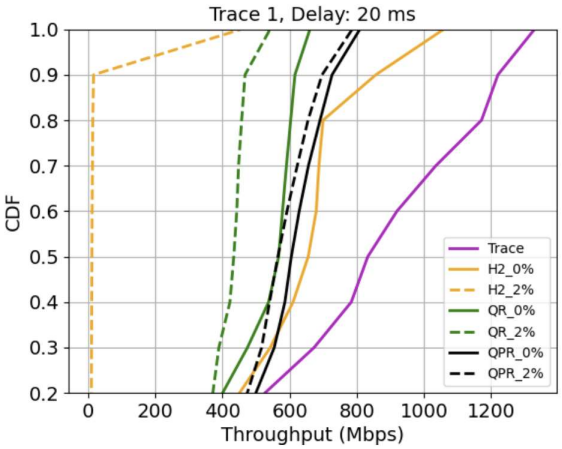
imec

ADAPTIVE PARTIALLY RELIABLE DELIVERY OF IMMERSIVE MEDIA OVER QUIC-HTTP/3

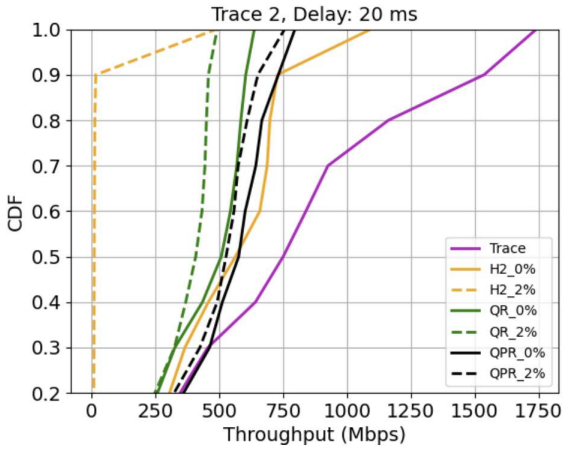


ADAPTIVE PARTIALLY RELIABLE DELIVERY OF IMMERSIVE MEDIA OVER QUIC-HTTP/3

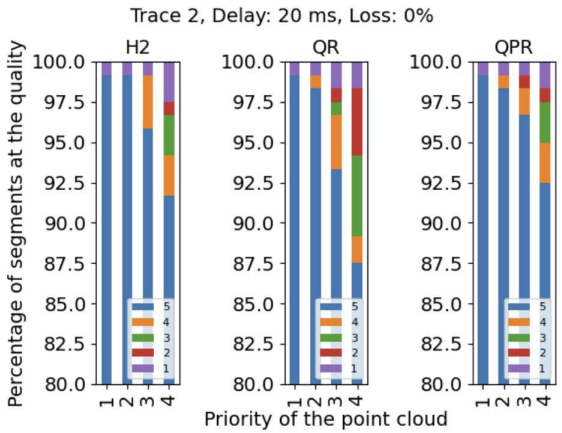
[1] H. K. Ravuri, M. Torres Vega, J. van der Hooft, T. Wauters, and F. De Turck, “Adaptive Partially Reliable Delivery of Immersive Media Over QUIC-HTTP/3”, IEEE Access, 2022.

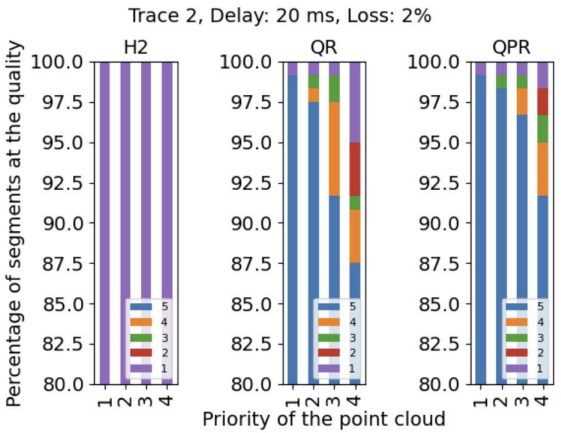


(a)



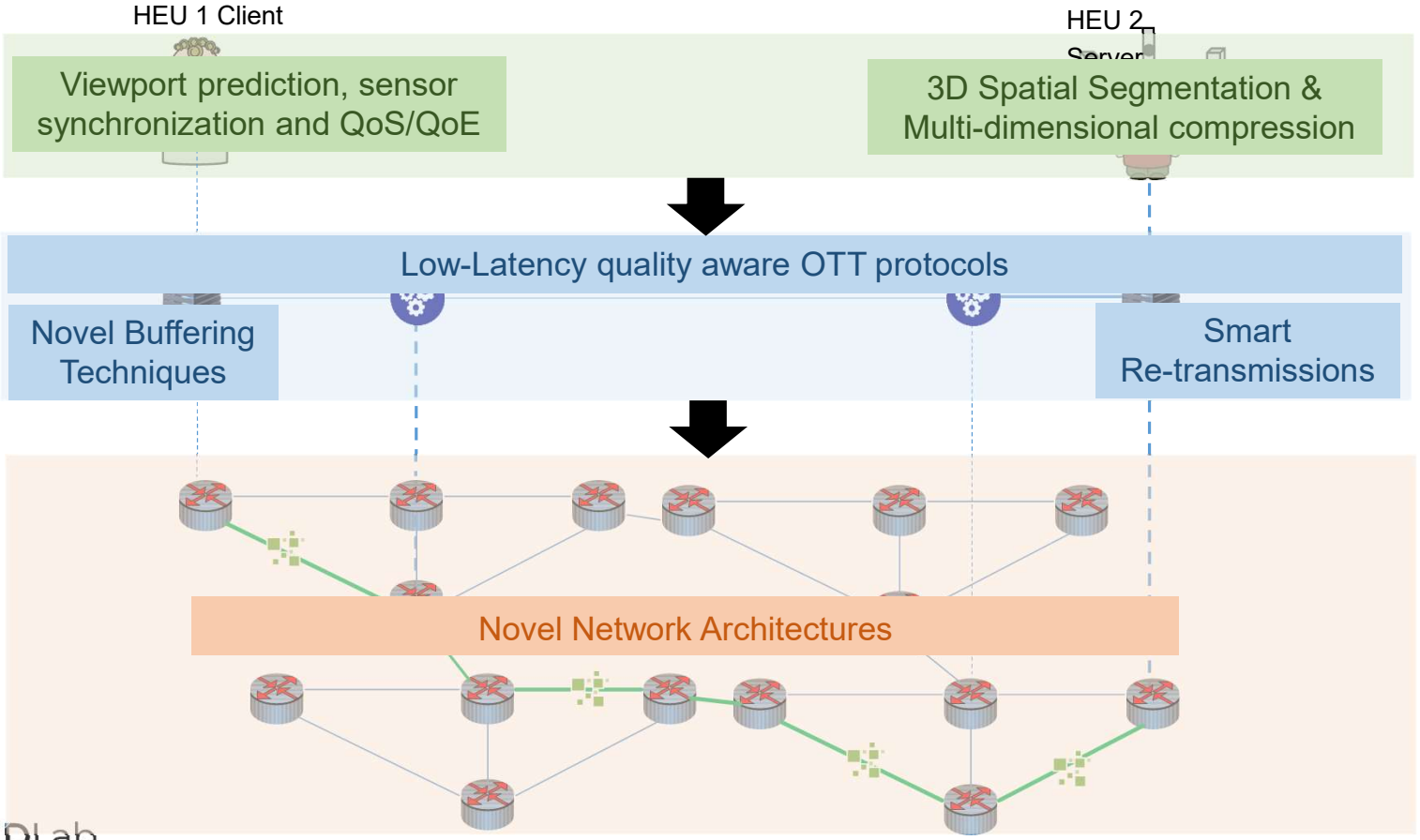
(b)



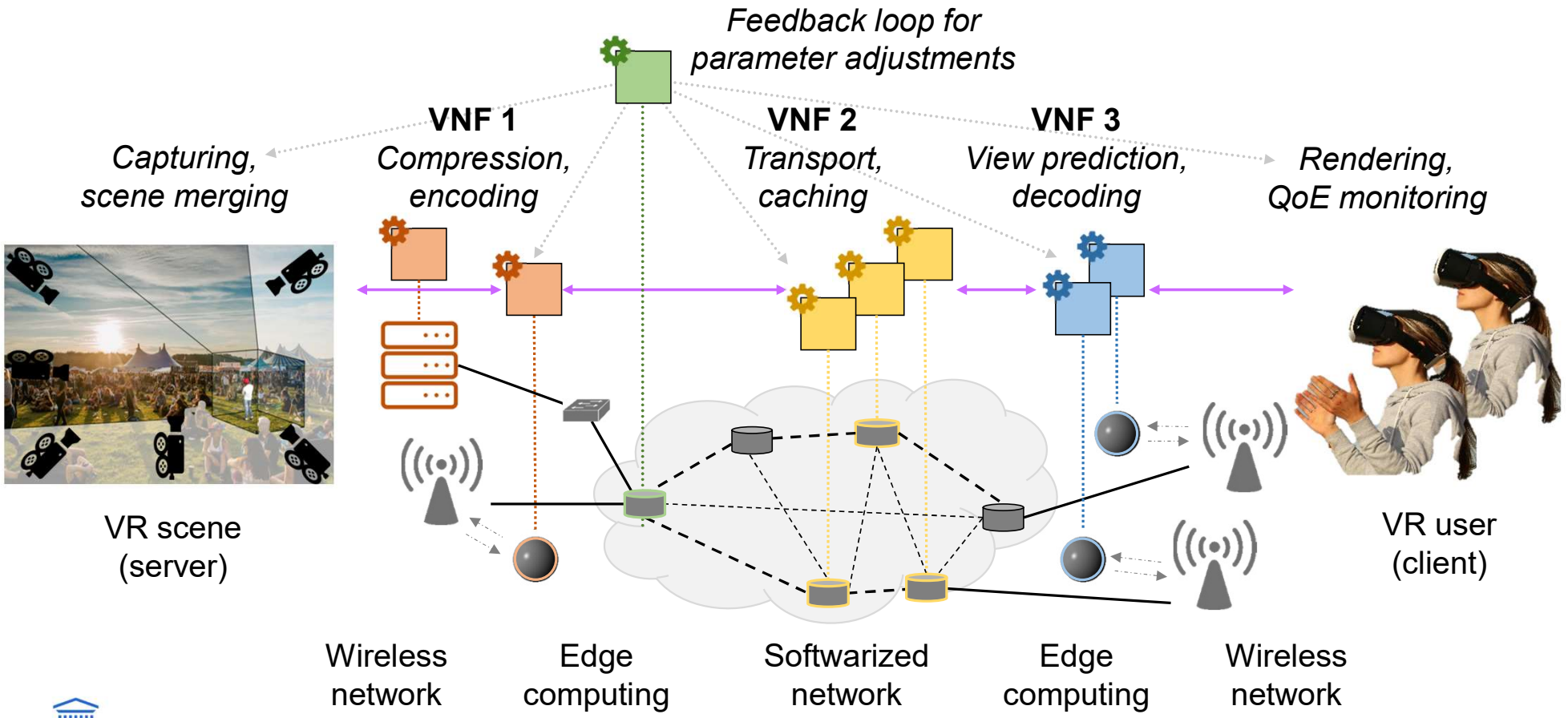


GHENT UNIVERSITY IDLab INTERNET & DATA LAB imec

TRULY IMMERSIVE HTC: A CROSS-LAYER APPROACH

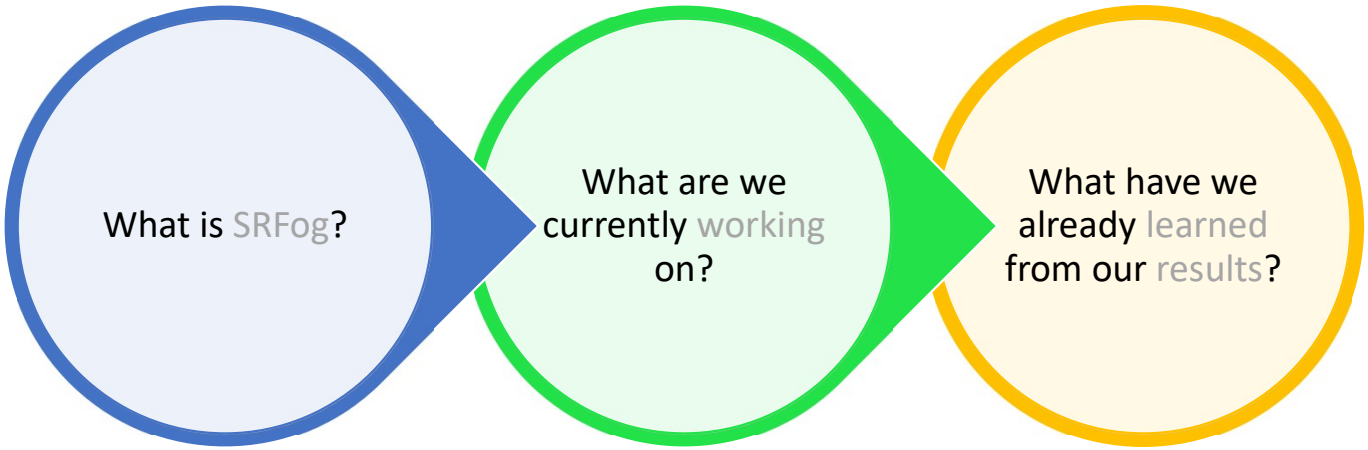


SERVICE FUNCTION CHAIN OPTIMIZATIONS



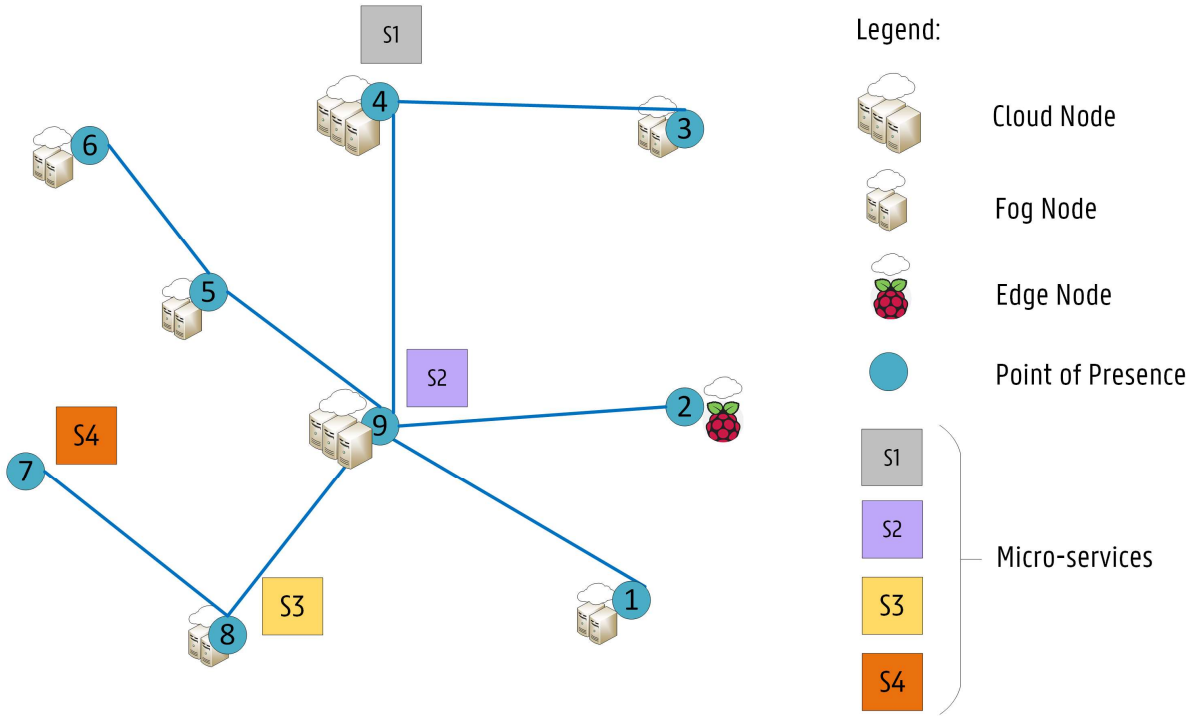
SRFog: A flexible architecture for Virtual Reality content delivery through Fog Computing and Segment Routing

SRFOG: OVERVIEW



GHENT UNIVERSITY IDLab INTERNET & DATA LAB imec

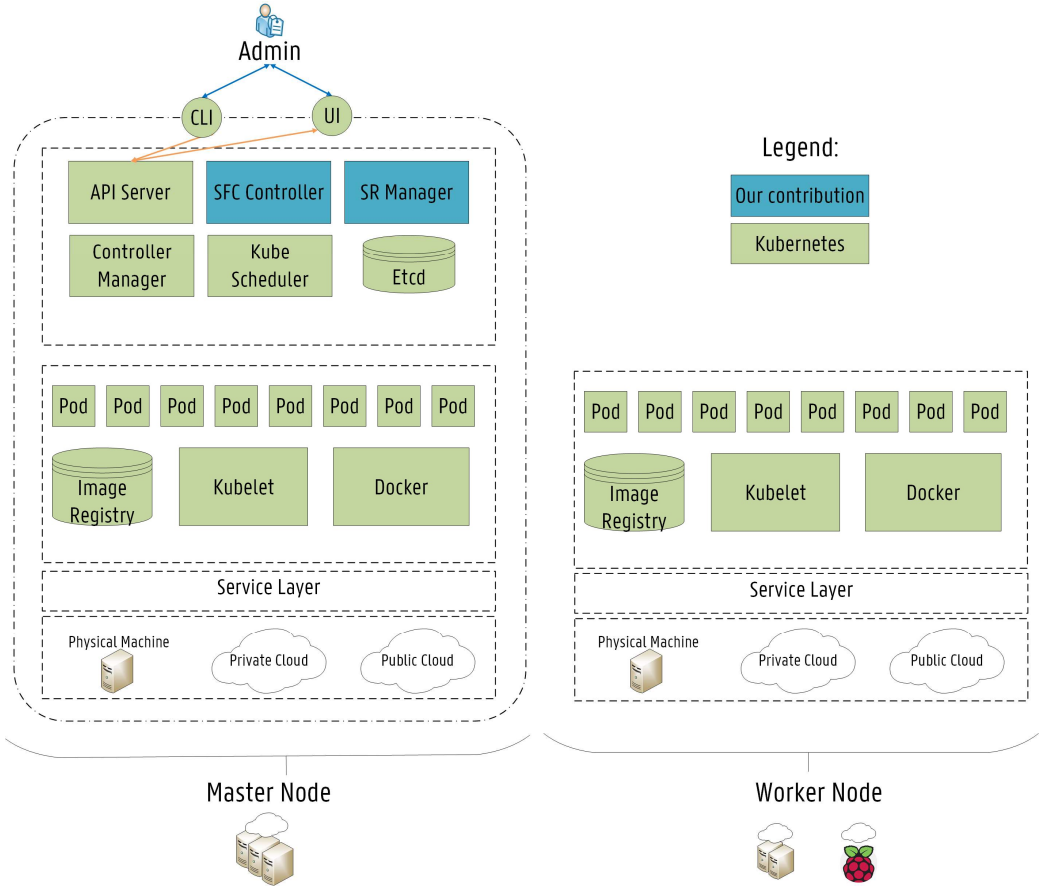
SRFog considers Fog Computing and micro-services



GHENT UNIVERSITY IDLab INTERNET & DATA LAB

imec

SRFog follows the Kubernetes architectural model

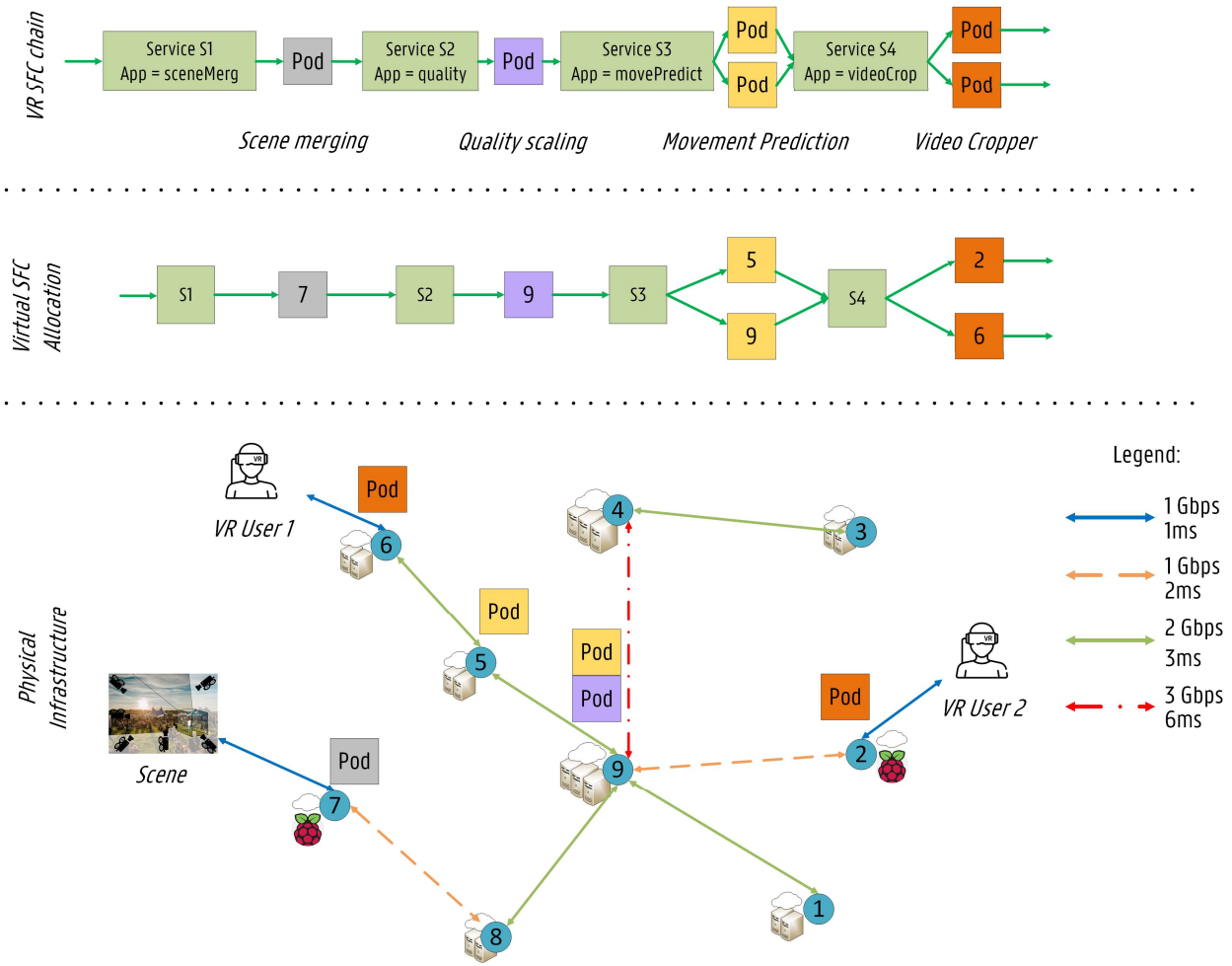


GHENT UNIVERSITY IDLab INTERNET & DATA LAB imec

SRFog adopts container-based service chains and traffic flow optimization based on Segment Routing

GHENT UNIVERSITY IDLab INTERNET & DATA LAB imec

High-level view of an illustrative SFC allocation



GHENT UNIVERSITY

IDLab
INTERNET & DATA LAB

imec

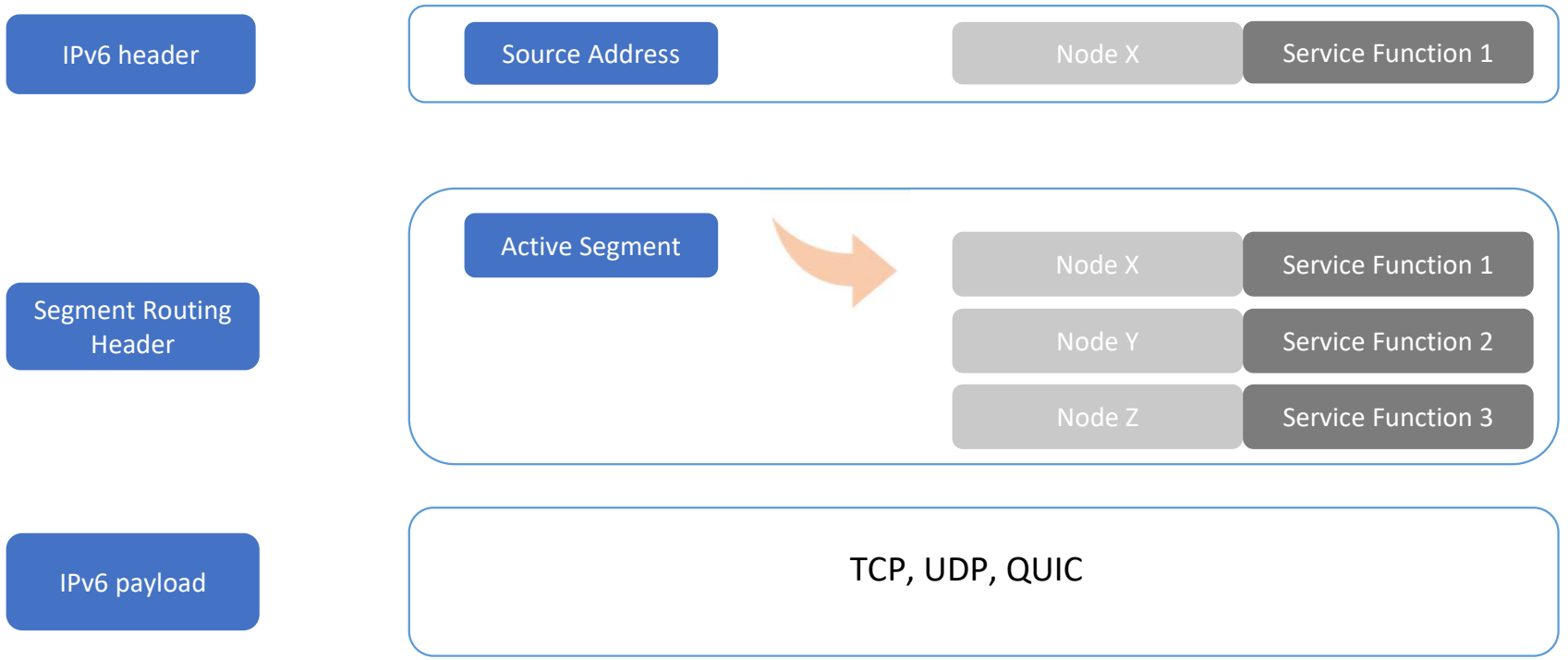
Segment Routing provides several advantages

- **Source Routing** → Topological and service chain path encoded in the packet header
- **Scalability** → The network fabric does not hold any per-flow state for Traffic Engineering
- **Simplicity** → Protocol elimination: LDP, RSVP-TE, NSH, VXLAN…

IETF standardization in progress:
– https://tools.ietf.org/html/draft-ietf-spring-srv6-network-programming-08

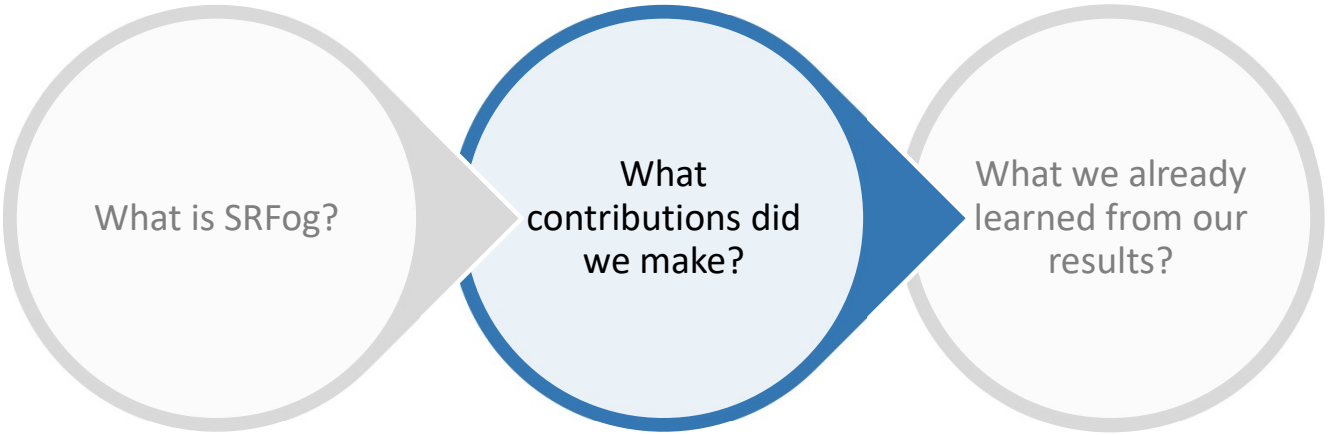
GHENT UNIVERSITY IDLab INTERNET & DATA LAB imec

Service Chaining with SRv6 (Packet Header)



GHENT UNIVERSITY IDLab INTERNET & DATA LAB

imec


What is SRFog?
What contributions did we make?
What we already learned from our results?

VR service chain allocation

Contributions:

- SFC controller for Kubernetes:

  Extending Kubernetes scheduling features to consider location and network-aware information

- MILP model for Fog-cloud infrastructures:

  Optimal service placement considering SFC, micro-service concepts in SRFog

GHENT UNIVERSITY IDLab INTERNET & DATA LAB imec

The MILP model considers different latency sources

$$\sum_{a \varepsilon A} \sum_{id \varepsilon ID} \sum_{s \varepsilon S} \sum_{\beta_i \varepsilon \beta} \sum_{n \varepsilon N} P_{s,\beta_i}^{a,id}(n) \times \varpi_n \times \omega_s \times \gamma_s \times \delta_s$$

MILP model (CPLEX)

Main Extension

Different Latency sources

$$\forall u \varepsilon U: \underbrace{\epsilon_u}_{\text{E2E latency}} = \underbrace{\zeta_U}_{\text{Scene}} + \underbrace{T_u}_{\text{Computing}} + \underbrace{\Lambda_u}_{\text{SFC}} + \underbrace{\psi_u}_{\text{User}}$$

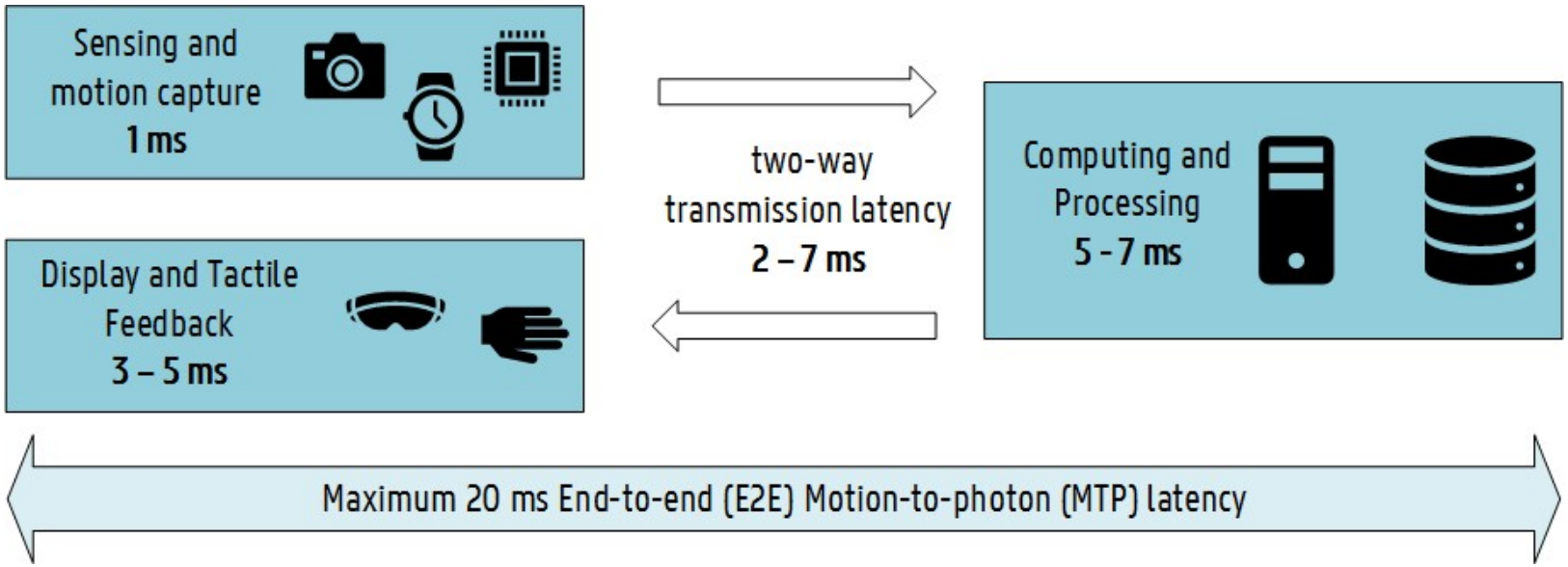
Objectives

- MIN. E2E Latency
- MIN. Deployment Costs
- MAX. User Video Quality

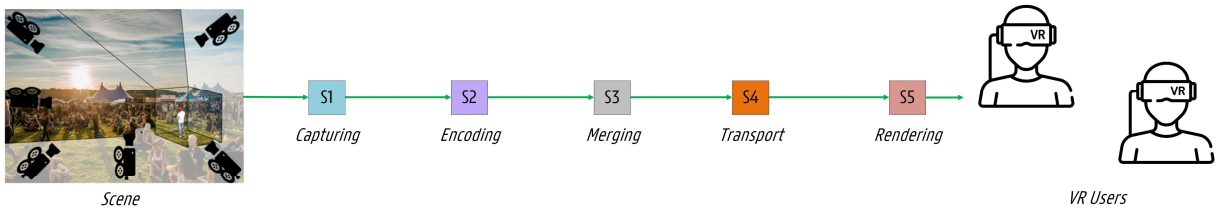
GHENT UNIVERSITY IDLab INTERNET & DATA LAB

imec

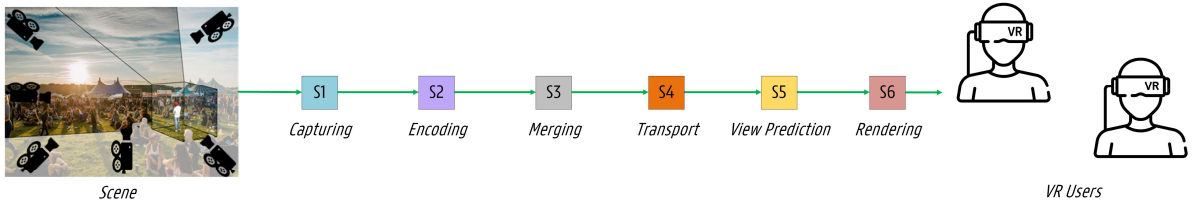
Future VR systems should meet the 20 ms E2E latency



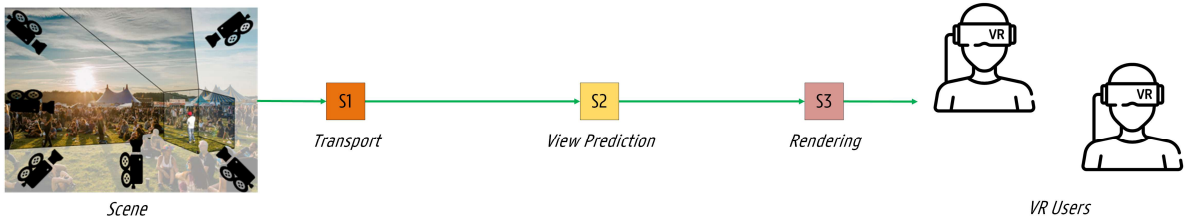
Three VR chains were designed



Live without View Prediction



Live with View Prediction



On-demand

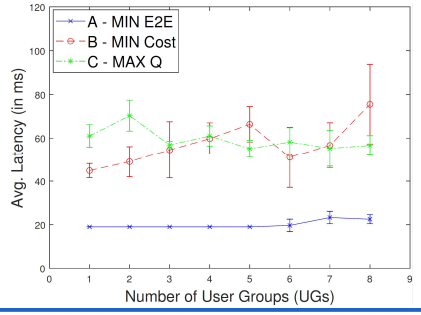
GHENT UNIVERSITY IDLab INTERNET & DATA LAB

imec

EFFECTS OF SFC ON E2E LATENCY [1]

Live without View Prediction

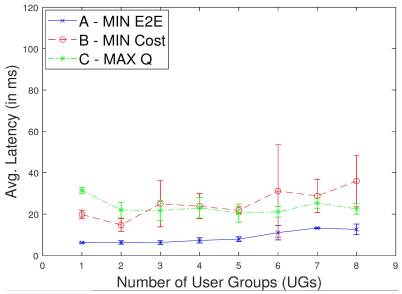
Capturing, encoding, merging, transport, rendering



Even optimizing latency (blue line), users experience an average latency of 19 ms even for a small number of user groups.

Live with View Prediction

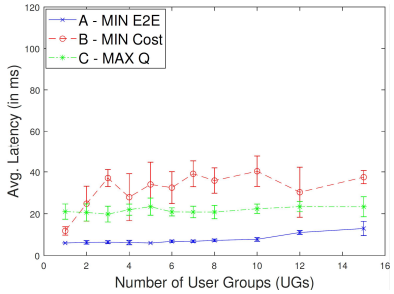
Capturing, encoding, merging, transport, view prediction, rendering



The addition of view prediction to the live scenario helps to reduce the user-perceived latency (blue line), but it brings additional costs (red line).

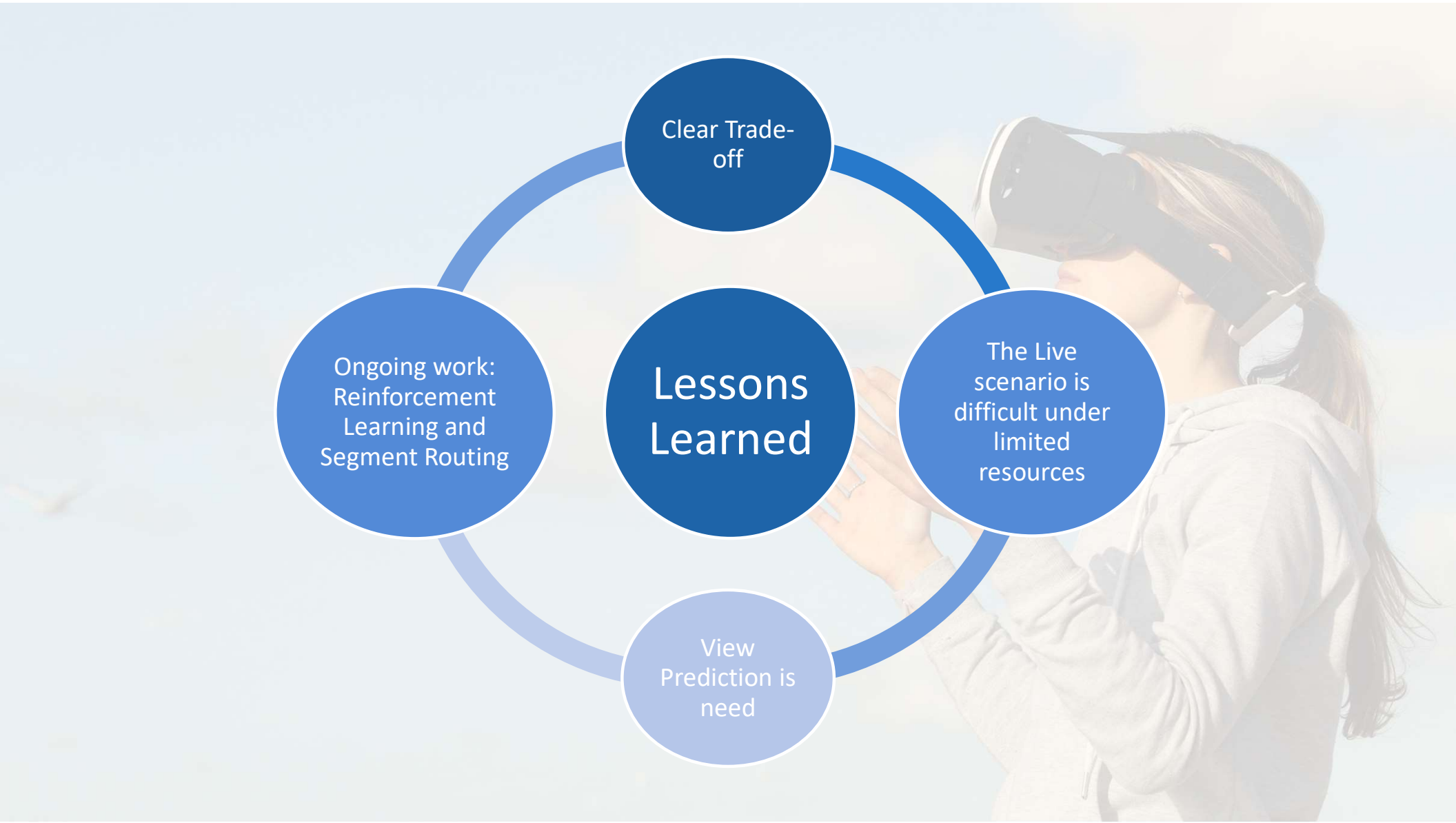
On-demand

Transport, view prediction, rendering

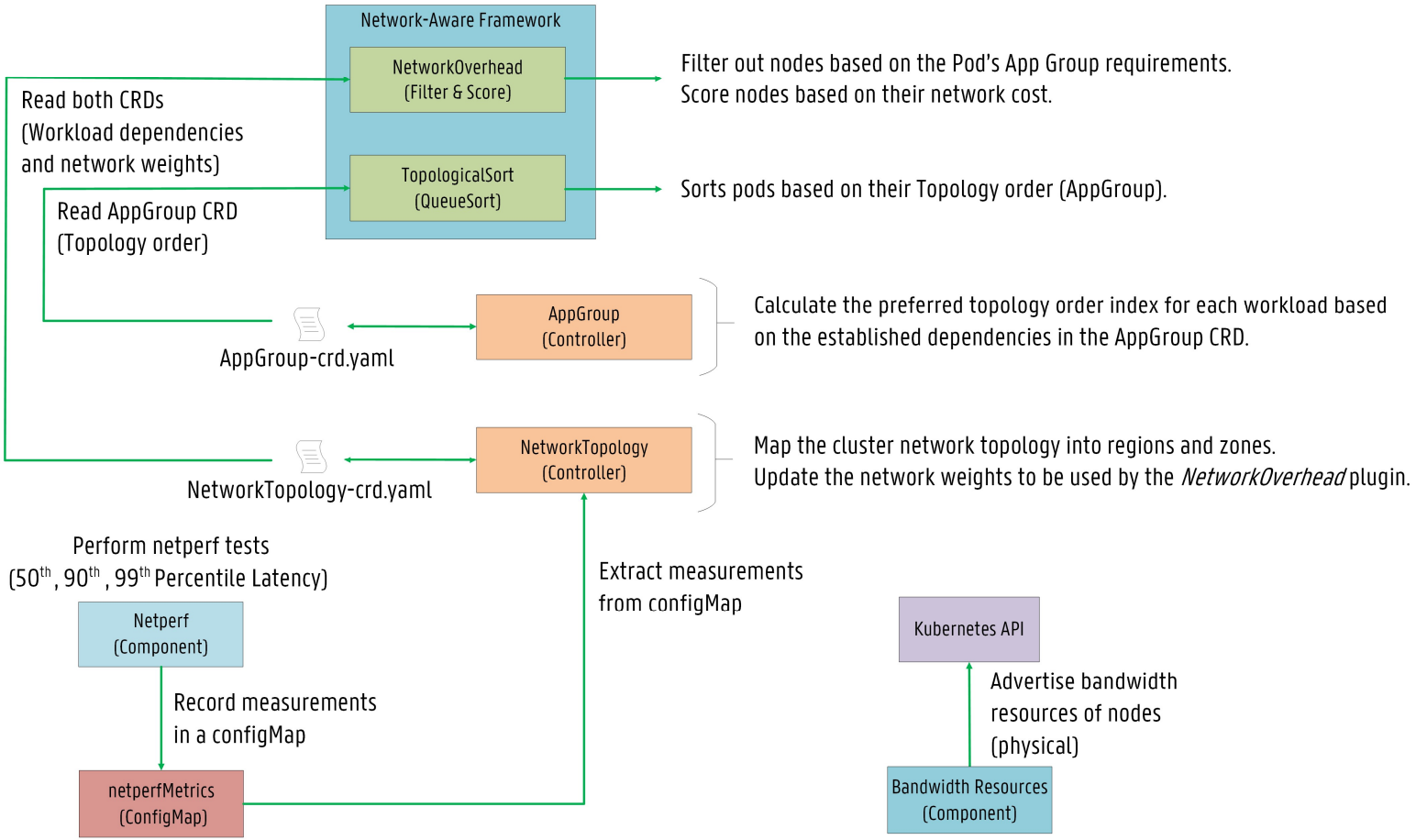


The on-demand scenario shows that deployment costs are considerably reduced, while keeping the E2E latency very low (3-4 ms), but it is not suited for live scenarios.

[1] J. Santos, J.van der Hooft, M. Torres Vega, T. Wauters, B. Volckaert, F. De Turck, Efficient orchestration of service chains in Fog Computing for immersive Virtual Reality, published in IEEE/IFIP/ACM CNSM 2021.


Clear Trade-off
The Live scenario is difficult under limited resources
View Prediction is need
Ongoing work: Reinforcement Learning and Segment Routing
Lessons Learned

Network-aware framework – Design Overview



This design introduces an end-to-end solution to model/weight a cluster's network latency and topological information, and leverage that to optimize the scheduling of latency- and bandwidth-sensitive workloads.

Netperf Component: https://github.com/jpedro1992/pushing-netperf-metrics-to-prometheus
Development repo: https://github.com/jpedro1992/scheduler-plugins/tree/KepDevWithNTController

Intent-Based Workflows to Take Cloud Native toward the Edge

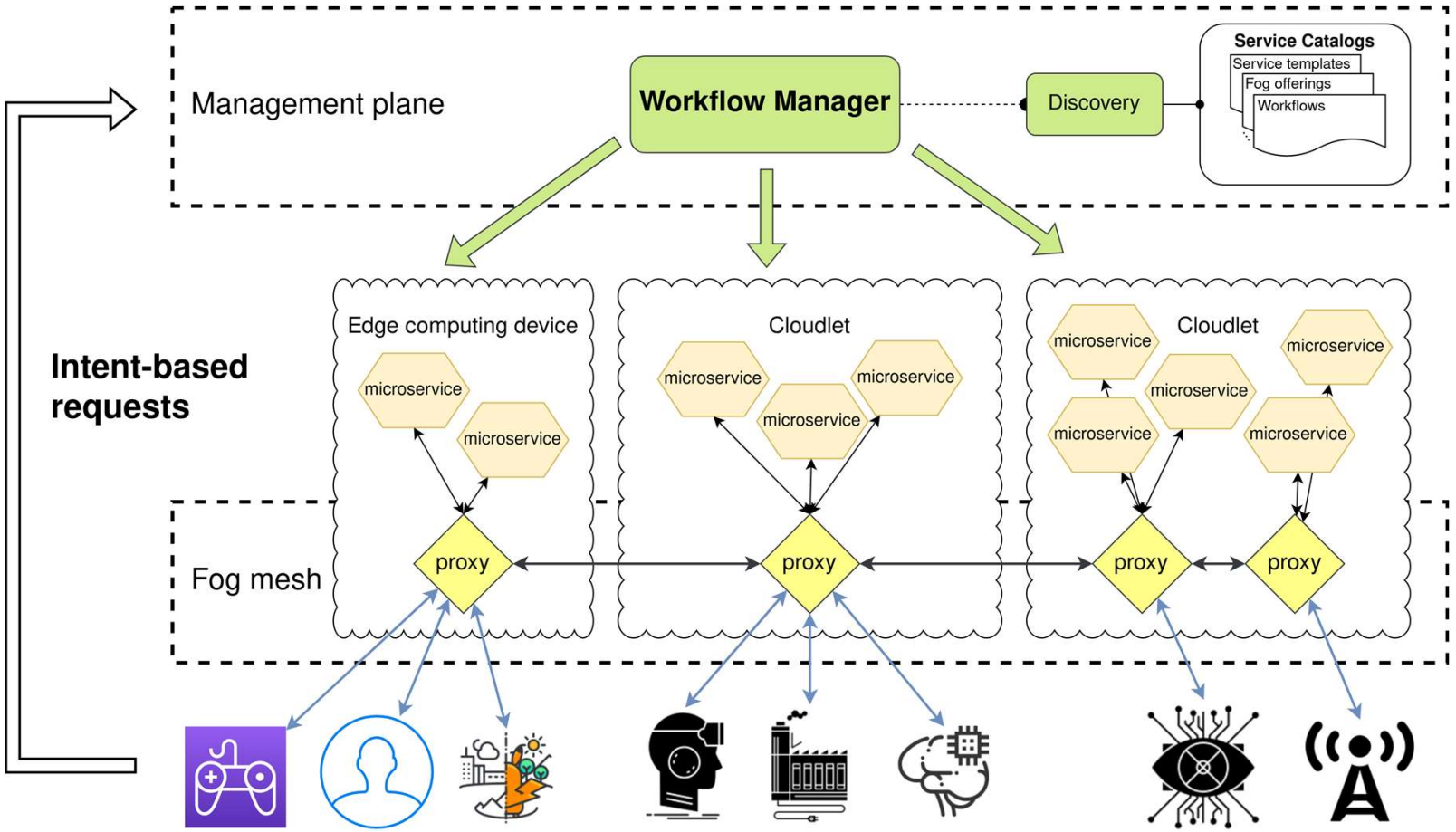
Cloud

Fog

Operational Complexity	Relatively homogeneous: seamlessly hides the specifics of the underlying infrastructure.	Inherently heterogeneous.
Resource constraints	Overabundance of relatively cheap resources.	Resource constraints become more apparent closer to the edge of the fog.
Distributed Nature	Typically offers low-latency and high-bandwidth communications.	Fog nodes are fully distributed with highly variant latency and bandwidth connections.
Geographical Complexity	The cloud has a relatively limited number of geographical locations.	Dispersed with vast geographical locations offered by edge tiers.

Solution: Microservice applications transformed into microservice workflows, constructed dynamically using intent-based matching of user requirements (e.g., location, specific user demands)

Designing a Fog Native Architecture

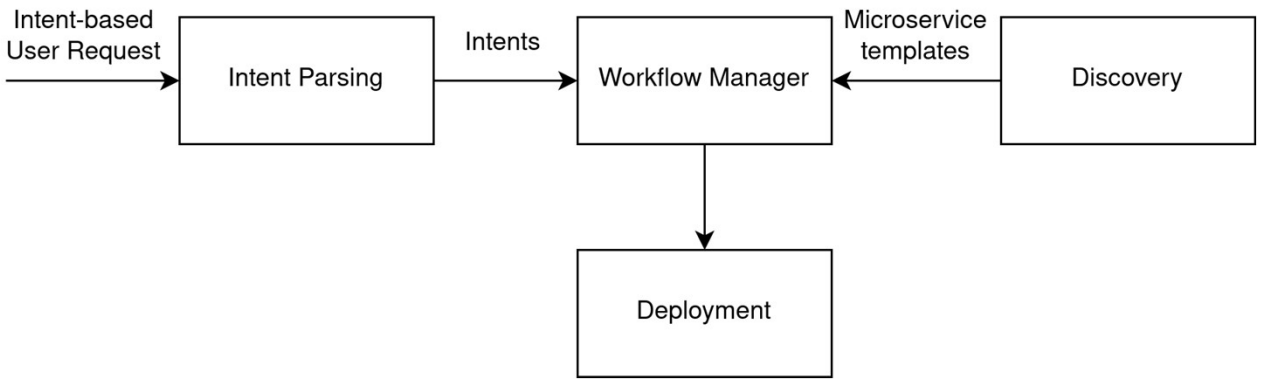


Intent-based Workflow construction

Discovery Service — Registers microservice templates enriched with metadata about functionality, characteristics, and dependencies.

Workflow Manager — Responds to user requirements by either identifying a running workflow or creating a new workflow from scratch to satisfy the user request.

Fog Mesh — Enables the communication between microservices and end users. Consists of a set of interlinked service proxies, providing network services to regional clusters.

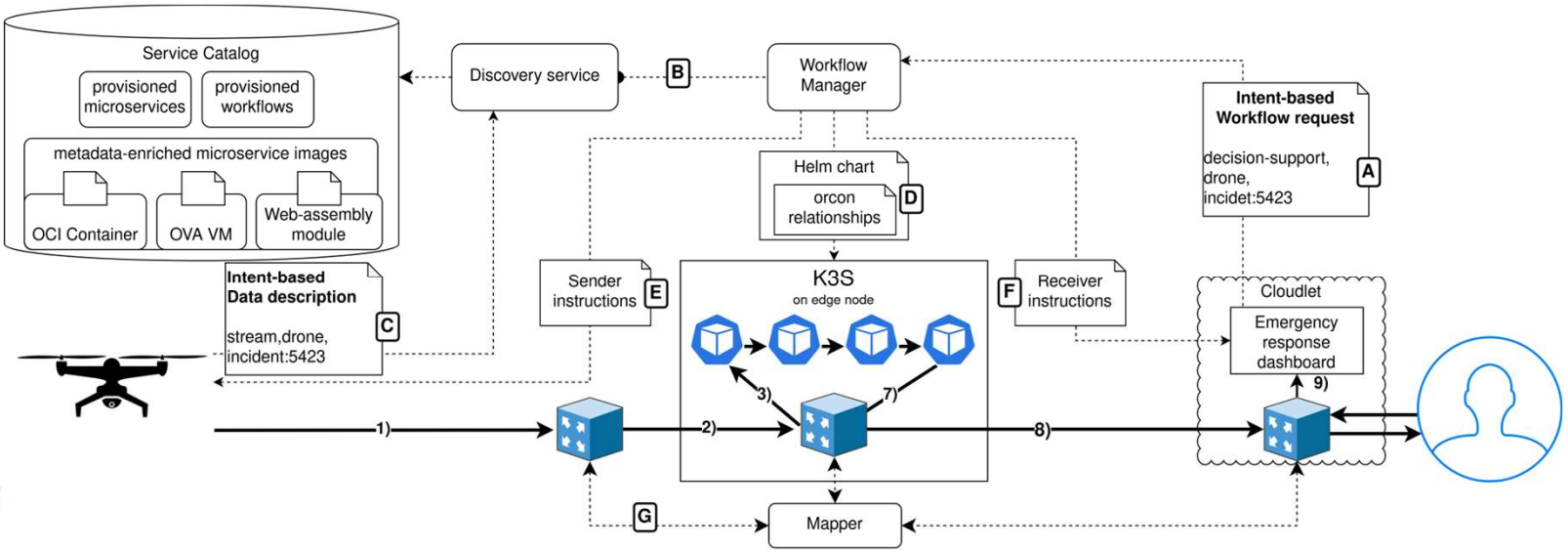


The intent-based workflow permits larger flexibility than the traditional desired-state approach since the orchestrator can change application topology and interchange application components depending on the user demand and available resources at a certain location.

Use Case: Decision-Support in the Fog

A dynamic decision support system based on live drone feeds for aiding first responders.

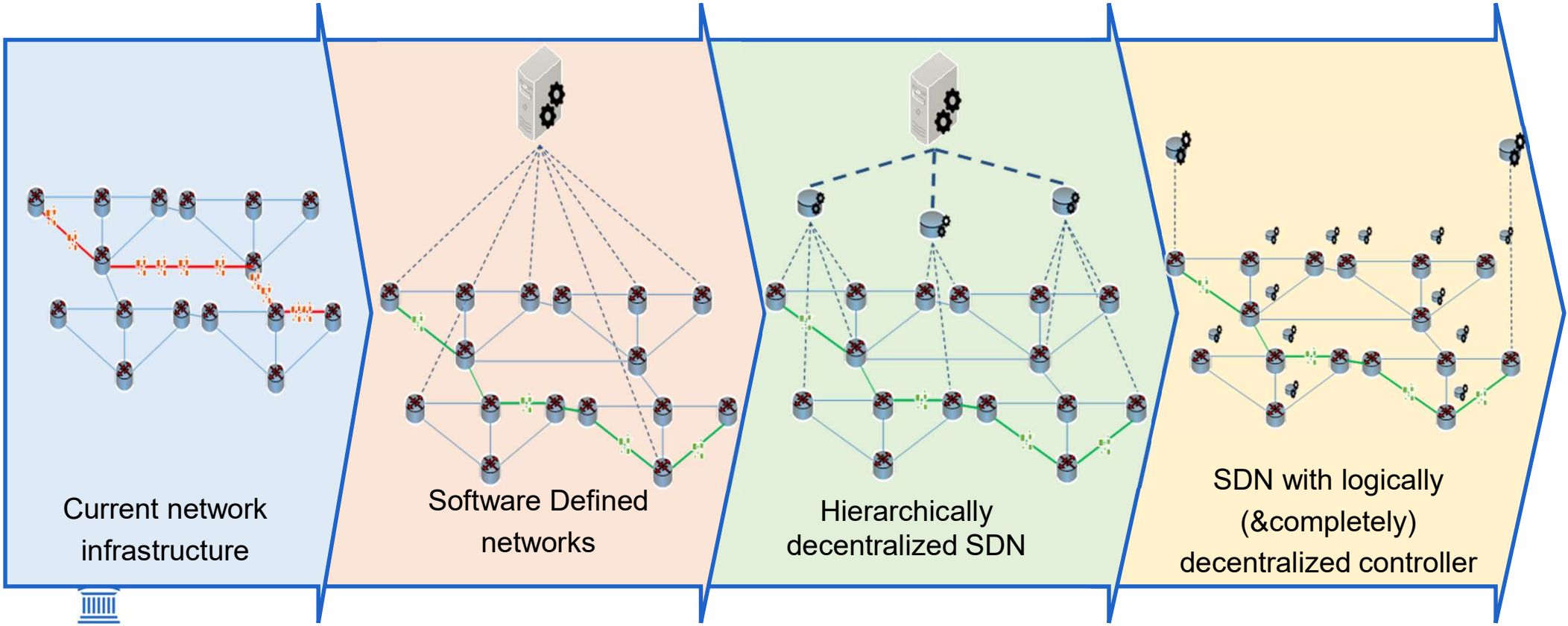
Due to the high bandwidth and low latency requirements of the decision support pipeline, running them on a centralized data center will negatively affect both the end-user experience and strain the network resources.
Thus, the system needs to dynamically design, and schedule local workflows based on responders' needs (i.e., intents).



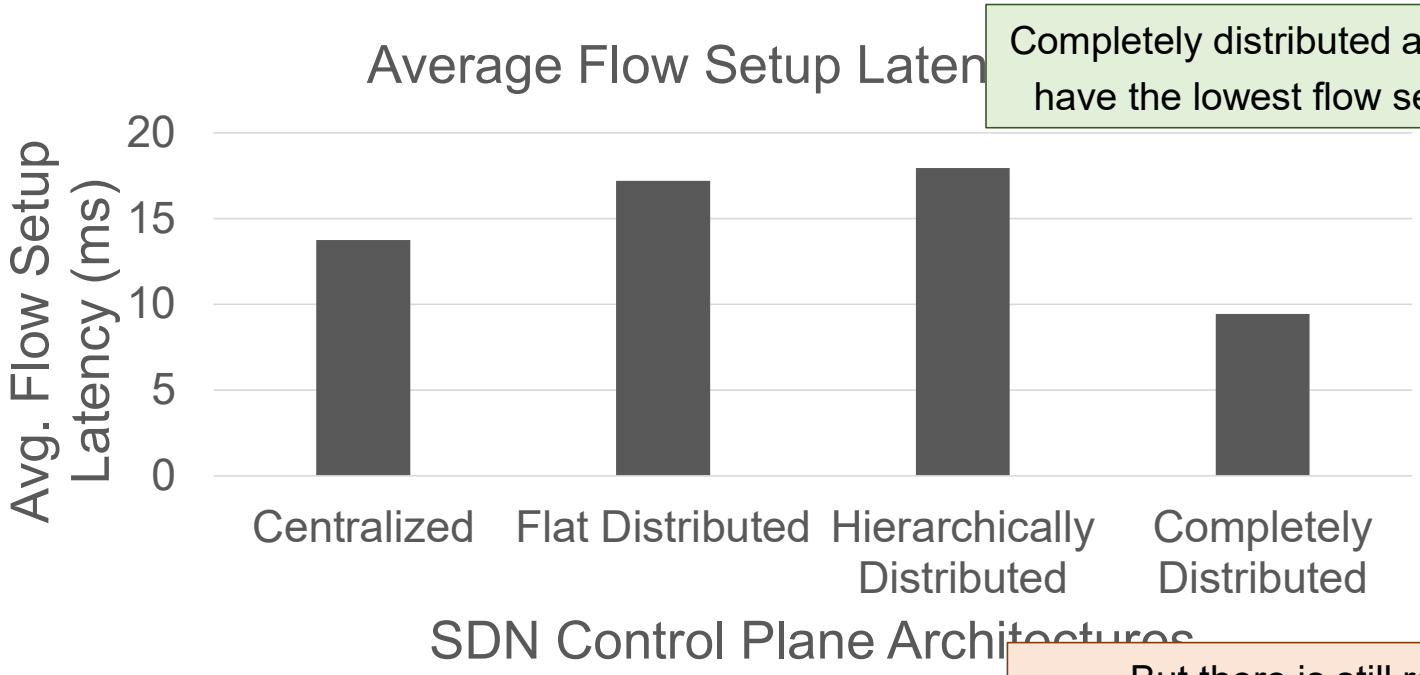
GHENT UNIVERSITY

imec

TOWARDS FULLY DECENTRALIZED NETWORKS…



FLOW SET-UP LATENCY ANALYSIS FOR DIFFERENT SDN ARCHITECTURES [1]



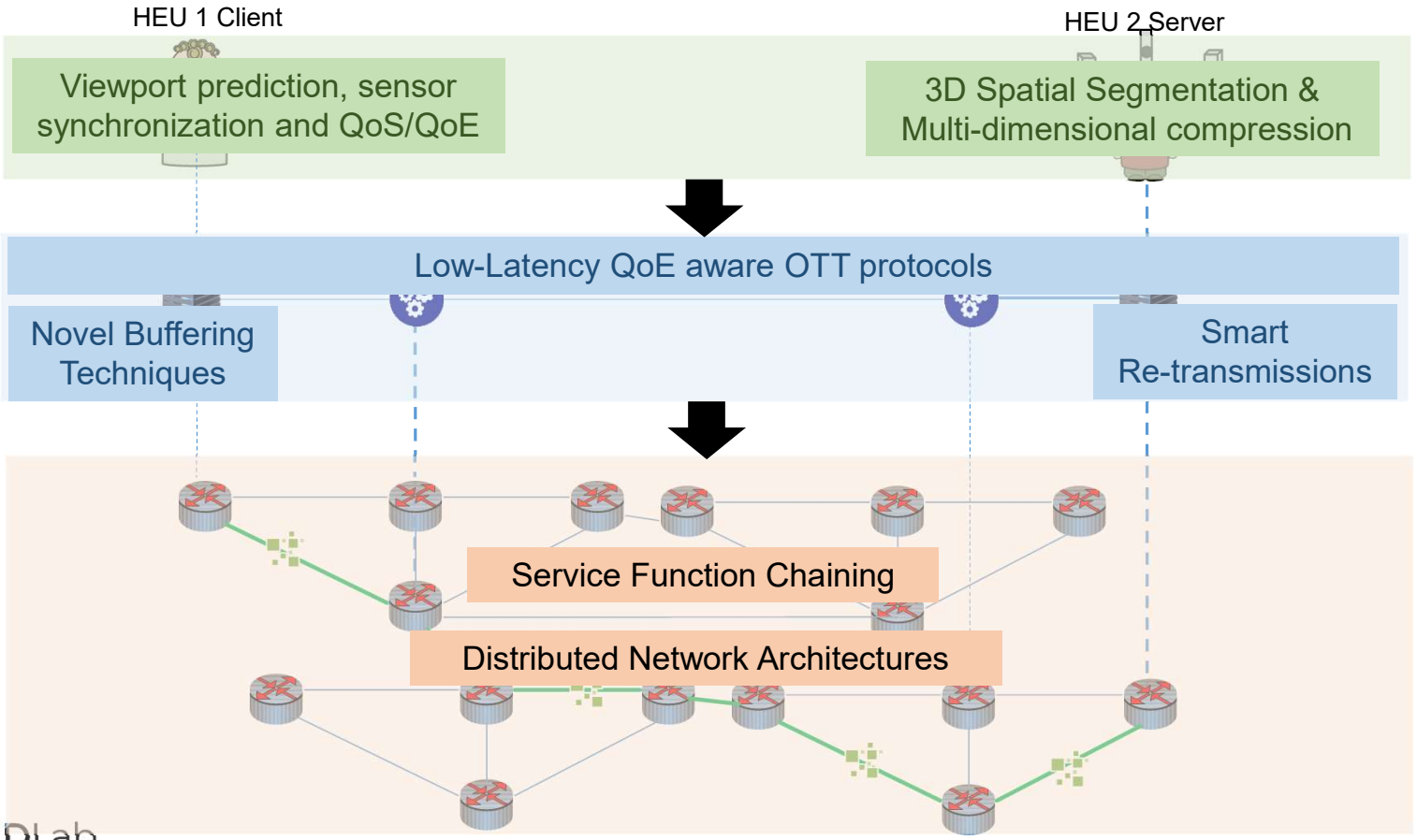
Completely distributed architectures have the lowest flow setup delay!

But there is still room for improvement

[1] H. K. Ravuri, M. Torres Vega, J. van der Hooft, T. Wauters, B. Da and F. De Turck, "On Architectures," 2020 11th International Conference on Network of the Future (NoF), 2020, pp. 23-27

[2] H. K. Ravuri, M. Torres Vega, J. van der Hooft, T. Wauters, and F. De Turck, "A scalable hierarchically distributed architecture for nextgeneration applications," Journal of Network and Systems Management, vol. 30, no. 1, pp. 1–32, 2022.

TRULY IMMERSIVE HTC: A CROSS-LAYER APPROACH



CONCLUSIONS AND CHALLENGES

OPEN CHALLENGES AND TOPICS

Network optimizations (to increase QoE by adapting quality / reducing latency)

- Protocols (DASH, WebRTC, QUIC) to be able to leverage the quality vs latency tradeoff.
- Softwarized networks to reduce latency: SDN/NFV, distributed flow setup
- Computational offloading to reduce latency and increase quality (Cloud/edge/fog, Mobile Edge Computing (MEC), Service Function Chaining (SFC))
- Transport enablers: Segment routing, packet scheduling / caching / prioritization / retransmission / dropping

Predictive orchestration and distributed networking

- For low-latency applications in future 6G Networks
- Intent-based networking
- Service Function Chaining
- Micro-Service allocation and resource management

Application optimizations (to increase QoE) (end-user and server side)

- Encoding strategies: Tiling, point clouds, light fields, dynamic meshes
- QoE modeling, haptics
- View prediction and content prefetch
- Adaptive Bit Rate (ABR) selection focused on what the user is looking at

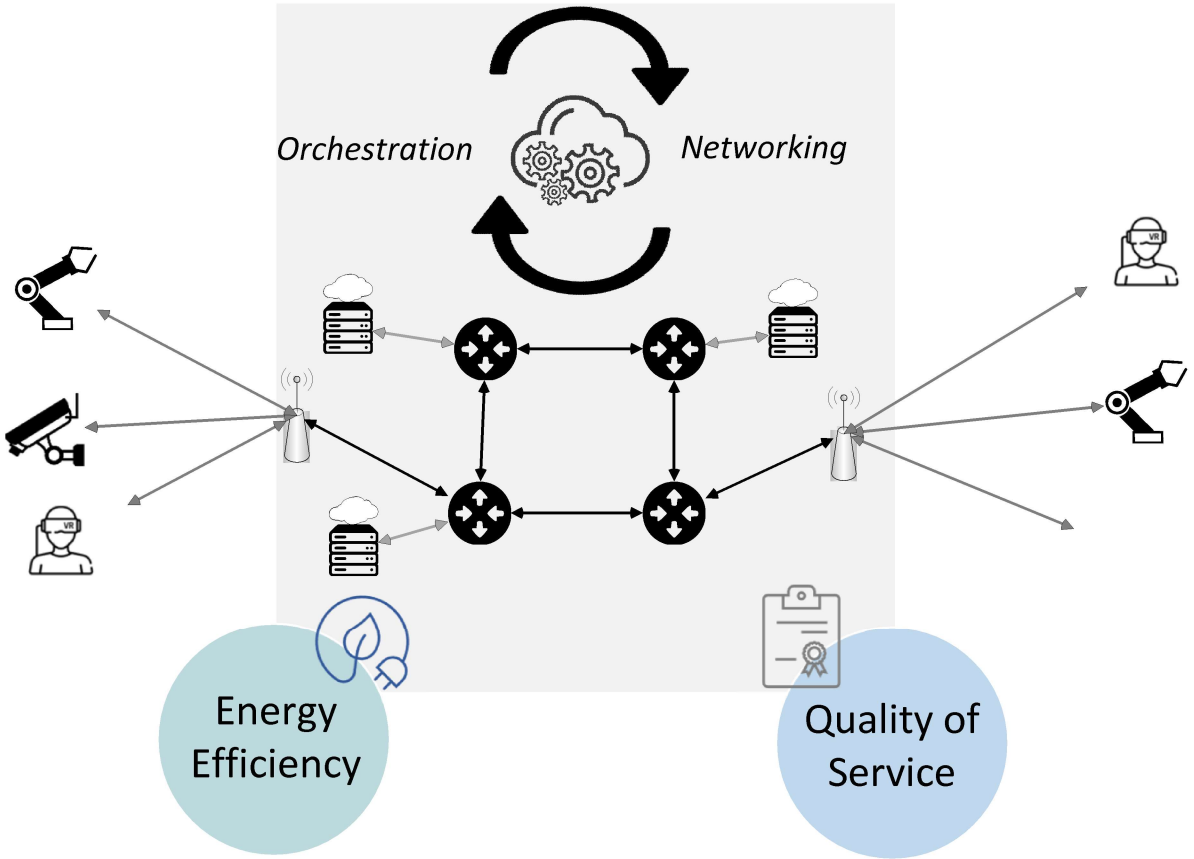
Predictive orchestration and distributed networking for low-latency applications in future 6G Networks

Reinforcement Learning (RL)

Service Function Chaining

Intent-based Networking

Segment Routing (SR)



Bandwidth

Latency

GHENT UNIVERSITY

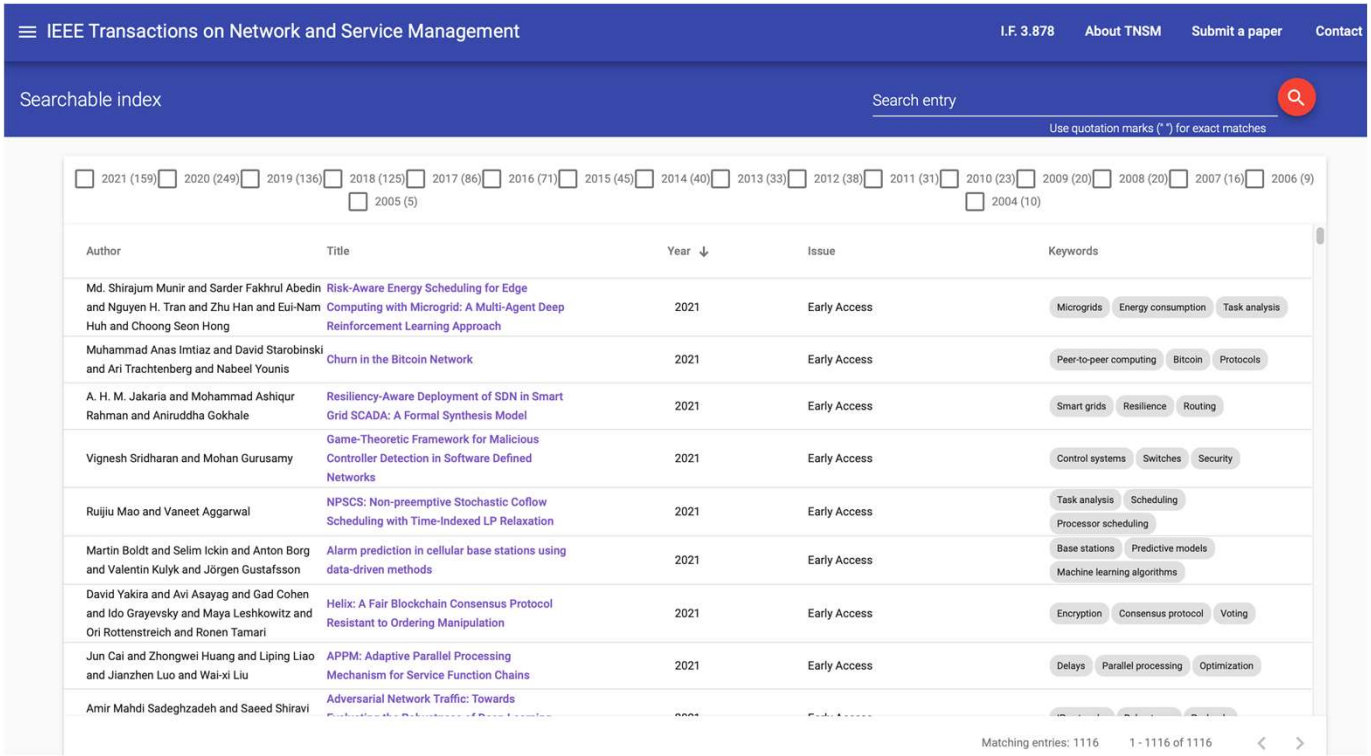
IDLab
INTERNET & DATA LAB

imec

IEEE TRANSACTIONS ON
NETWORK AND
SERVICE MANAGEMENT

IEEE Transactions on Network and Service Management (TNSM)

- Submission format: 12 pages + extra pages
- Regular call + special issues + extended versions of best papers selected from conferences
- Current issue: Volume 19, Issue 4
- JCR Impact factor: 4.758
- H-index: 57 (as of 12/22)



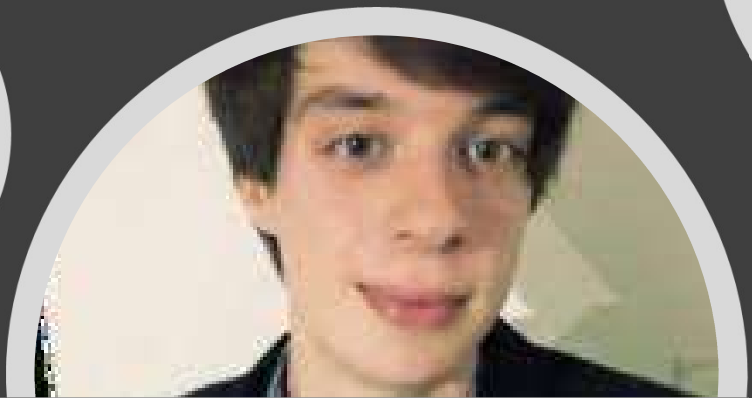
TNSM Overview – Analytics

http://www.tnsm-overview.org

Conferences on Network and Service Management

Thanks to the team

- Maria Torres Vega
- Tim Wauters
- José Santos
- Hemanth Kumar Ravuri
- Sam Van Damme
- Jeroen van der Hooft



"***No, you can't wipe `em off. They're holograms***." - Tobias Becket to Chewbacca in Solo (2018)
"***Holograms are the next video***" – Philip A. Chou

Thank you for your attention!

Any questions or comments?