Qualcomm

EURASIP Summer School on Metaverse Technologies

September 20, 2023

XR and the Metaverse: Technologies and Standards

Thomas Stockhammer

Senior Director Technical Standards, QSIO

Qualcomm Technologies, Inc.

@QCOMResearch

Qualcomm

Today's agenda

Metaverse – Introduction

Qualcomm's contributions to the Metaverse

Technologies for the Metaverse

Standards for the Metaverse

Invitation to collaborate

Joint work with **Dr. Imed Bouazizi,**
Director, Technical Standards, Qualcomm
Delegate to 3GPP SA4, MPEG, and Khronos
and many other people in Qualcomm and standards

Presenter

Dr. Thomas Stockhammer
Senior Director, Technical Standards
Qualcomm Europe, Inc., IEEE Fellow

Leading and driving among others

- DVB: 5G TF, DVB-I
- MPEG: MPEG-I, CMAF and DASH
- 3GPP: XR over 5G, 5G Video, 5GMS
- DASH-IF: Interop WG, Test
- ETSI & 5G-MAG: 5G Broadcast and 5GMS
- CTA WAVE: CMAF Device PB, Test
- Metaverse Standards Forum: Chair, Board

Introduction to the Metaverse

Metaverse

Metaverse **is a set of virtual spaces, where you can create and explore with other people in the same or different physical spaces**

Exercise / Games

Education

Shopping

Communication

Business

Courtesy: Joel Kim, Head of Cellular Standards and Ecosystems at Meta
Surfing Metaverse with 5G and Beyond, ICC 2022

Consumer

- Fitness
- Gaming
- Social
- Sports / Concerts

- Entertainment
- Navigation
- Shopping

VR

is here today - and will continue to grow

AR

is next - and will disrupt personal computing

- Corporate Training
- Education
- Medical

- Infinite Desktop
- Instructions
- Remote Assistance

Enterprise

Device Requirements

- **Ergonomics:**
  - **Wearable devices should be light enough, and not hamper movement of the users**
- **Power consumption:**
  - **Battery-friendly media and data processing to ensure high service quality and sufficient use time per charge**
- **Heat dissipation:**
  - **Heat from processors and displays worn around the face needs to be avoided or properly insulated**

Network Requirements

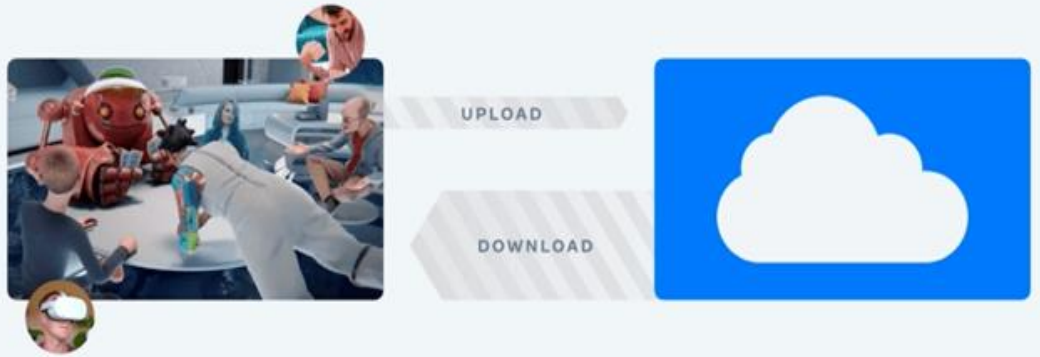
Consistent quality of experience

Ensure network availability can enable immersive experiences

Reduce latency

Reduce the time it takes for information to travel from point A to point B

Symmetric bandwidth



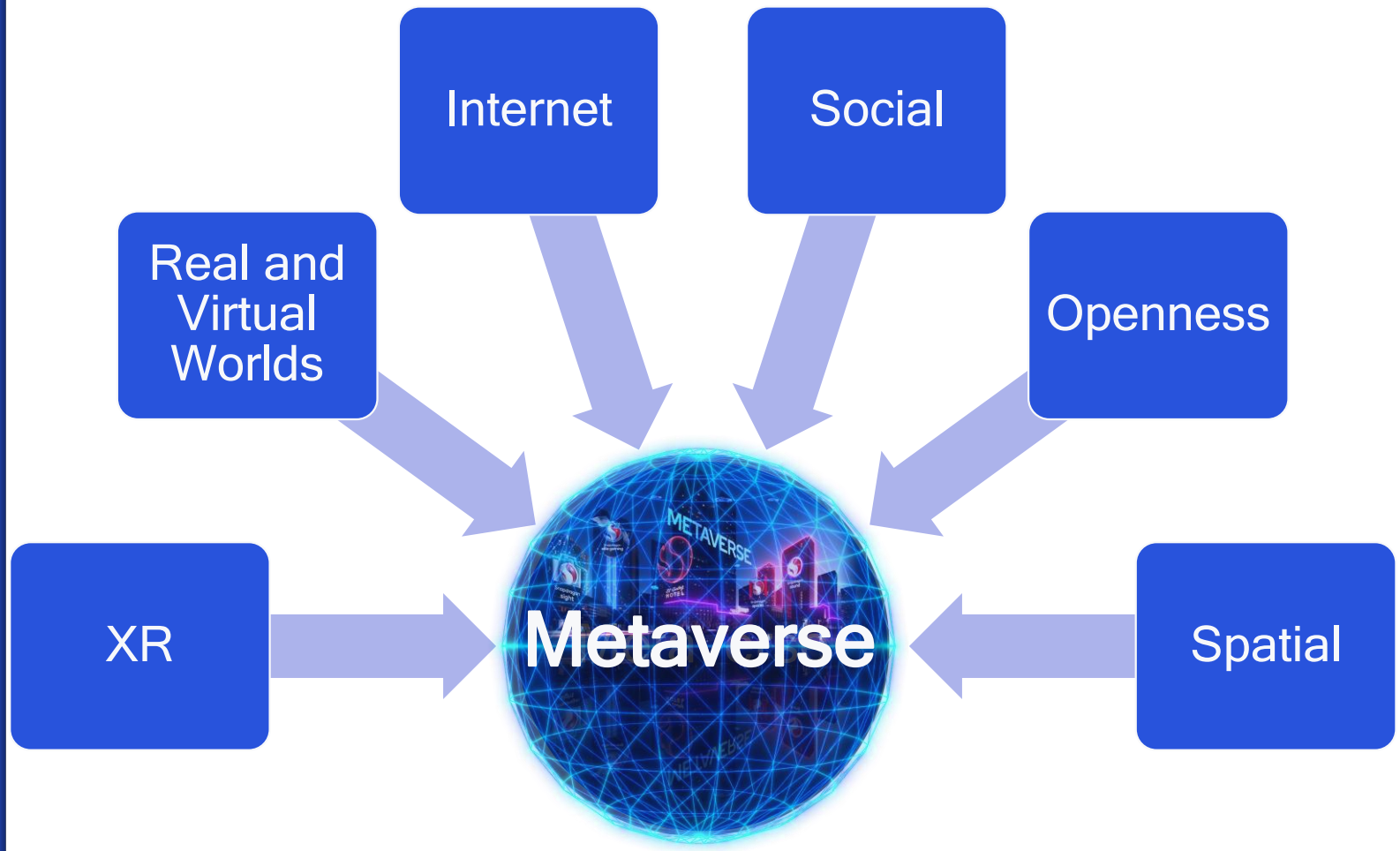
Ensure users can send as much information as they can receive

Metaverse

Persistent spatial internet with personalized digital experiences

Spans both physical and virtual worlds

Shared virtual space in VR today, evolving to digitally enhanced physical space with AR & MR



Snapdragon

Your ticket to the metaverse

Snapdragon Spaces is a product of Qualcomm Technologies, Inc. and/or its subsidiaries.

Qualcomm Technologies are enabling the Metaverse

Platforms

Snapdragon
XR2+ platform
Gen 1

Snapdragon
XR2 platform
Gen 1

Snapdragon
XR1 platform
Gen 1

Software & algorithms

6DoF/SLAM

Object Rec & Tracking

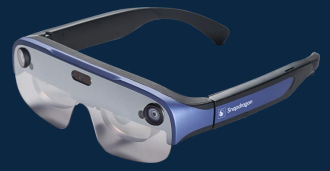
Eye Tracking

Pass Through (video see through)

Hand Tracking

3D Reconstruction

Reference designs



Ecosystem

Snapdragon spaces

Qualcomm Technologies is enabling the XR industry

Accelerating VR

Pico Neo 4

Meta Quest Pro

Accelerating AR

Tethered Smart Viewer Reference Design with Snapdragon XR1

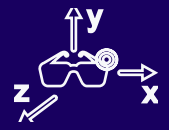
Wireless Smart Viewer Reference Design with Snapdragon XR2

Snapdragon
spaces

Snapdragon
spaces

Spatial mapping & Meshing

Scene understanding



Positional tracking

Object recognition & Tracking

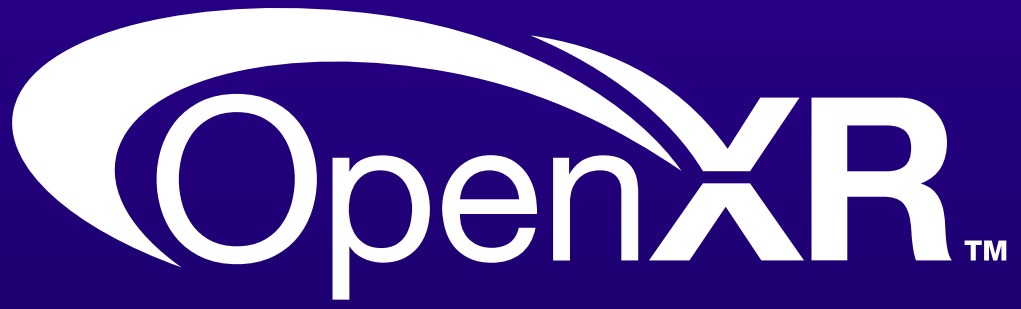
Local anchors & Persistence

Snapdragon spaces

Hand tracking

Image recognition & Tracking

Plane detection


OpenXR™

Snapdragon spaces

OpenXR™

EPIC GAMES

Unity®


Snapdragon
spaces

Snapdragon spaces

AR

VR and MR

Snapdragon spaces

Snapdragon branded products are products of Qualcomm Technologies, Inc. and/or its subsidiaries.

TRIPP®

Snapdragon spaces

Digital Accelerator · arvizio · Beem · CAREAR™ A Xerox Company · ENGAGE REAL BUSINESS DONE VIRTUALLY · FUNDAMENTALVR · glue. · Hevolus | Innovation · HOLO-LIGHT · holo|one · OSSO VR · PIXO · room · Roundtable LEARNING · SCOPE AR · STRIVR · TALESPIN · TeamViewer · TECHVIZ · transfr · uptale · VICTORYXR · VIEWAR™ · XRHealth

AR VR MR

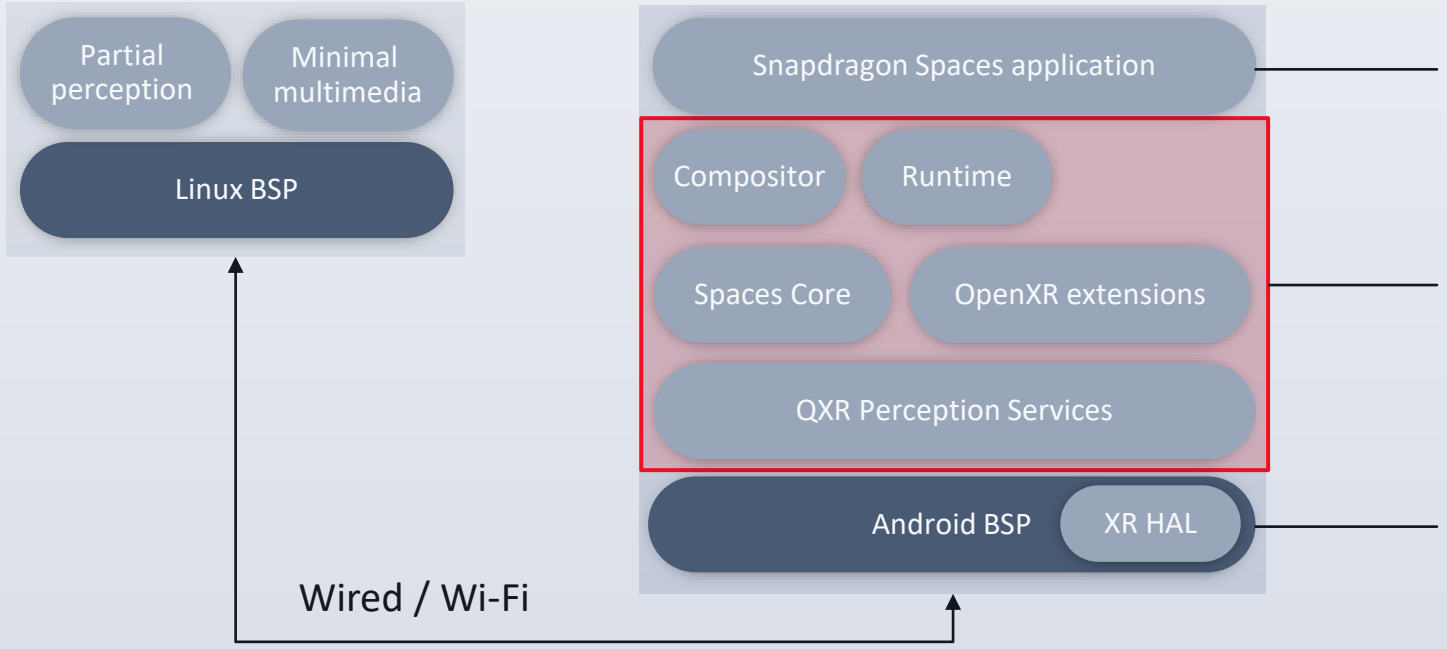
Adobe

Microsoft

NIANTIC

SNAPDRAGON SPACES READY

motorola · nubia · ONEPLUS · REDMAGIC · SHARP · mi · ZTE



Snapdragon Spaces application

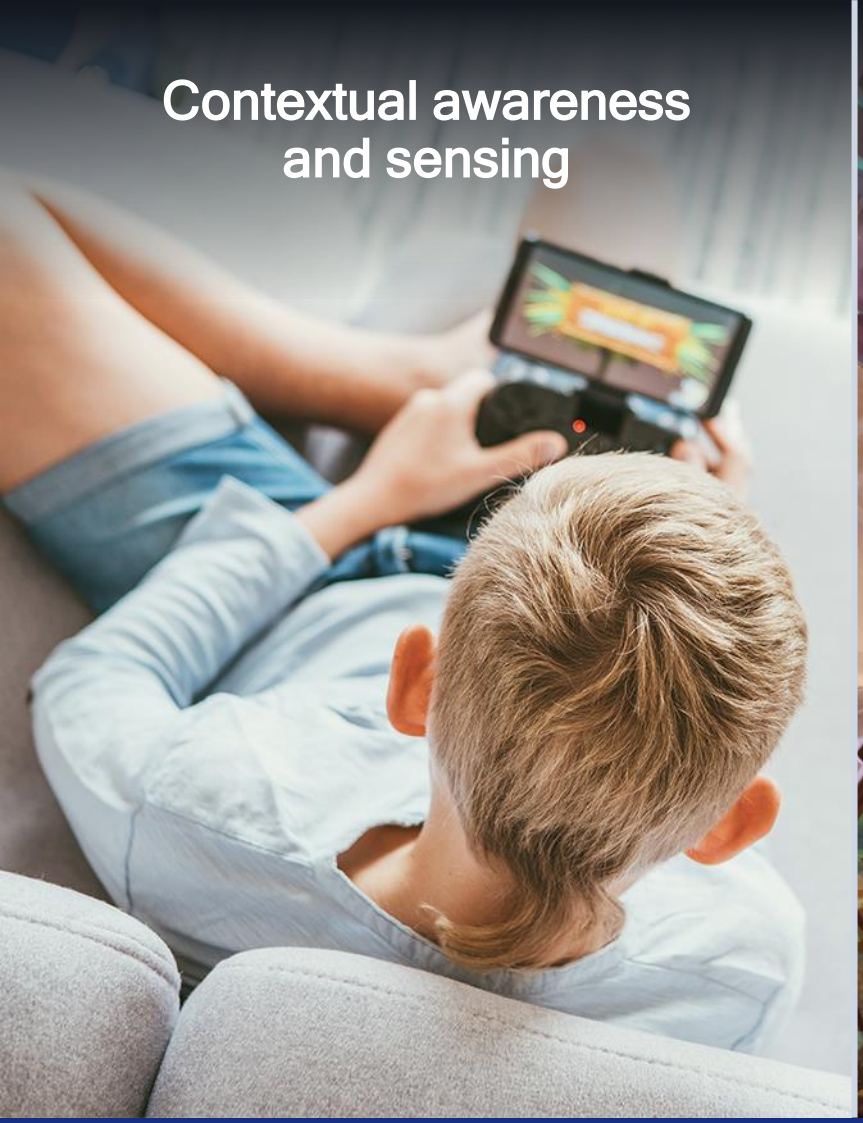
Published by developers

Snapdragon Spaces Services APK

Downloadable through Google Play Store

Snapdragon Spaces Ready HAL

Delivered as part of smartphone baseline
SP (software product)



Technologies for the Metaverse: XR, AI, 5G

AR/VR/AI Work Flow

Simultaneous Complex Processing for Seamless AR/VR Experience

3 **Recognize and understand images** for symbolic information

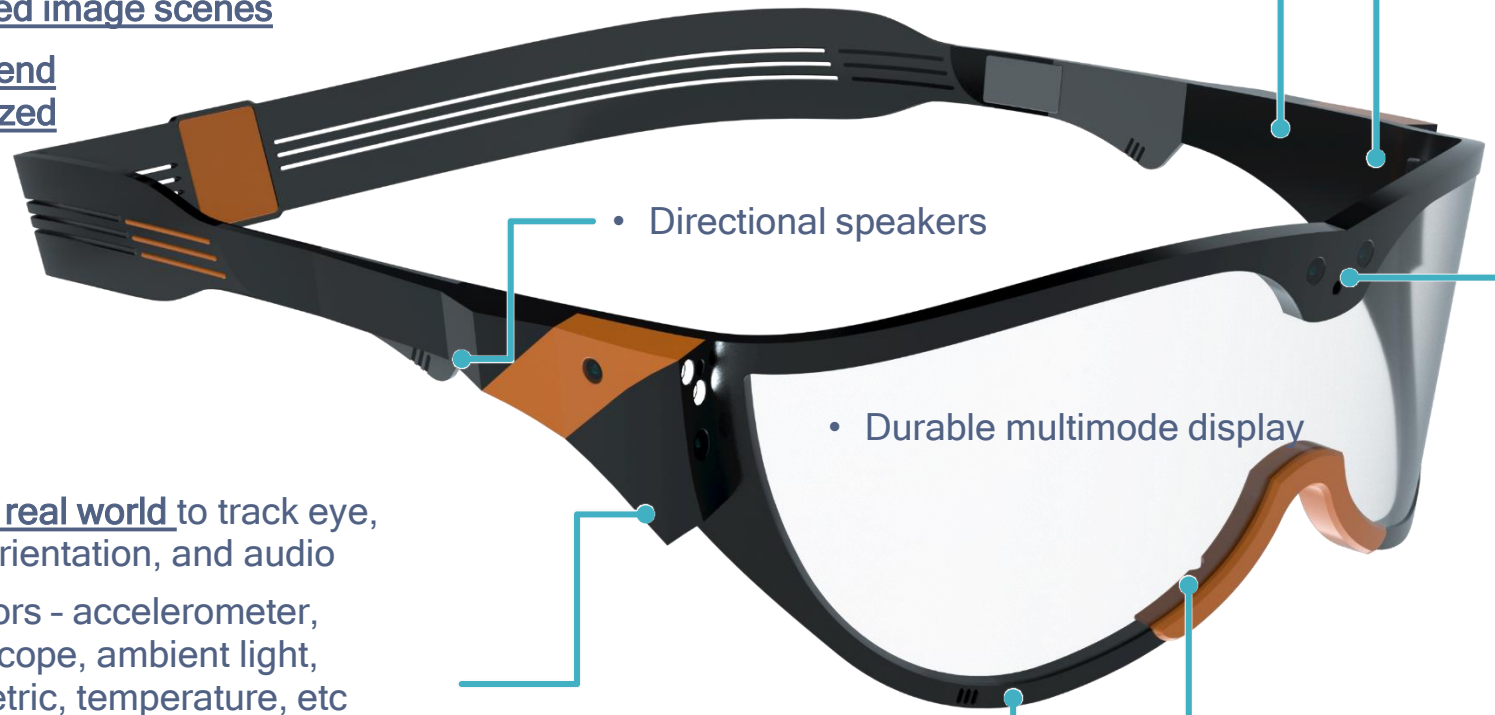
- Highly efficient and lower power CV processing

5 **Overlay virtual objects/data with real or virtually created image scenes**

6 **Recommend personalized actions**

4 **Collect virtual contents and data** from cloud network

- Multimode connectivity (4G, 5G, Wifi)
- GPS for location information
- Coding compression
- Video transport



- Directional speakers
- Durable multimode display

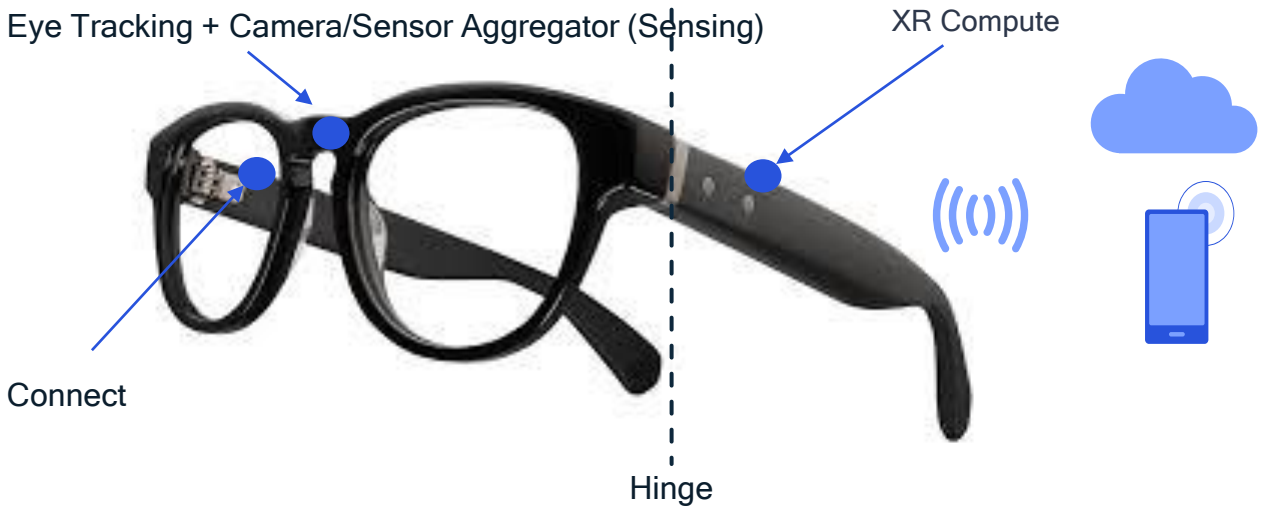
1 **Sense the real world** to track eye, position, orientation, and audio

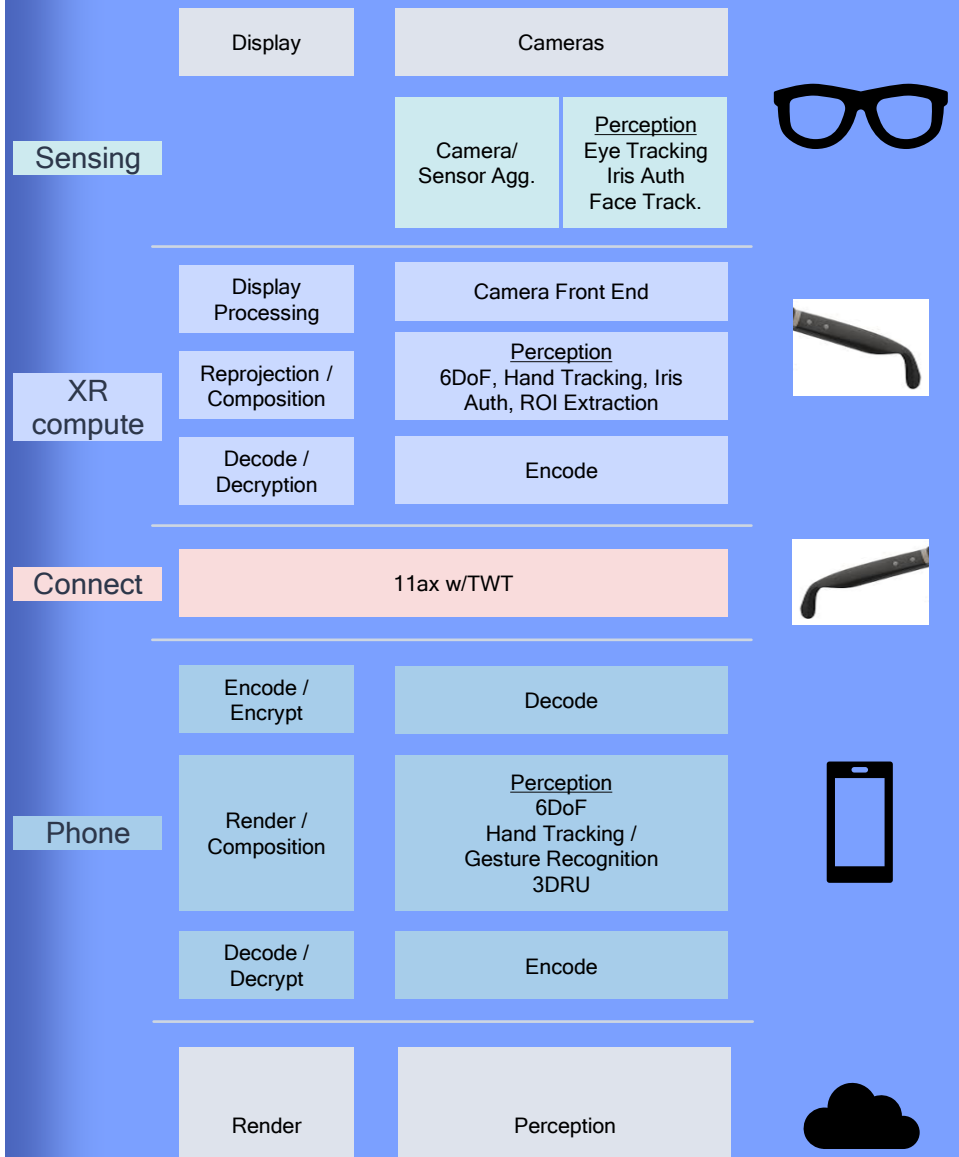
- Sensors – accelerometer, gyroscope, ambient light, biometric, temperature, etc
- Multiple audio microphones
- Eye tracking cameras

2 **See the world around** you by collecting and monitoring visual images for depth info, object detection, face detection

- Cameras with fisheye and telephoto lenses
- Depth sensing camera (TOF, Structured Light, etc)
- Optoelectronic night vision and thermal imaging sensors

AR System Architecture / Partitioning





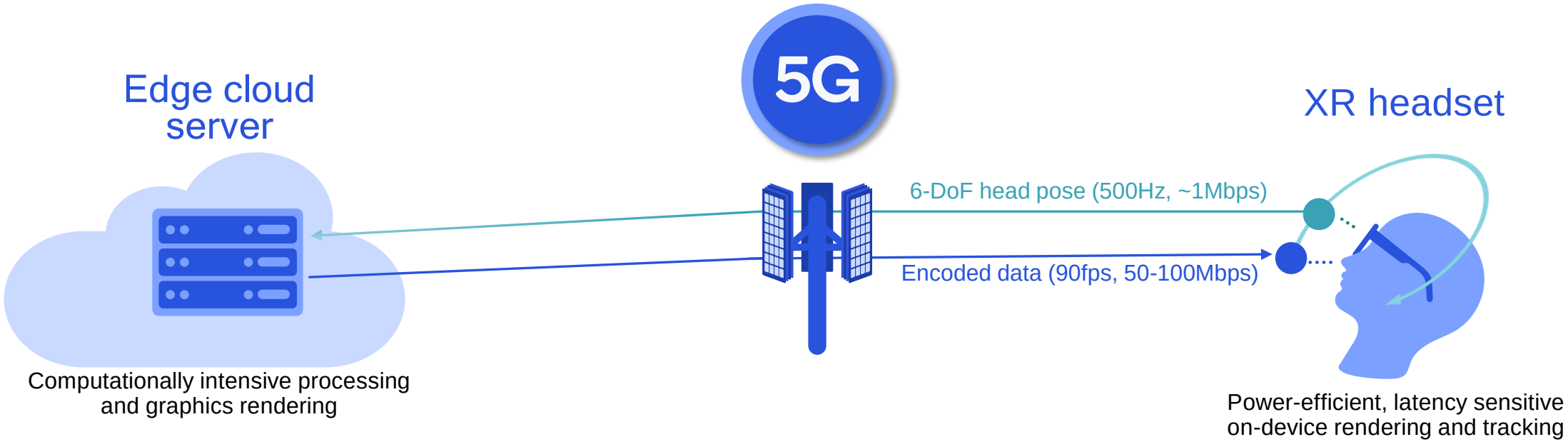
Key Goals/Constraints for Chipset

- Minimize power per thermal island – divide workloads to multiple chips in headset
- Minimize overall power for battery life
- Minimize wire across hinges
- Partition workloads to remote devices / cloud to balance power loading
- Satisfy end-to-end latency requirements
- Conform to physical size constraints

A distributed compute architecture enables rich XR user experience

M2R2P = Edge processing + 5G round-trip time + Device processing

Boundless XR over 5G



Distribute computation between the edge cloud server and device

Leverage low-latency and high-capacity 5G

Motion-to-render-to-photon (M2R2P) animation video

Boundless XR over 5G

Latency[1]
M2R2P latency < 70ms
5G RTT < 20ms

Frame rate
2kx2k per eye
at 90 frames per second

Throughput[1]
Reliable average downlink throughput of 50-100 Mbps
Reliable uplink throughput of ~1 Mbps, 500 Hz pose

Immersive VR
Photorealistic visuals
6-DoF mobility
Robust and reliable

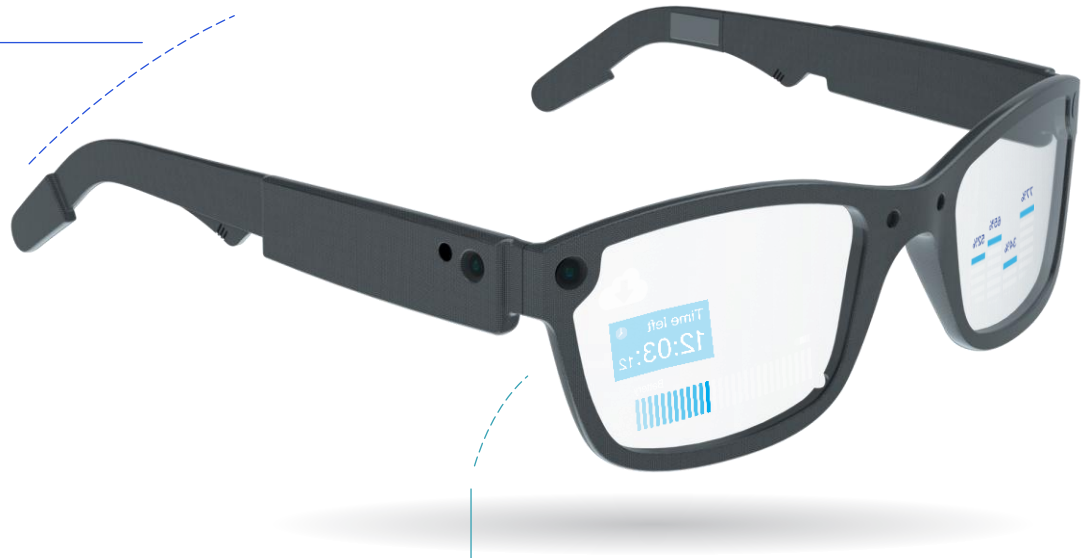
Achieved initial KPIs for at-scale 5G boundless VR deployments

Boundless AR

Cloud server

Smartphone

AR glasses



Latency
M2R2P latency of 50-100ms
5G+Wi-Fi RTT < 30ms[1]

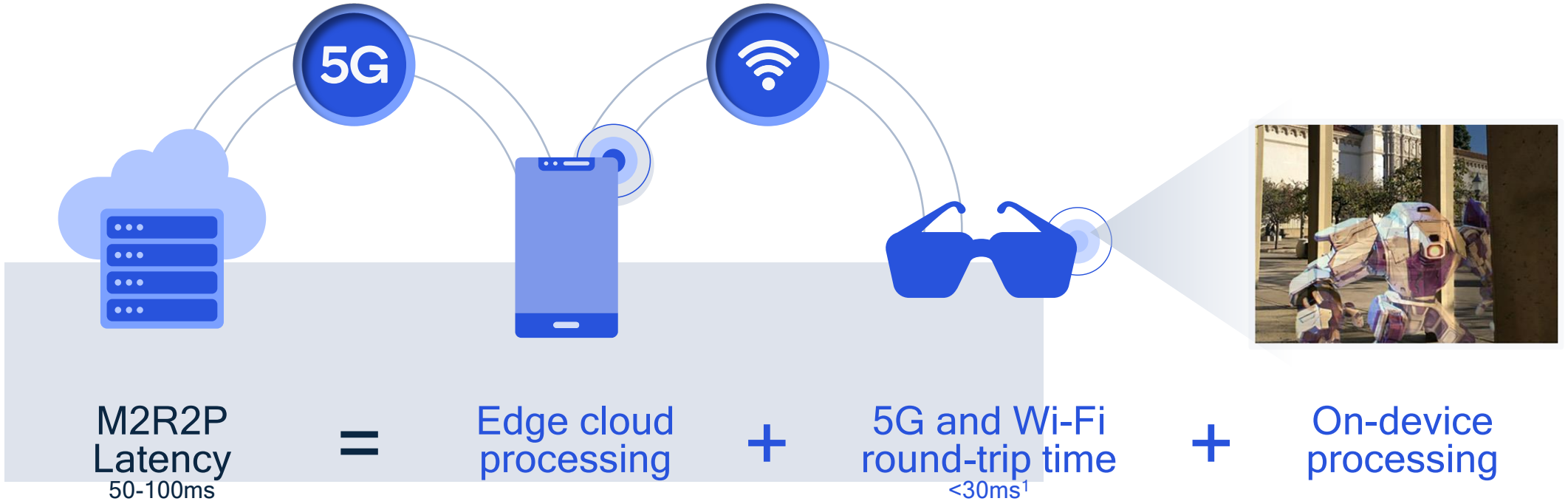
Frame rate
2kx2k per eye
at 90 frames per second

Throughput
Reliable average downlink throughput of 10-40 Mbps
Reliable uplink throughput of 3-20 Mbps

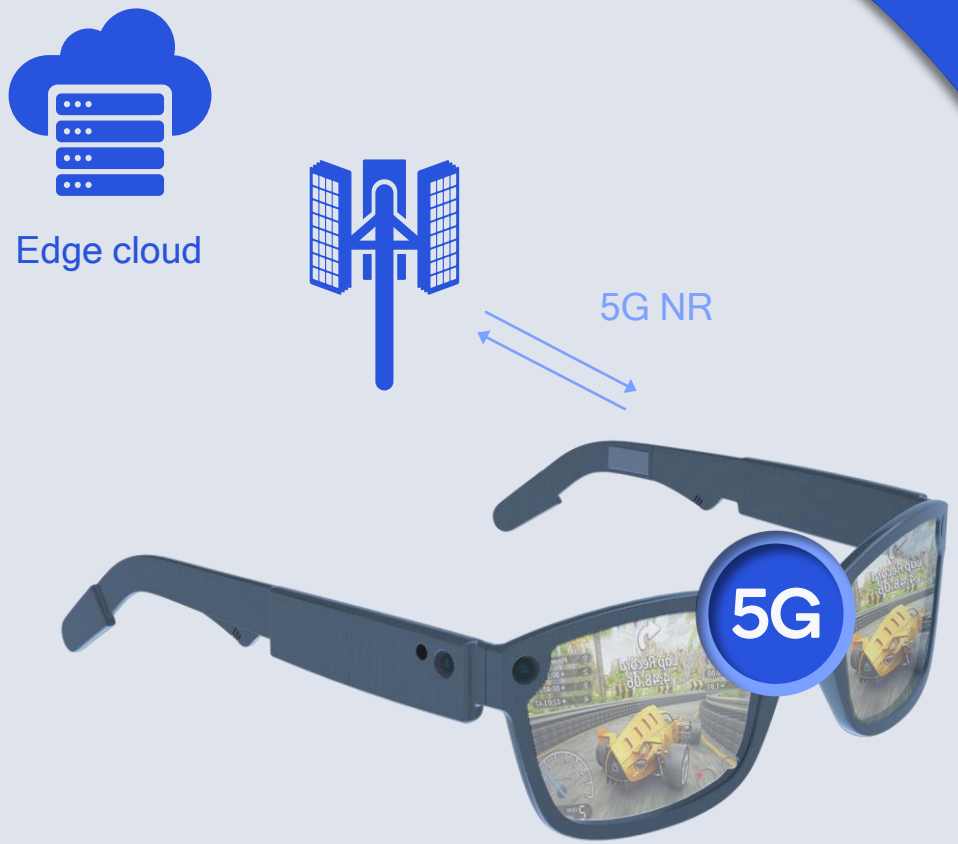
Immersive AR
Photorealistic visuals
6-DoF mobility
Robust and reliable

KPIs for cloud-to-phone-to-glass deployments

Boundless AR over 5G and Wi-Fi



Distribute the computation
between the edge cloud server, phone, and AR glasses



Enabling 5G-powered AR glasses

Optimized edge processing

Migration from central cloud to local edge

Improved Infra schedulers[1]

Delay aware schedulers to meet latency QoS

Low-power, low-latency 5G

3GPP based features

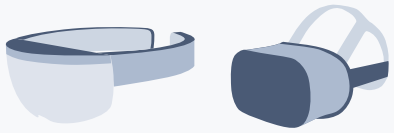
5G modem APIs

Enabling low latency on-device optimizations

Enabling applications to adapt to RF/network conditions

XR evolution

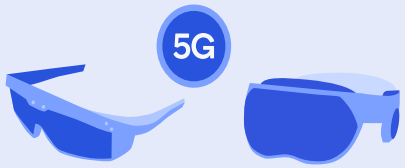
Standalone
VR and AR



Today

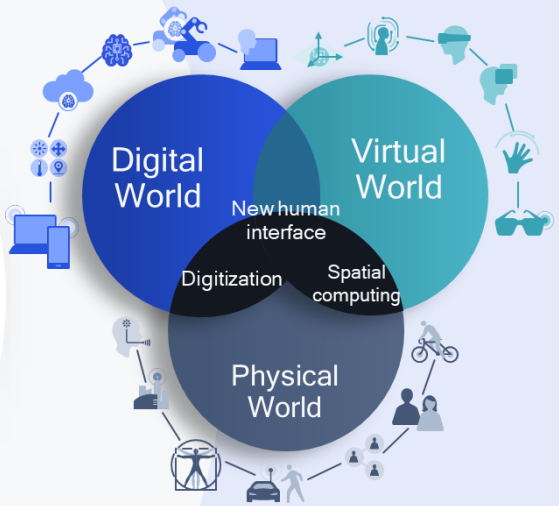
Viewer VR & AR
cabled

Standalone
VR and AR



1 - 4 years

Viewer VR & AR
wireless



Metaverse
6G research vector

6G will enable the merging of our worlds

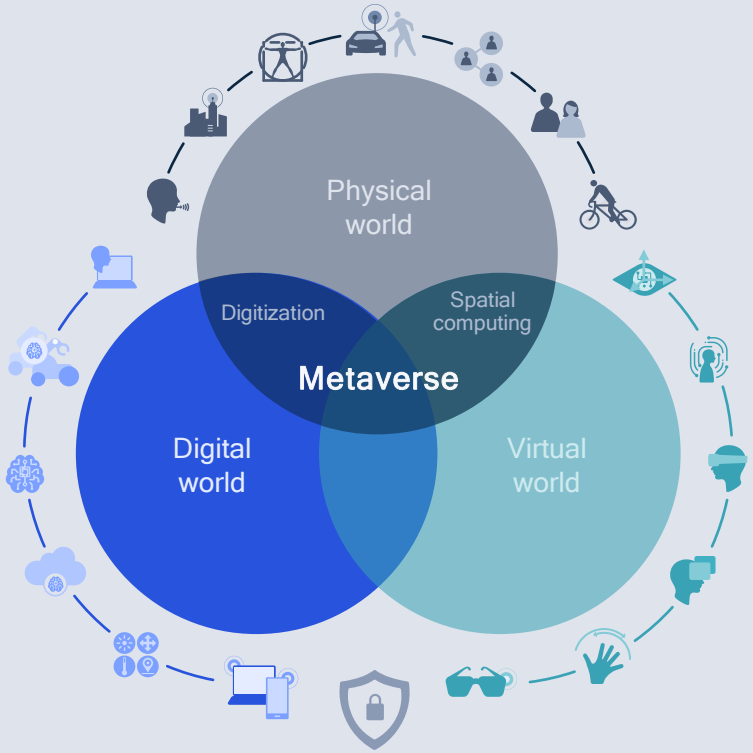
Physical, digital, virtual, immersive interactions taking human augmentation to next level via ubiquitous, low-power joint communication and sensing

AI-native E2E communications

Data-driven communication and network design, with joint training, model sharing and distributed inference across networks and devices

Air interface innovations

Waveform/coding from MHz to THz, advanced duplexing, Giga-MIMO, mmWave evolution, passive MIMO, satellite comm., system energy efficiency



Spectrum expansion and sharing

Expanding to THz, wide-area expansion to higher bands, new spectrum sharing paradigm, dynamic coordination with environmental awareness

Scalable network architecture

Disaggregation and virtualization at the Connected Intelligent Edge, use of advanced topologies to address growing demand

Communications resiliency

Multifaceted trust and configurable security post quantum security, robust networks tolerant to failures and attacks

Standards for the Metaverse

Open and Global Standards for the Metaverse

Metaverse STANDARDS FORUM™

3GPP A GLOBAL INITIATIVE

MPEG

KHRONOS GROUP

- XR Architectures
- XR Split Rendering
- Tethered AR Glass
- XR Conferencing
- IVAS Speech Codec
- XR Traffic QoS, Power Savings, Capacity Enhancement

- XR System: Scene Description
- Coding/Compression for CGC/3D content
- Haptics, Audio, Video
- Coding for Machines

Vulkan

OpenXR™

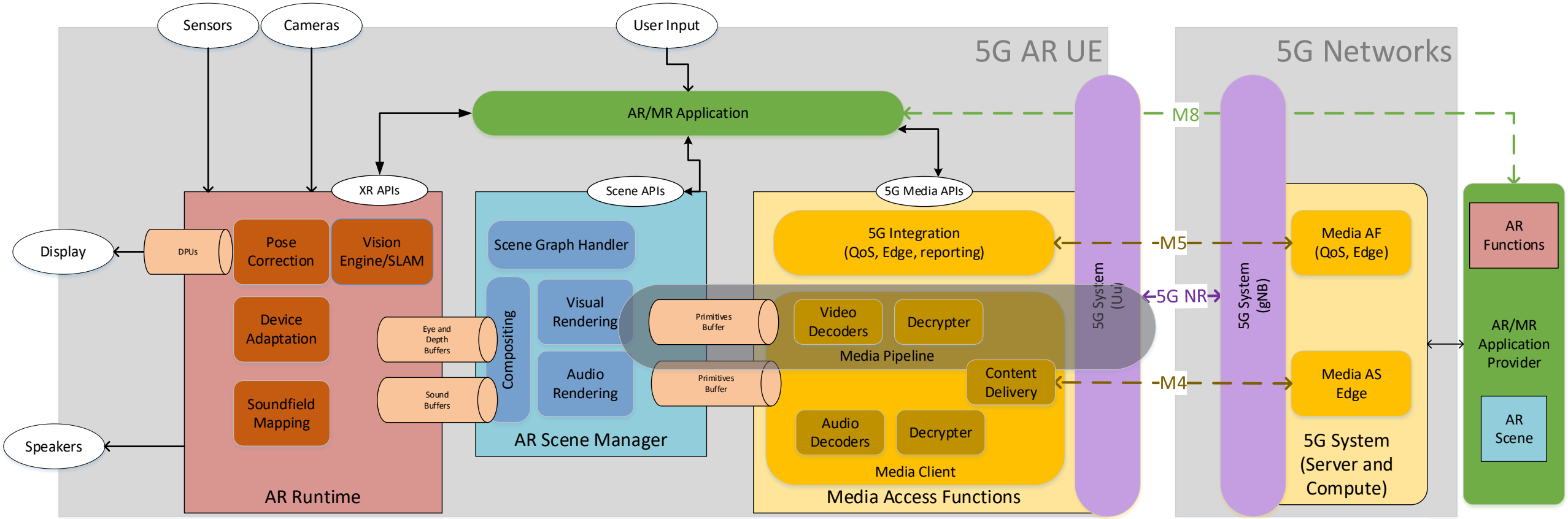
glTF™

QC driving or contributing

Additional selected organizations

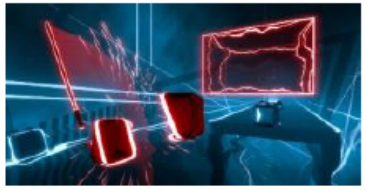
XR Standards – Optimizations, Systems and Workflows

Formalizing architectures, workflows and APIs for highest quality and lowest power consumption



OpenXR handles communication to and from an application and an XR platform

- Image(s) to display
- Audio
- Haptic responses



Application

OpenXR™

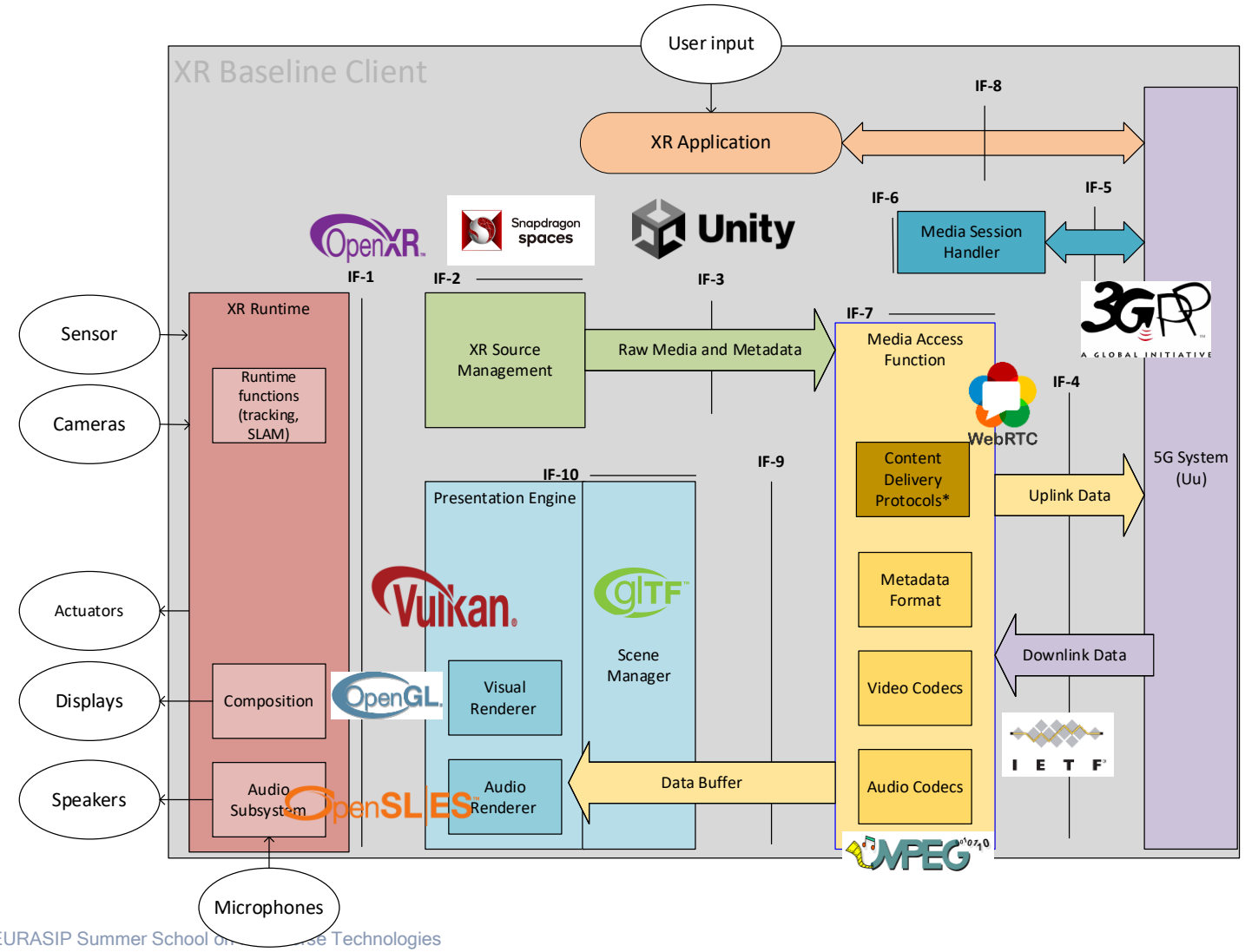
XR device

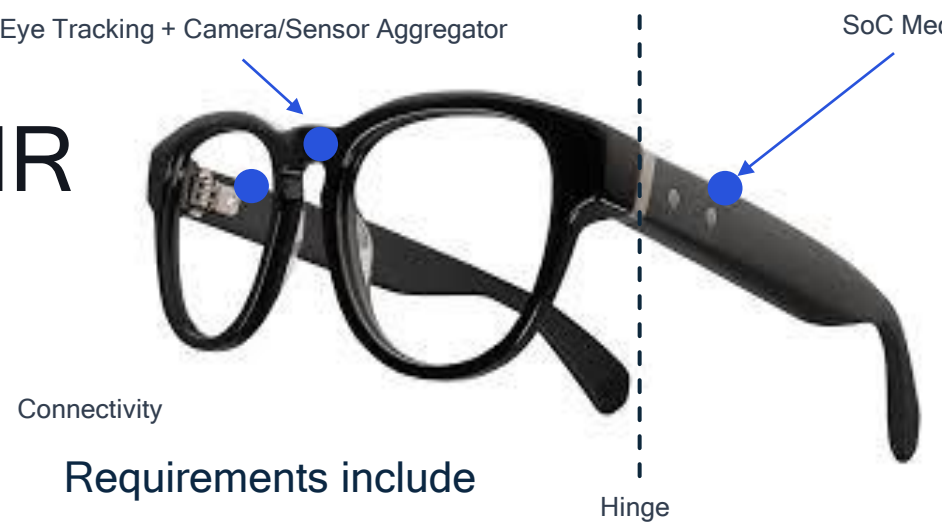
- Head & hand pose information
- Controller/input state
- Display configuration / form factor

KHRONOS GROUP

3GPP Baseline Architecture for AR/MR

Defining baseline requirements for an lightweight AR and XR device





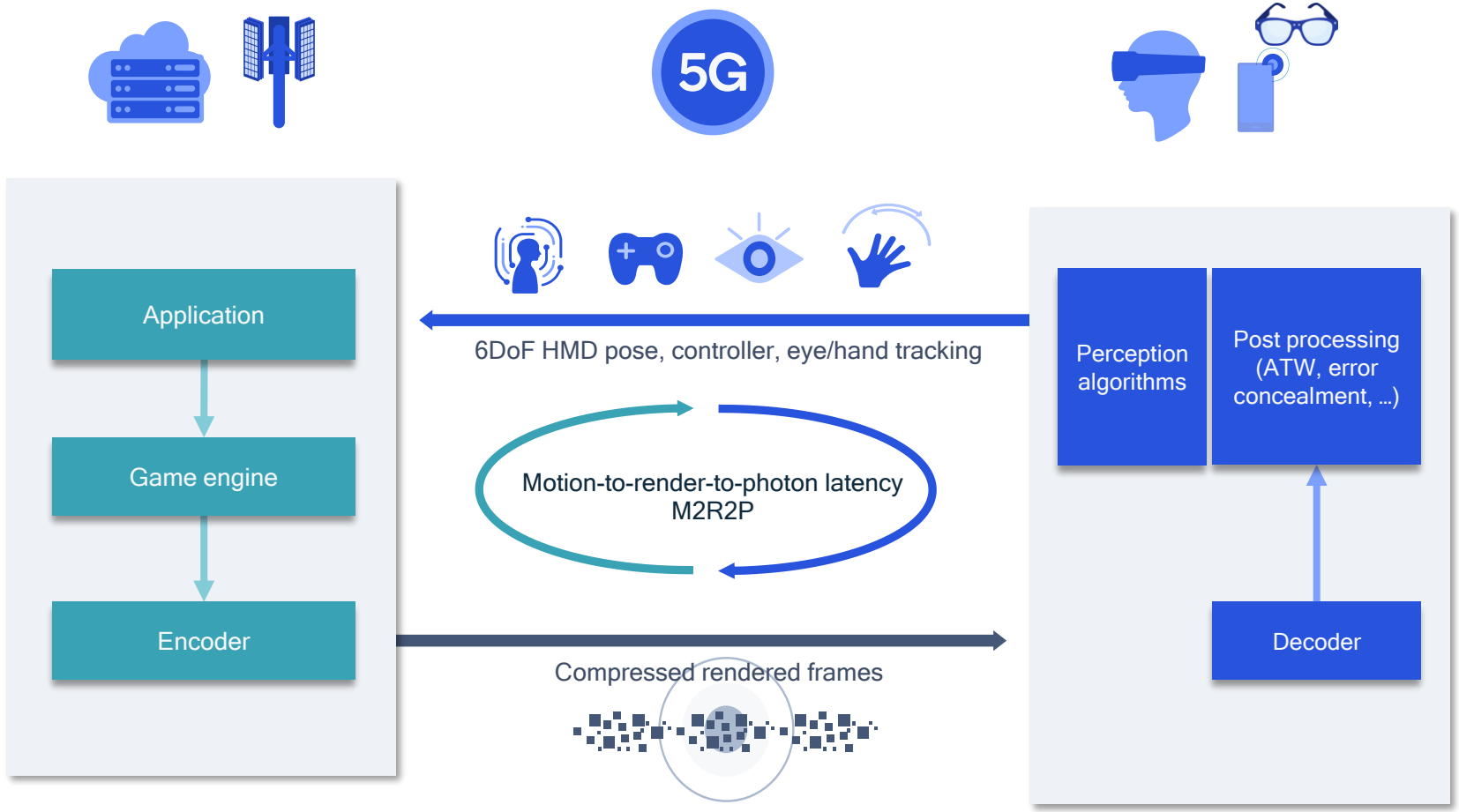
Requirements include

- OpenXR Core APIs
- rendering capabilities
- capturing capabilities
- audio and video codecs
- scene description
- Wire formats for metadata
- Profiles for split rendering and stand-alone rendering

Additional work

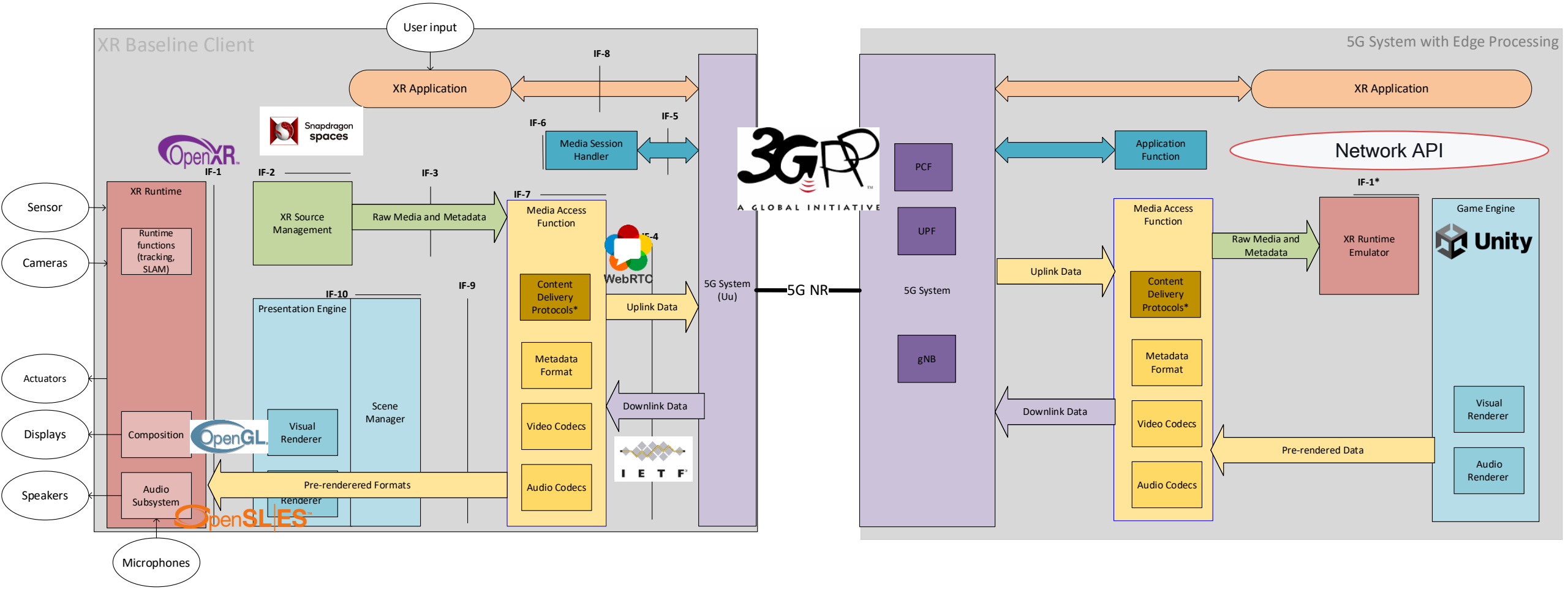
- Split Rendering
- XR Metrics
- 5G system optimizations for XR QoS and power consumption

A distributed compute architecture enables rich XR user experienc



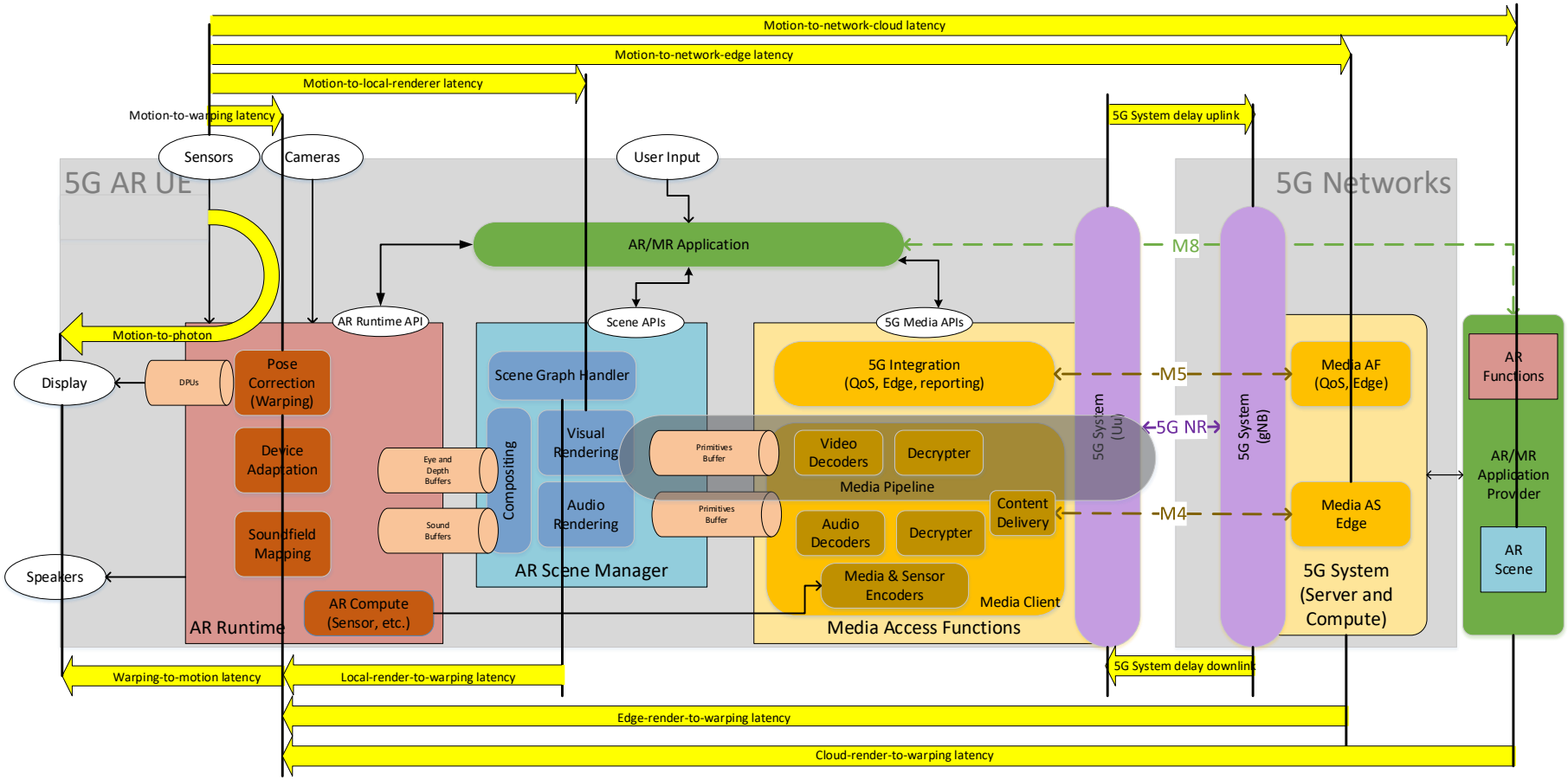
M2R2P = Edge processing + 5G round-trip time + Device processing

Split Rendering Workflow Architecture



Latency Considerations for Split Rendering

- motion-to-photon latency being less 20ms, but preferably even single digit latency below 10ms.
- pose-to-render-to-photon latency: as small as 50-80ms



Improving XR experience with 5G and 5G Advanced

Rel-16: Low power modes

Rel-16: Uplink enhancements

Rel-17: XR burst handling

Release 16, 17, 18

Lower latency
Lower power
Higher capacity

Align transmission to multimedia cadence

Enhanced CDRX and configured grant

Sleep after low latency uplink transmission

Retransmission-less configured grant

QoS based on multimedia payload

Define QoS based on PDU sets

Staggering UE traffic arrivals at gNodeB

Improved scheduler

Low latency mobility

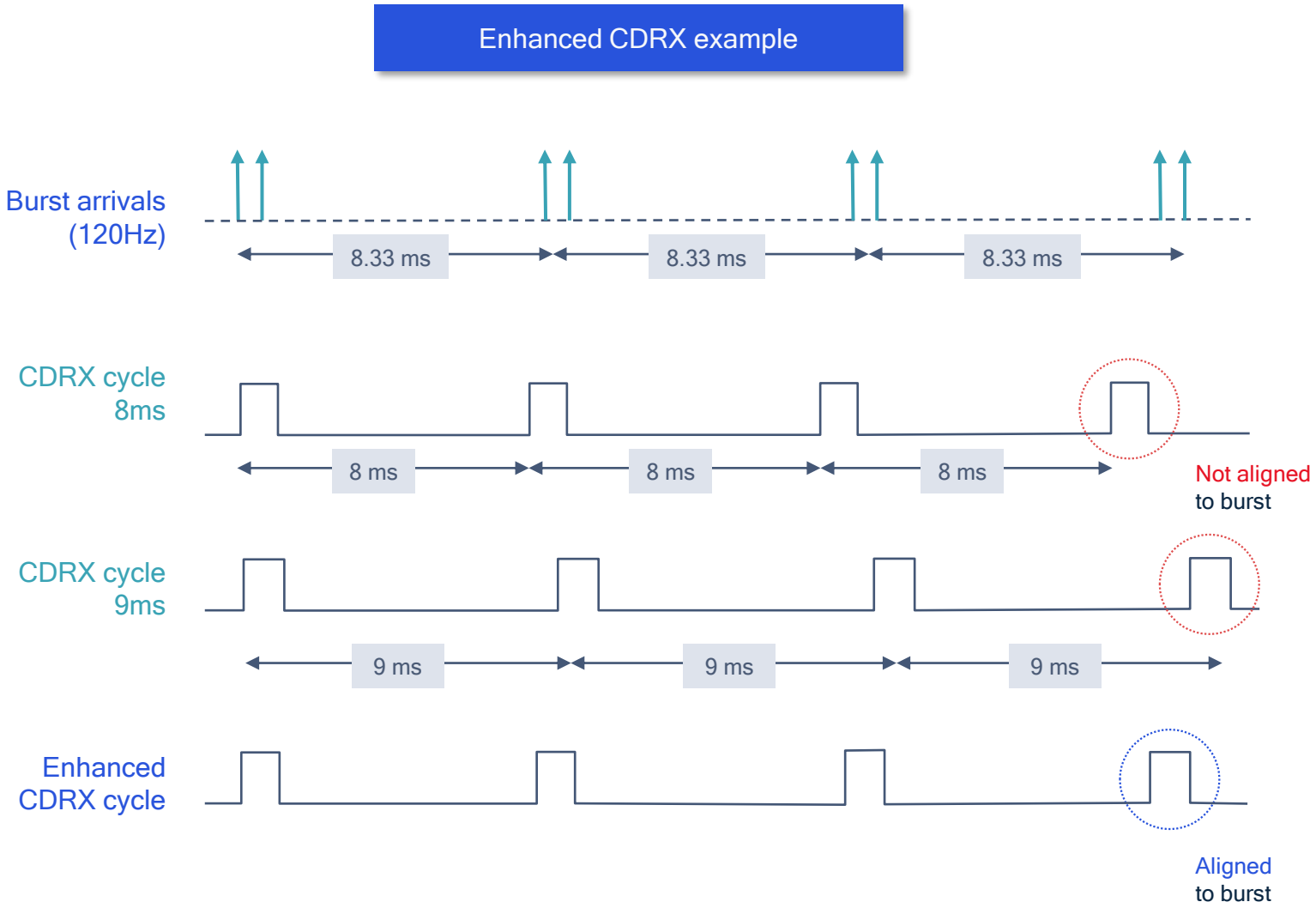
Using L1/L2 signaling for handoffs

Rel-18 aligns transmission times to the multimedia cadence

Enhanced CDRX eliminates drift between DRX-ON and XR traffic

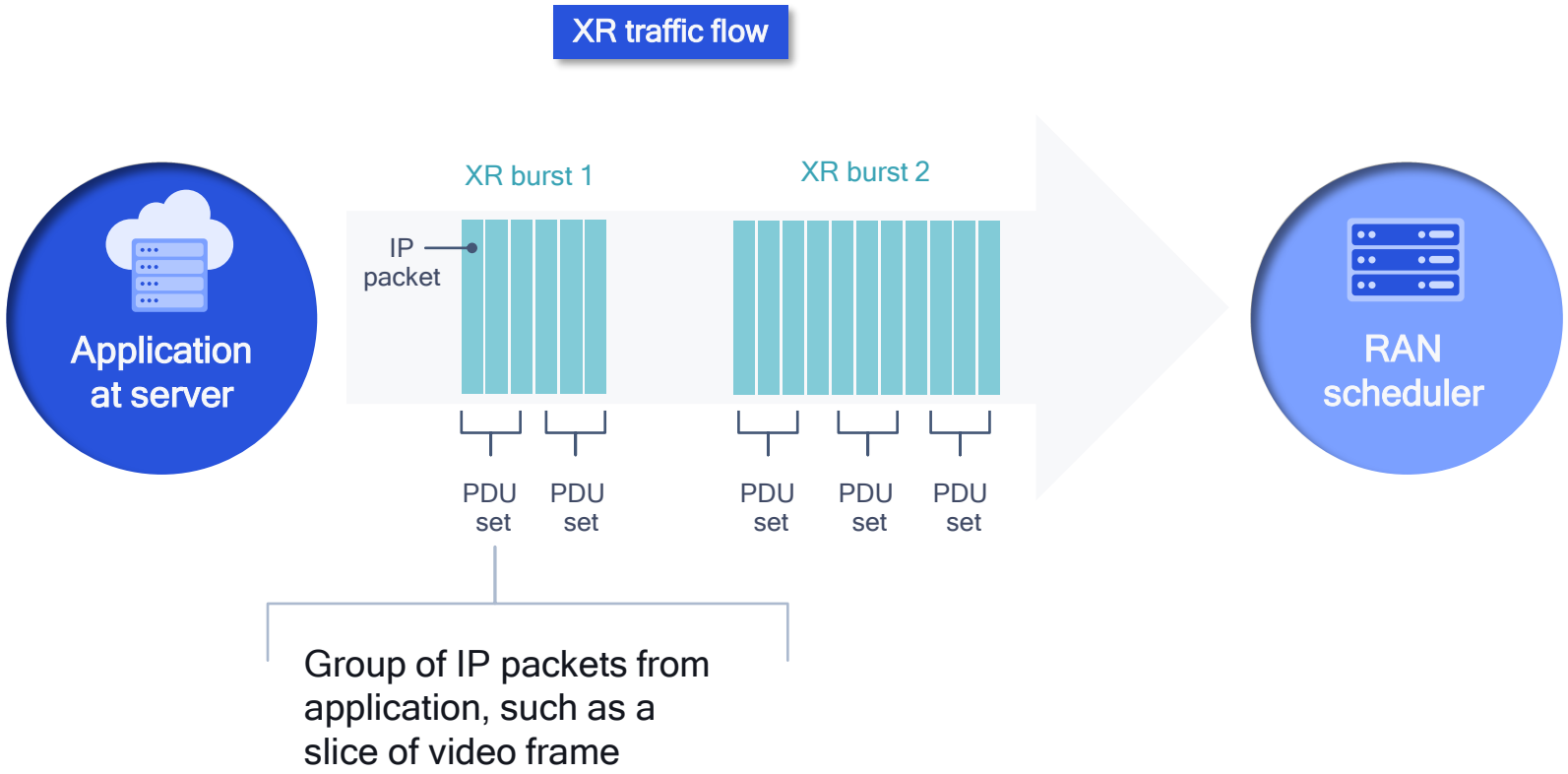
Enhanced CG eliminates drift between CG and XR traffic

Reduces latencies and device power consumption



Rel-18 enables QoS based on multimedia traffic payloads instead of IP packets

- QoS defined for PDU-sets
- QoS parameters include error rate and delay
- Enables RAN schedulers to satisfy multimedia QoS requirements



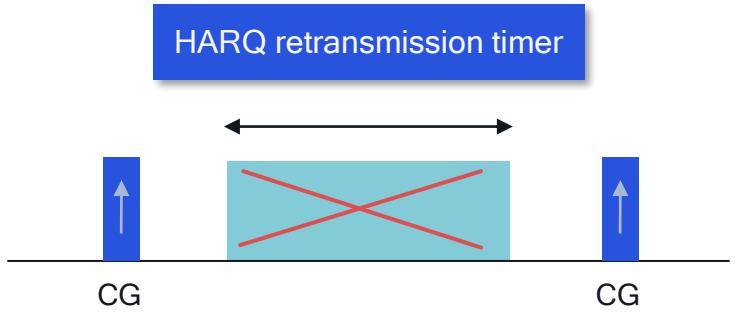
Rel-18 enables devices to sleep after uplink transmission

Retransmission-less Configured Grant:

1. Uses conservative MCS to improve reliability of first transmission
2. Avoids the need for the UE to monitor control channel after CG
3. Allows longer sleep cycles reducing device power consumption

Legacy device is ON for potential retransmissions

Rel-18 device can sleep



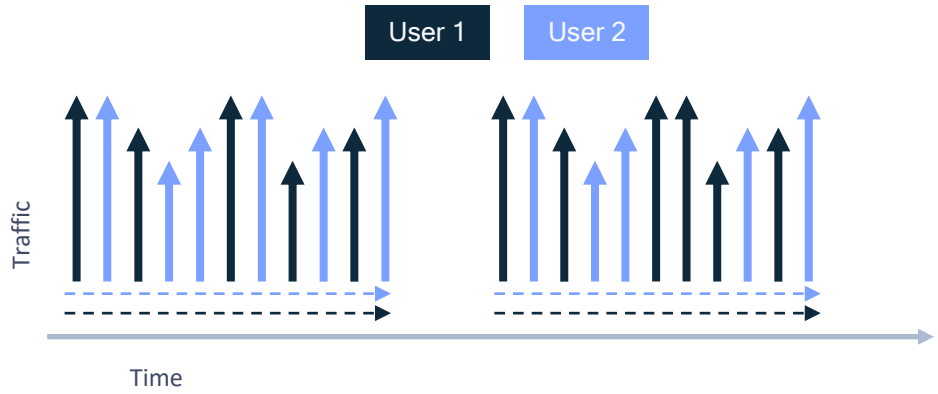
Low-latency, low-bandwidth transmissions
(e.g., 100-byte pose)

Rel-18 staggers UE application traffic arrivals from the server to the base station

Base station signals the user burst arrival time offset to the server

Server staggers user traffic to the base station

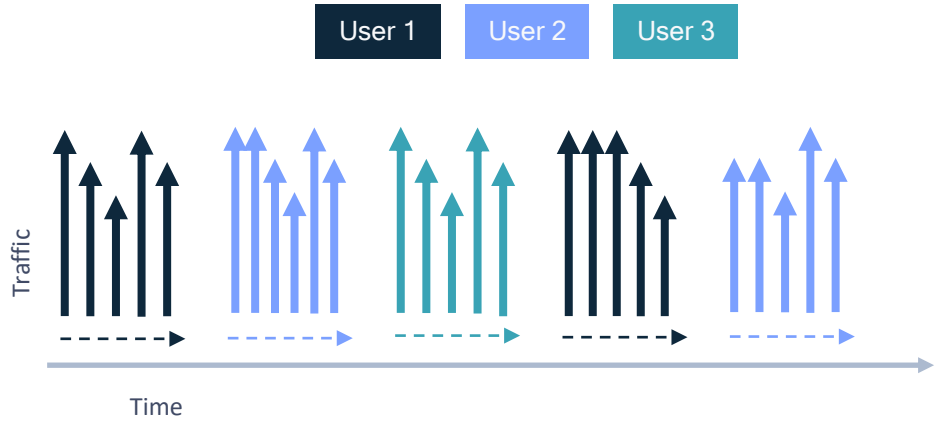
Improves base station scheduler



No UE application traffic staggering

Inefficient scheduler

- Higher latencies
- Higher power / device ON-time
- Lower capacity



UE application traffic staggering

Improved scheduler

+ Lower latencies
+ Lower power / device ON-time
+ Higher capacity

6G protocols can natively support distributed compute

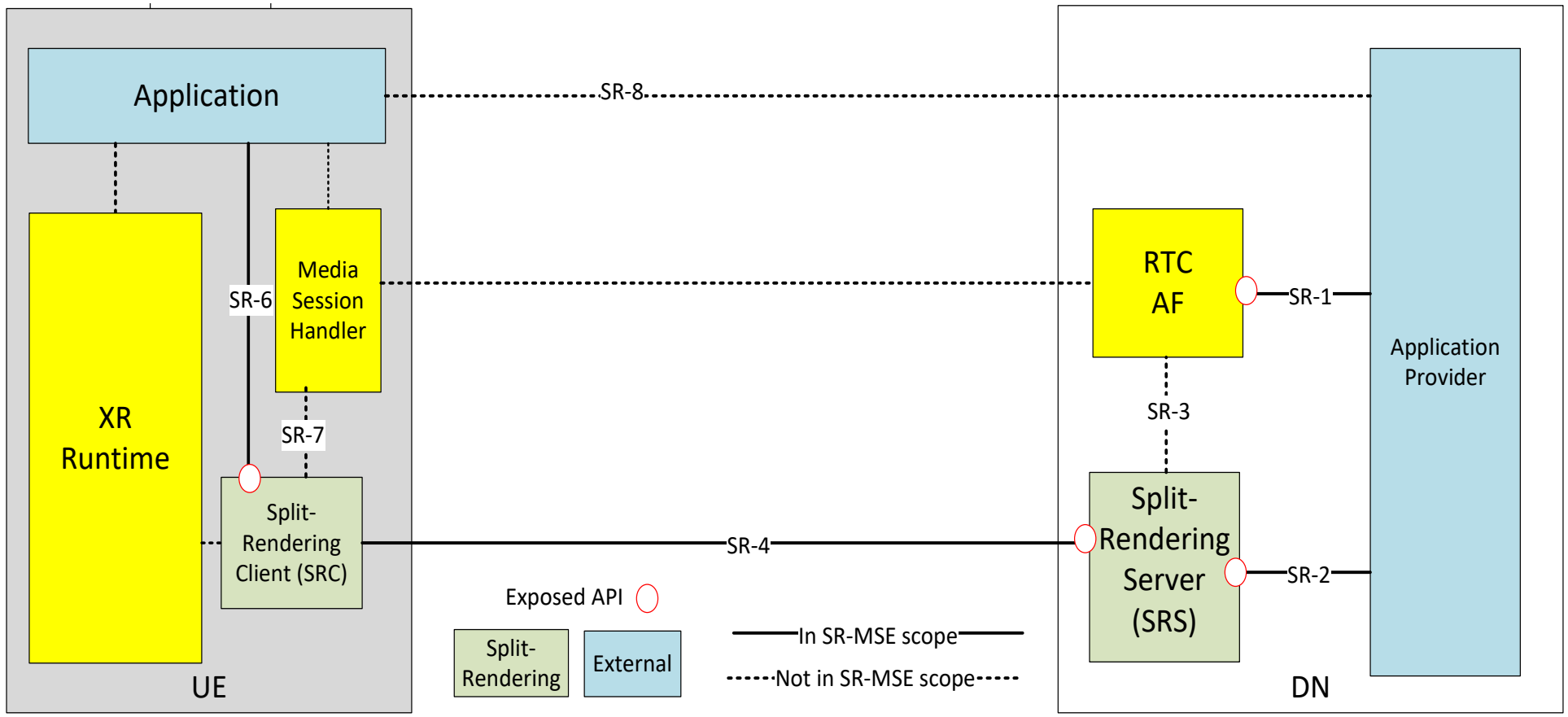
Dynamically distributing workloads

Delivering low power and low latencies

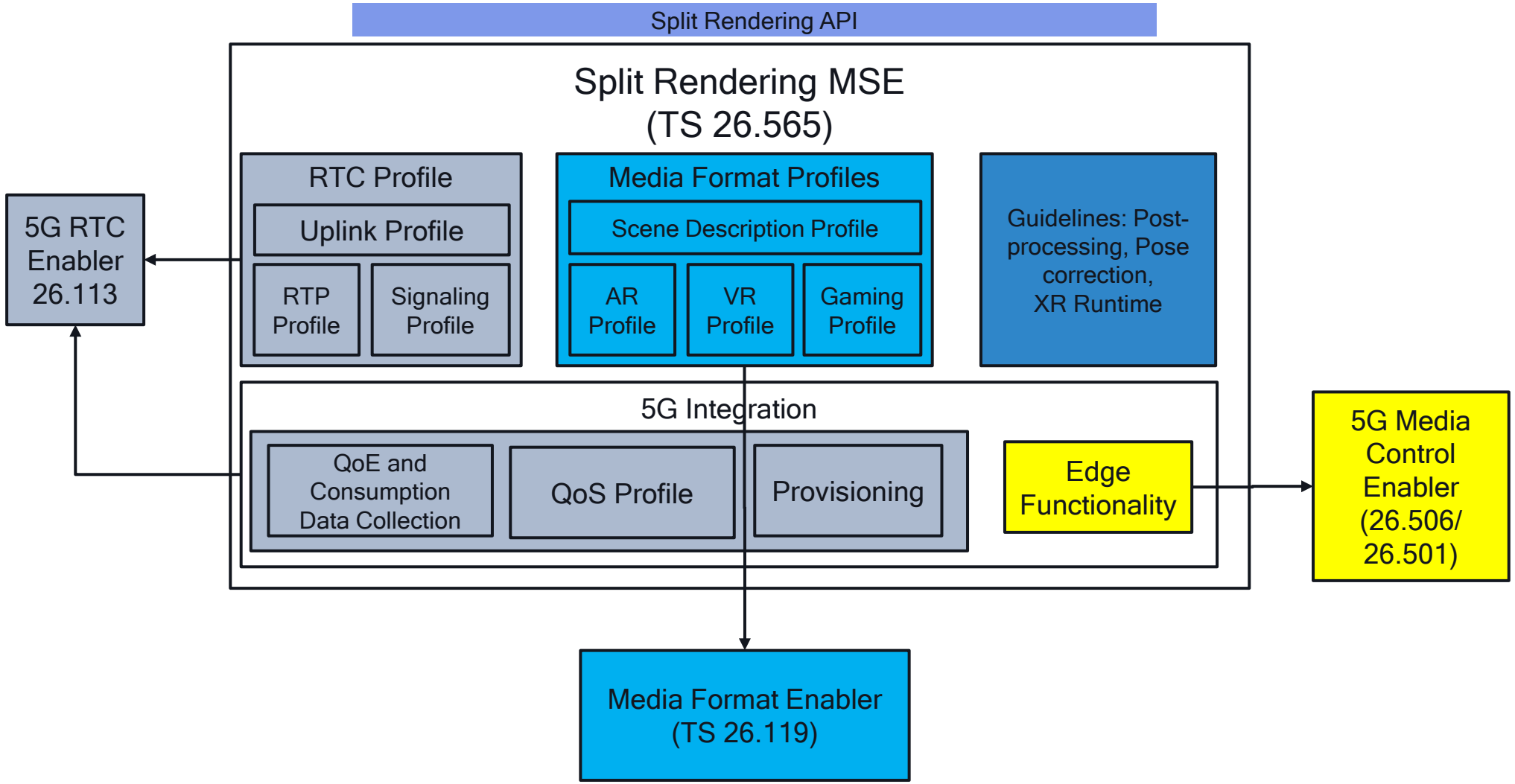
Incorporating e2e QoS and session handling

- 6DoF/SLAM
- Hand Tracking
- Facial Expression
- Eye Tracking
- 3D Graphics Rendering
- Scene Understanding
- Objected Rec and Tracking
- Camera Processing

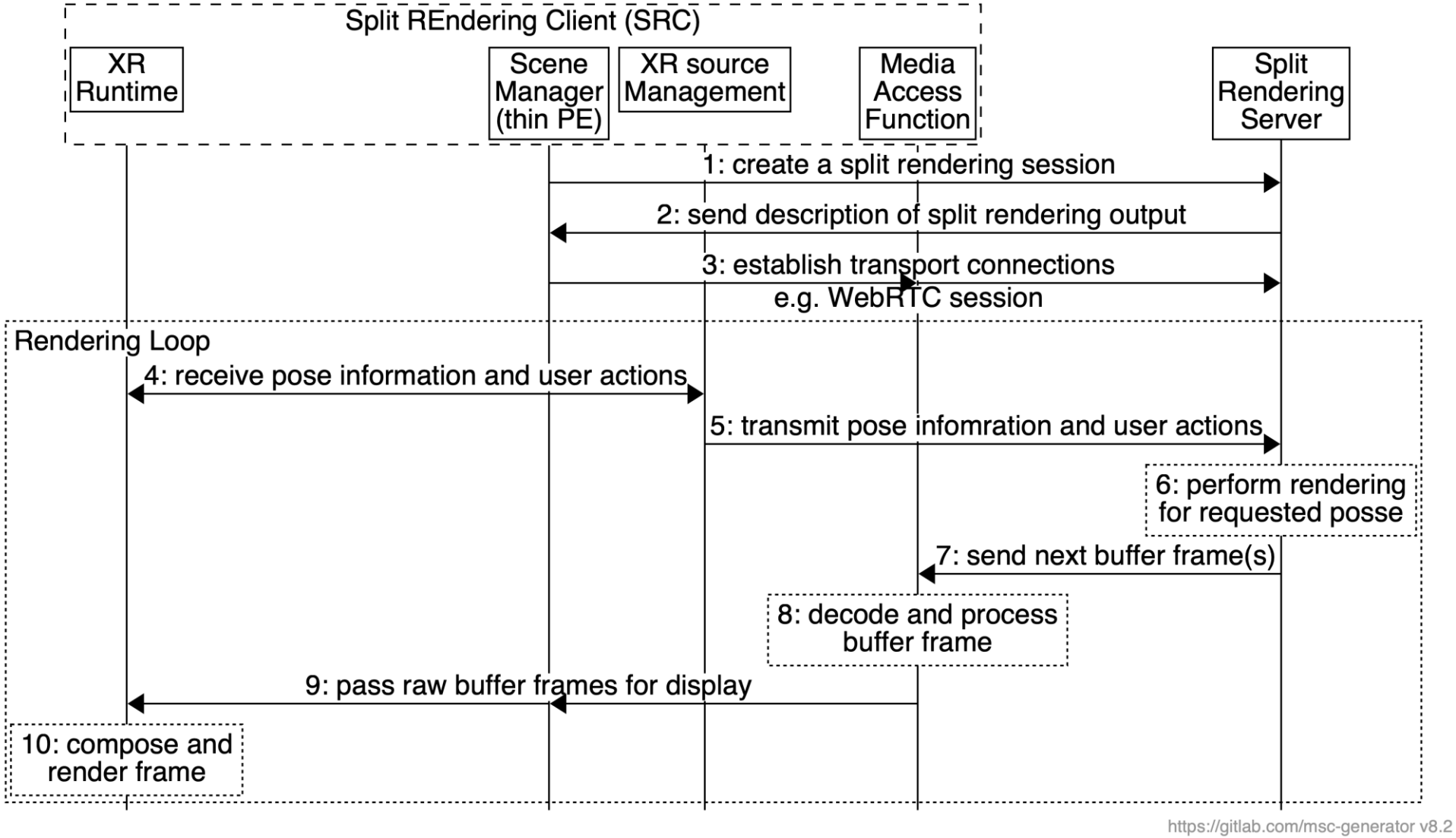
Split Rendering Interfaces



Standardizing Split XR in SA4 – Media Service Enabler

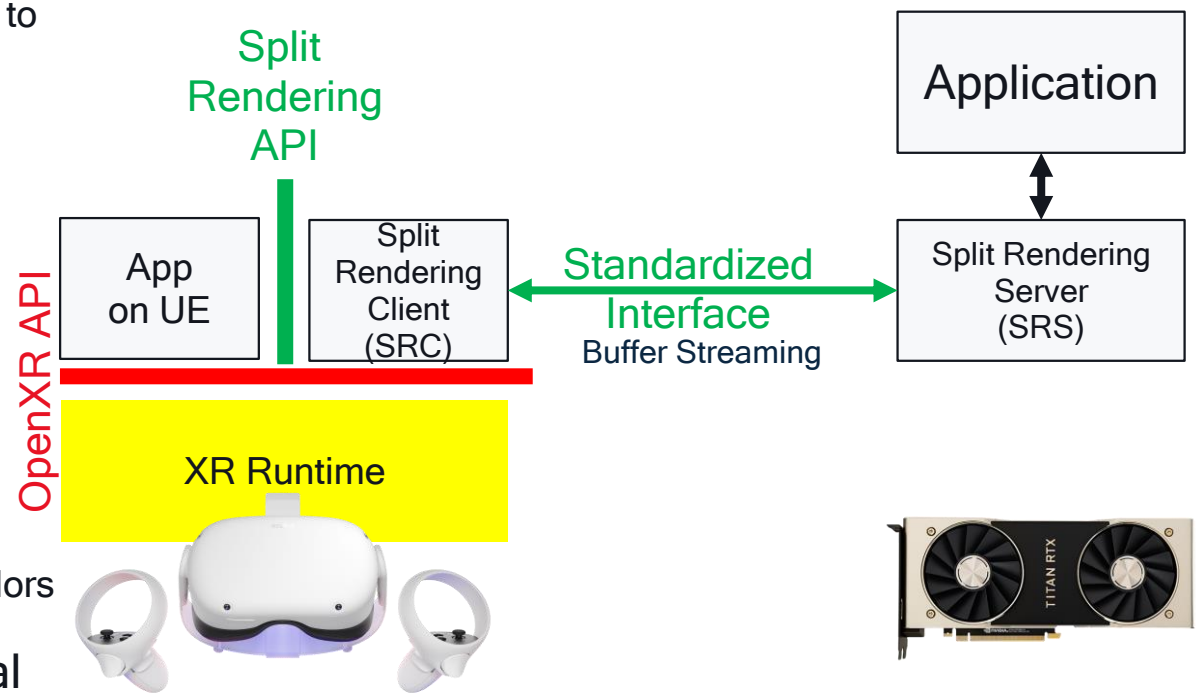


Call Flows


Split REndering Client (SRC)
XR Runtime
Scene Manager (thin PE)
XR source Management
Media Access Function
Split Rendering Server
1: create a split rendering session
2: send description of split rendering output
3: establish transport connections
e.g. WebRTC session
Rendering Loop
4: receive pose information and user actions
5: transmit pose infomration and user actions
6: perform rendering for requested posse
7: send next buffer frame(s)
8: decode and process buffer frame
9: pass raw buffer frames for display
10: compose and render frame
https://gitlab.com/msc-generator v8.2

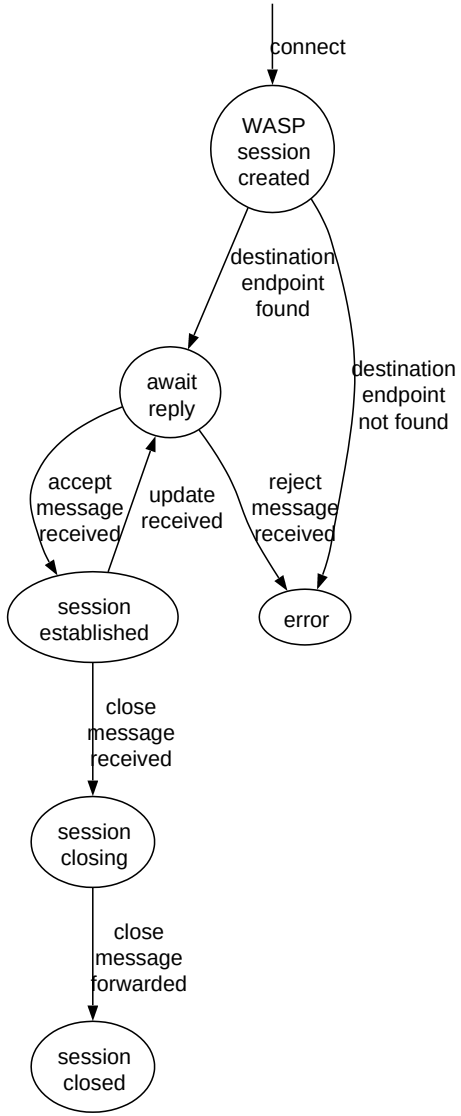
Split Rendering and OpenXR

- SR Server
  - Typically a Game Engine plugin
  - Configures Application rendering (e.g. sets it to stereo or cubemap to match OpenXR projection configuration)
  - Emulates game input
  - Captures/encodes/transmits rendering output to UE
  - Syncs up with SRC on UE
- SR Client
  - Runs on the UE/HMD
  - Discovers and connects to SRS on edge
  - Application/SRC on UE owns the OpenXR session
- Interoperable Design
  - Streamer Service and Streamer App may come from different vendors
- Robust as App on UE is always able to fallback to local rendering



Signaling Protocol

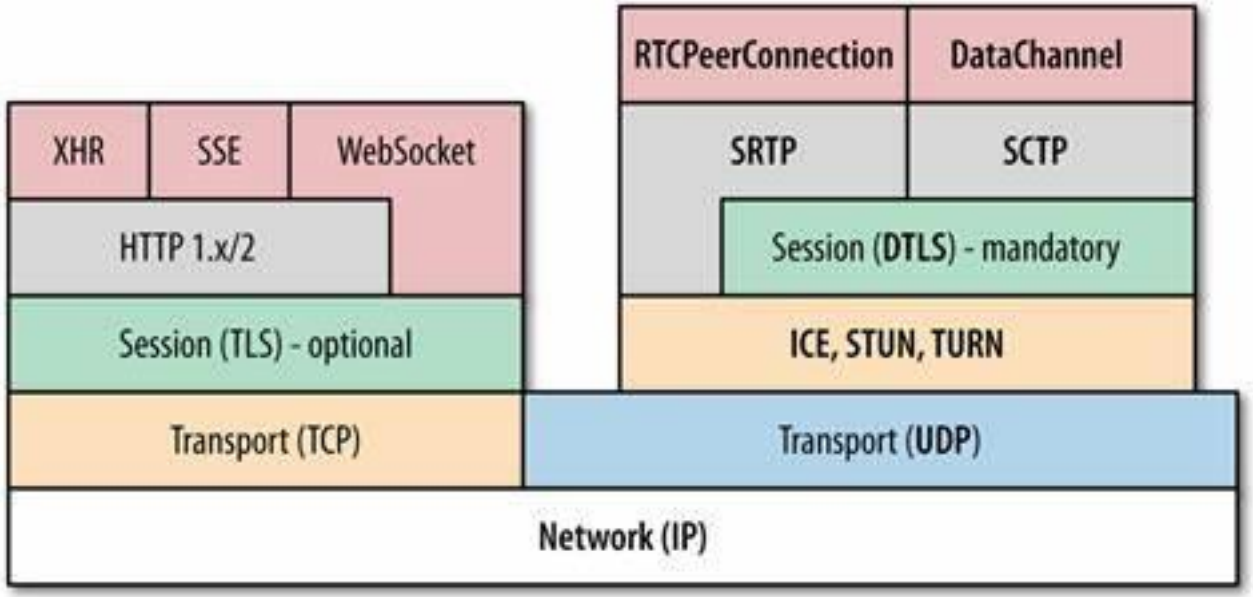
- Negotiate rendering split between SRC and SRS
  - SRC supplies the XR session configuration to the SRS
  - SRS responds with a selected split rendering description
- Establish a WebRTC connection between SRC and SRS
- SWAP is a WebSockets-based protocol
- Messages are JSON formatted
- Split Rendering information is exchanged through application-specific messages
- SWAP Server may pass information to the AF
  - QoS related information extracted from offer/answer negotiation
  - RTP header extension configuration for PDU Set and Data Bursts
- Specified as part of iRTCW in TS26.113



https://gitlab.com/msc-generator v8.2

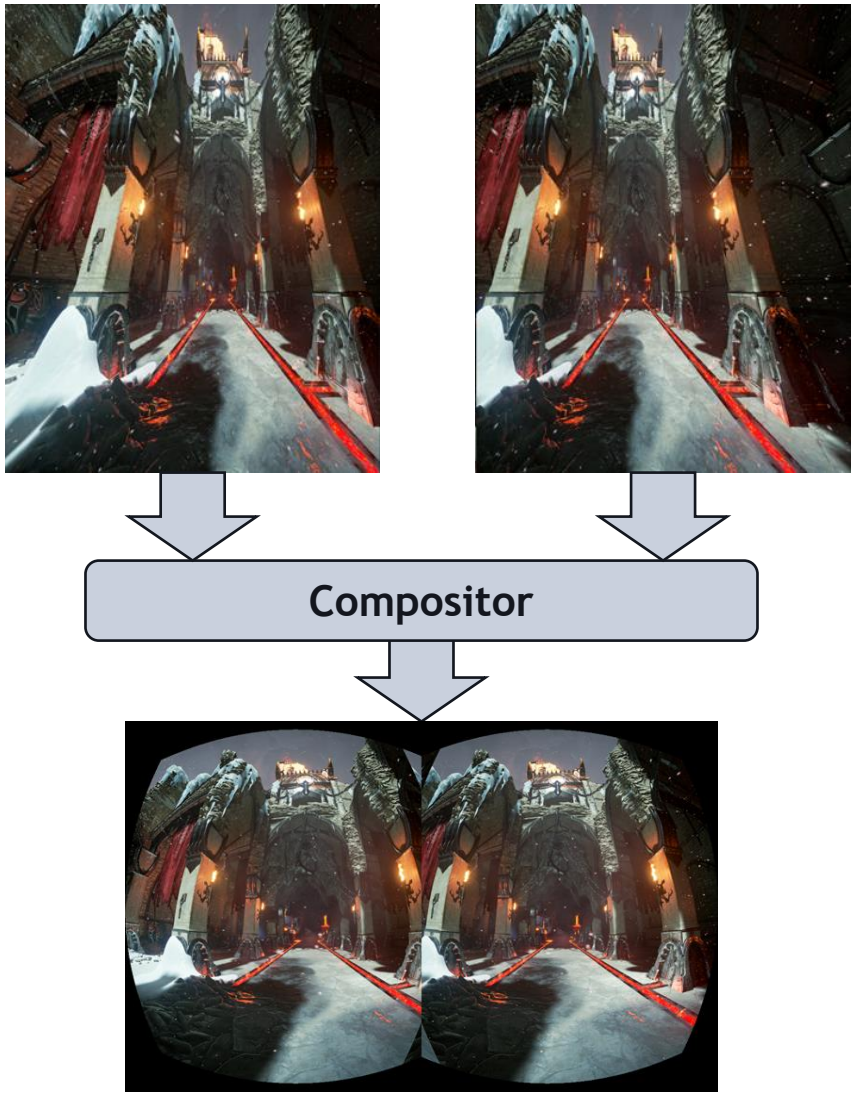
Transport Protocols

- WebRTC framework is used as basis but native implementations are expected
- RTP header extensions are defined for:
  - Marking of PDU Sets and Data Bursts
  - Signaling of rendering pose and action list
- Data channel
  - Used mostly for uplink metadata
  - Transports pose and action information to the SRS
- Guidelines to ensure low latency WebRT stack
  - Codecs specified by 3GPP
  - Congestion control aligned with QoS negotiation
- Specified in 5G_RTP TS26.522



Split Rendering Negotiation

- Split Rendering Client maintains XR session
- SRC tells SRS about configuration of the XR session and its rendering capabilities
- SRS replies with a description of the rendered format
  - May cover a wide range of configurations from 2D to 3D
- Configuration information may include:
  - View configuration
  - Composition layer configuration
  - Swapchain resolution and level of detail
  - Rendering capabilities
- SR Description is proposed to be a glTF + related extensions
  - Allows for alternative operation points

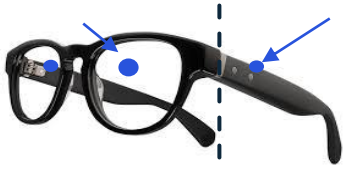


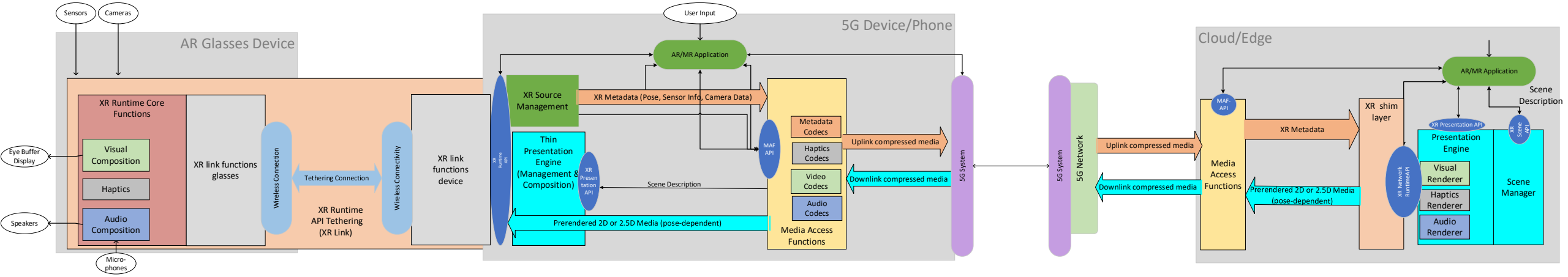
Summary of what all needs to be defined

- Wire-formats for the metadata
  - Spaces and coordinate systems
  - Render pose in downlink
  - Uplink predicted pose information and XR Runtime actions
- Pre-rendering formats
  - Multiple video buffers (left & right eye, depth, overlays, projection formats) to support pose correction
  - Multiple audio buffers (pre-rendererd formats for binaurilization)
- Compression and codecs
  - Support for multiple concurrent video and audio decoders – with minimum capabilities (resolution, formats, frame rates)
- Content Delivery Protocol in uplink and downlink including adaptation, security/privacy & metadata
- Session Establishment, capability exchange, edge resource establishment, etc.
- QoS framework to support the latency, reliability and bitrate requirements
- QoE framework to continously monitor and measure the quality.

Variants - Smartly Tethering AR Glasses (SmarTAR)

Study item in 3GPP SA4





- Issues:
  - Non-provisioned additional tethering link (e.g. WiFi) ➔ QoS and latency measurement and control
  - End-to-end formats to avoid transcoding and security issues
  - Combination of UE/Phone and network rendering
  - Distribution of compute resources across glasses, UE and edge/cloud

glTF 2.0 Extensions in MPEG and 3GPP – Real-time exchange formats for 3D Experiences

Imed Bouazizi

Thomas Stockhammer

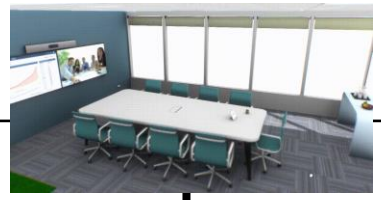
Use Case 1 – Accessing Dynamic and Interactive 3D Scenes for AR/MR/VR

- Examples: Live Entertainment, Movies, Games, etc.
- Typically,
  - Collection of static and dynamic objects to be accessed and presented in parallel
  - huge volume and data sizes
  - Interactive and immersive
  - Different media types
- Access through download, or preferably smart streaming
- Adaptation to device capabilities

Use Case 2 - AR Call/Conferencing

- Extend existing IMS and WebRTC-based telephony services
- Multi-party calls and conferencing with AR
- Users join with realistic avatars and contribute 2D/3D content
- Create shared spaces where call participants can interact

High Level Procedure



Scene of shared space

Scene of shared space

Add nodes to scene

Add nodes to scene

Send scene updates

Send scene updates

Pose and animation updates

Pose and animation updates

Send scene updates

Send scene updates

Update Loop

Interchange Formats for Portable Dynamic and Interactive 3D Scenes

MPEG is developing core compression and representation formats and the distribution/system integration of those

MPEG organization under ISO/IEC JTC1/ SC29

WG 2
MPEG Technical Requirements

Igor Curcio, WG2 Convenor

WG 3
MPEG Systems

Youngkwon Lim, WG3 Convenor

WG 4
MPEG Video Coding

Lu Yu, WG4 Convenor

WG 5
Joint Video Coding Team with ITU

Jens-Rainer Ohm, WG5 Convenor

WG 6
MPEG Audio Coding

Schuyler Quackenbush WG6 Convenor

WG 7
MPEG Coding for 3D Graphics and Haptics

Marius Preda, WG7 Convenor

WG 8
MPEG Genomic Coding

Marco Mattavelli, WG8 Convenor

Jörn Ostermann, AG2 Convenor

AG 2
MPEG Technical Coordination

Kyuheon Kim, AG3 Convenor

AG 3
MPEG Liaison and Communication

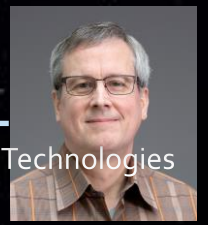
Peter Schelkens, AG4 Convenor

AG 4
JPEG and MPEG collaboration

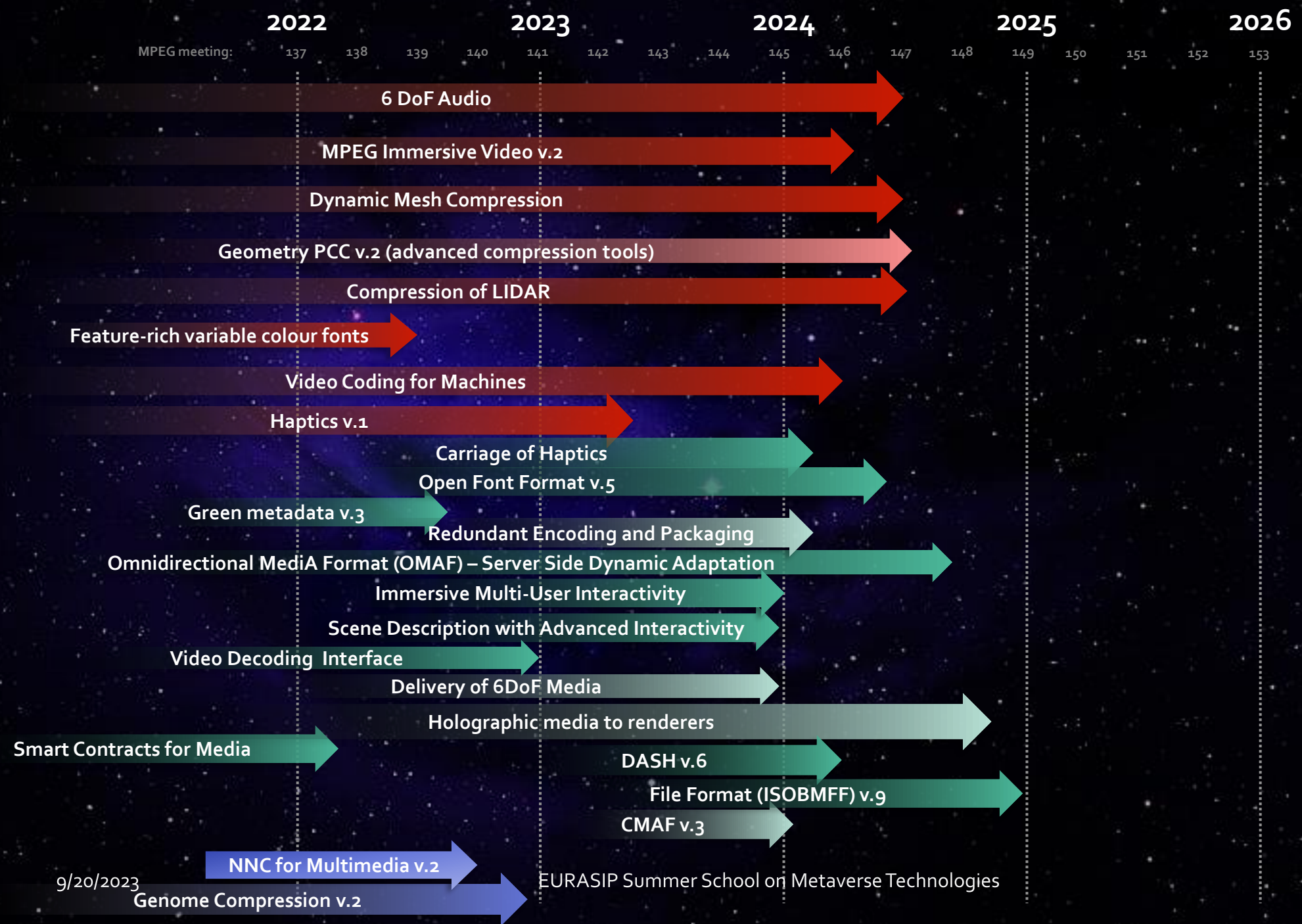
Mathias Wien, AG5 Convenor

AG 5
MPEG Visual Quality Assessment

SC 29



Gary Sullivan, SC29 Chairman

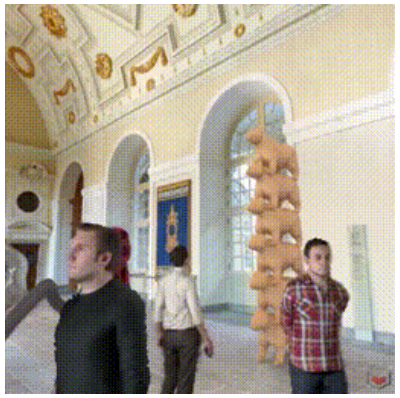

2022
2023
2024
2025
2026
MPEG meeting:
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
6 DoF Audio
MPEG Immersive Video v.2
Dynamic Mesh Compression
Geometry PCC v.2 (advanced compression tools)
Compression of LIDAR
Feature-rich variable colour fonts
Video Coding for Machines
Haptics v.1
Carriage of Haptics
Open Font Format v.5
Green metadata v.3
Redundant Encoding and Packaging
Omnidirectional MediA Format (OMAF) – Server Side Dynamic Adaptation
Immersive Multi-User Interactivity
Scene Description with Advanced Interactivity
Video Decoding Interface
Delivery of 6DoF Media
Holographic media to renderers
Smart Contracts for Media
DASH v.6
File Format (ISOBMFF) v.9
CMAF v.3
NNC for Multimedia v.2
Genome Compression v.2
Media Coding
Systems and Tools
Beyond Media

From MPEG-I to Dynamic Scenes

MPEG-I is a collection of representation formats for 2D and 3D objects of different media type

Included technologies in MPEG-I ISO/IEC 23090

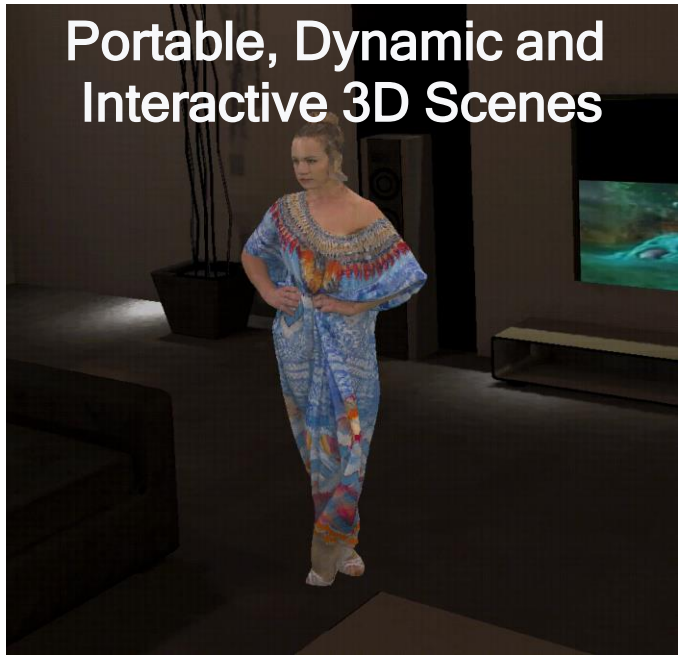
- File and application formats with metadata
- Immersive Audio and Haptics
- Video Decoding Interface
  - Leverage and optimize 2D decoders for 3D and immersive media
- Dynamic Point Clouds and V-Mesh
  - Compress 3D formats such as point clouds and dynamic meshes
- MPEG Immersive Video
  - Support multi-view and light field displays
- 2D Video Compression
  - H.266/VVC
  - AI/ML-based and VCM
  - H.267/ECM



MPEG studied different options in 2019 for Scene Descriptions, including defining its own.

Scene Description

How?



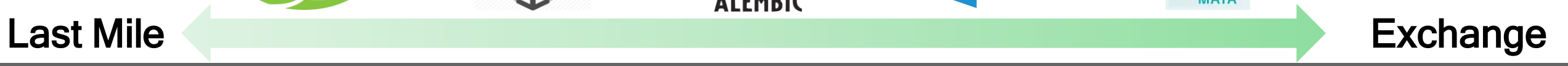
Choosing the right format - Considerations

Last Mile

- Low Complexity
- Flat Hierarchy
- Compressed components
- Adaptive and network friendly
- Support for Texture/Light Baking

Exchange

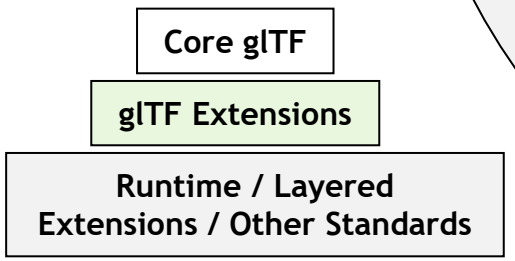
- High fidelity
- Superstructure
  - Hierarchical
  - Distributed
  - Preserves author's intents/choices
  - Documents authoring process
- Lossless
- Preserves asset's metadata/versioning

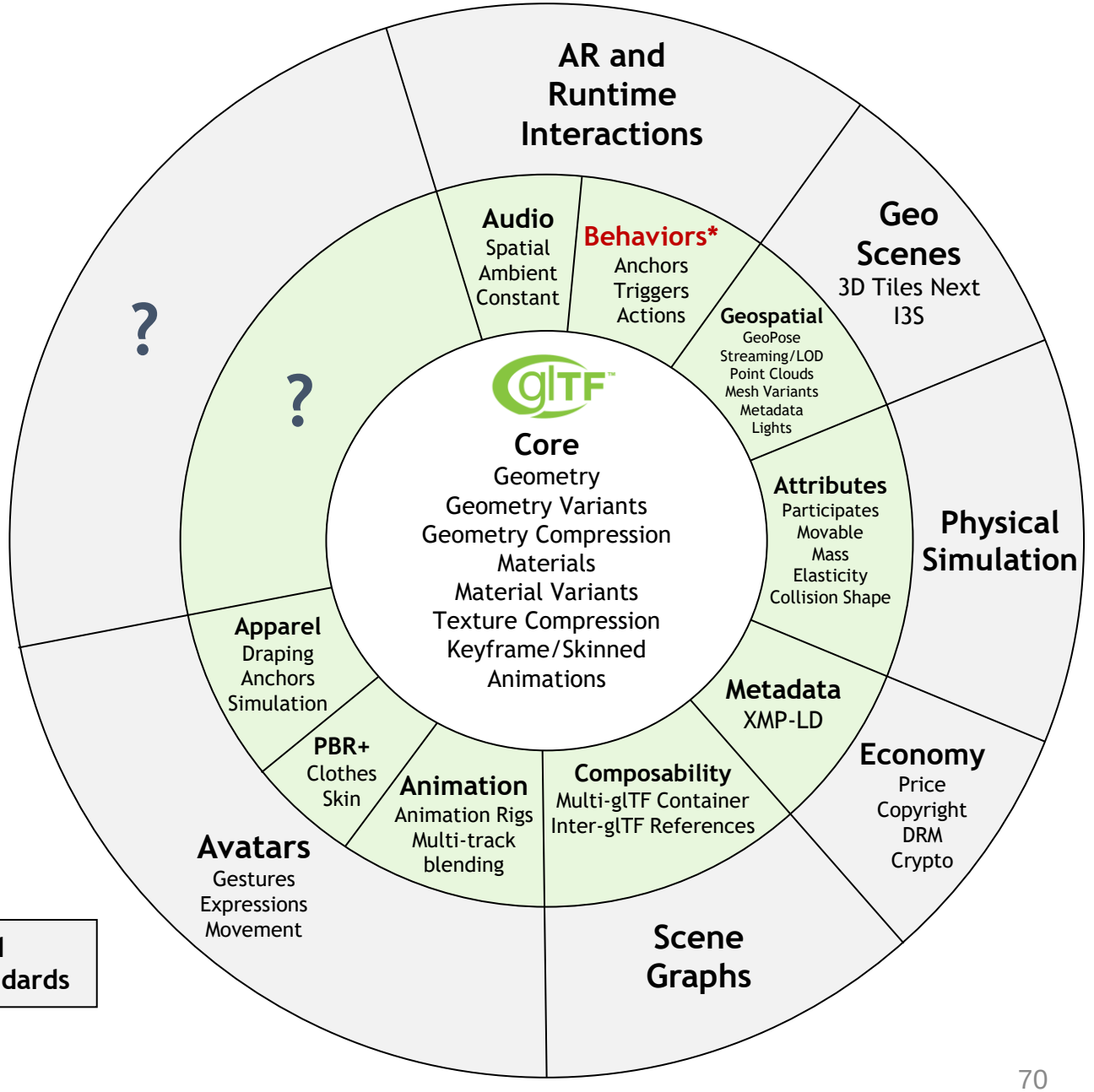


glTF Roadmap

The metaverse is driving many key glTF use cases and requirements

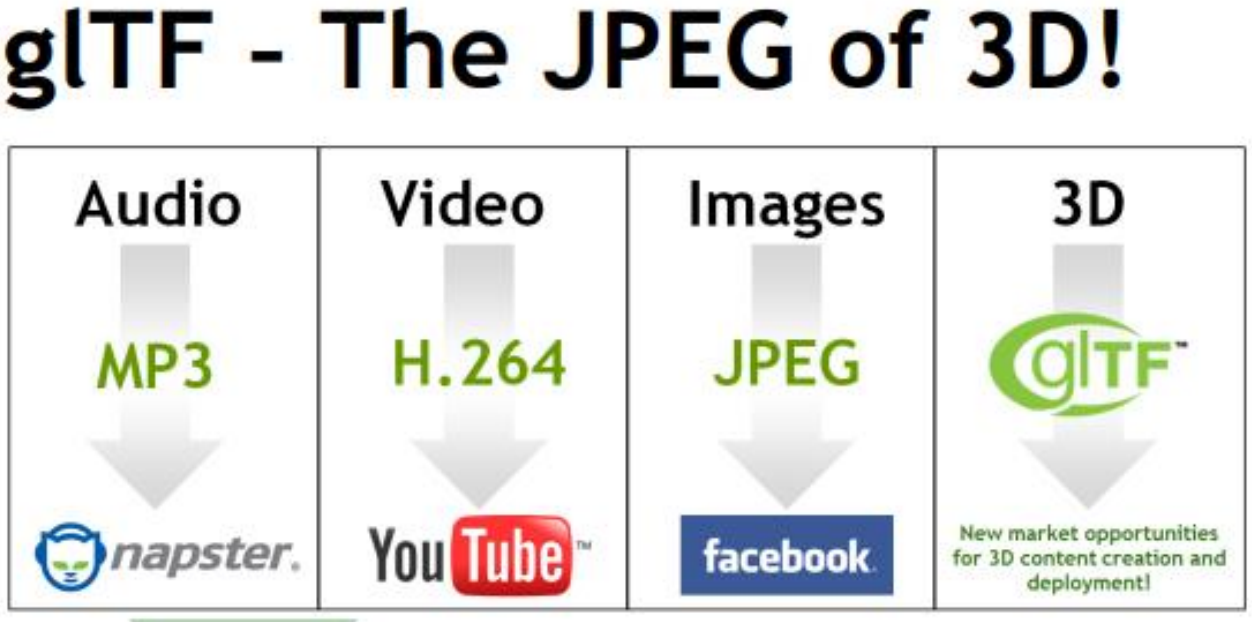
KHRONOS GROUP





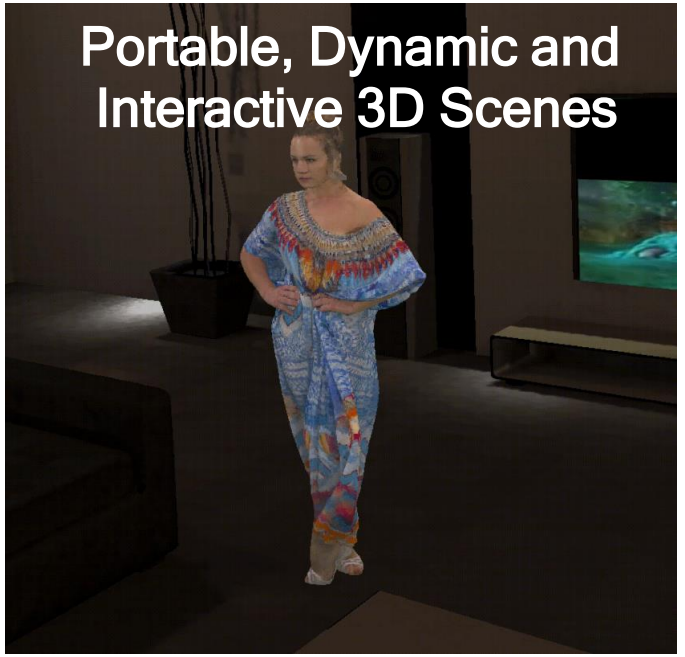
From glTF to Dynamic Scenes

glTF is a standard file format for three-dimensional scenes and models. An open standard developed and maintained by the Khronos Group, it supports 3D model geometry, appearance, scene graph hierarchy, and animation. It is intended to be a streamlined, interoperable format for the delivery of 3D assets, while minimizing file size and runtime processing by apps. As such, its creators have described it as the "JPEG of 3D."



Time Dimension

How ???



The Idea: Combine glTF and MPEG-I ...

... in order to create the "HTML-5 for XR/Metaverse"

Multimedia Elements

- are built-in in HTML5
- no plug-ins needed
- Streaming and DRM content

HTML vs HTML

Mobile support

- User's Geographical Location
- Simplicity
- Re-use of hardware codecs

Client Server Communication

- full duplex communication client/server
- web sockets/webRTC

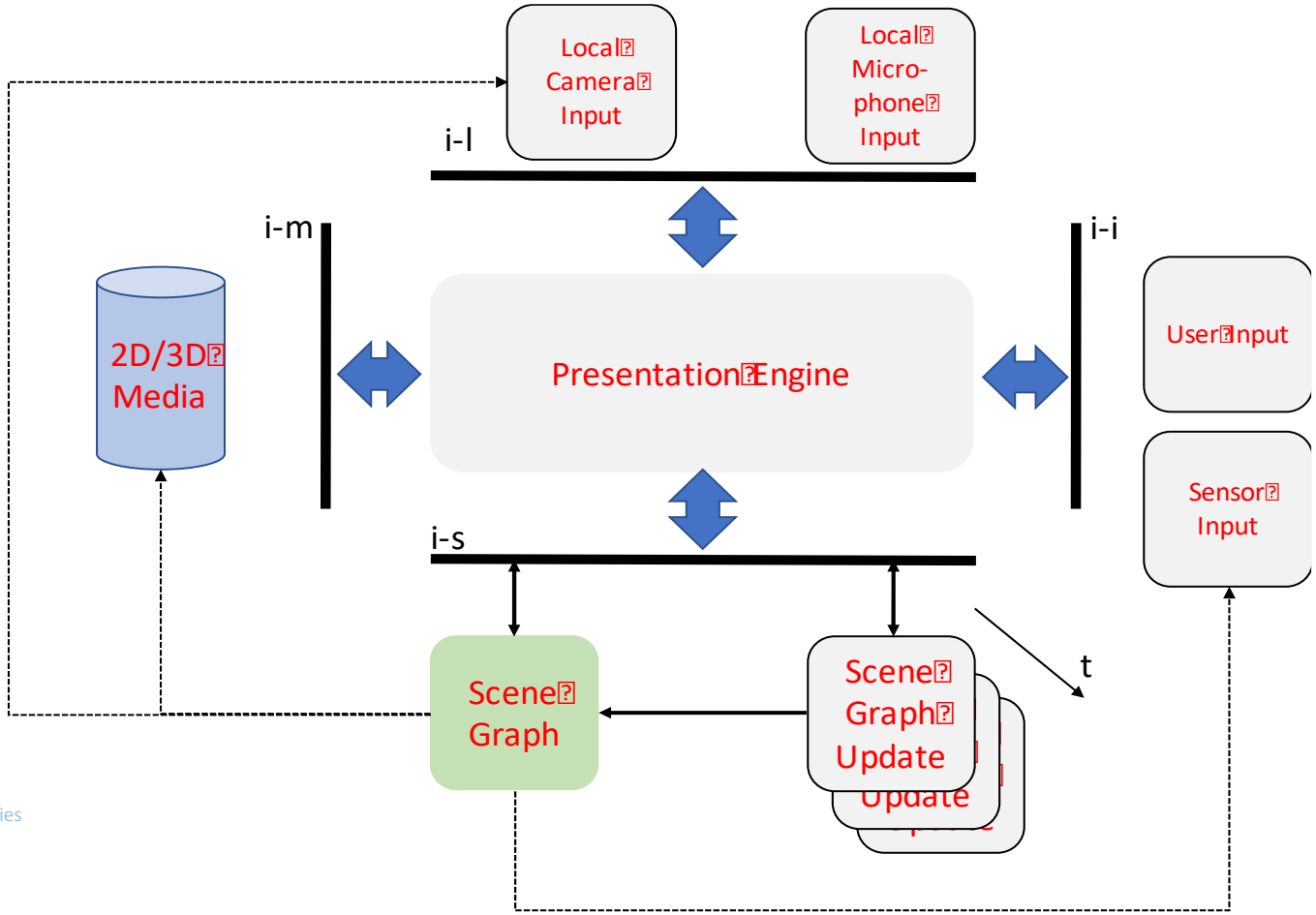
Advanced Scripting

- JavaScript support
- Multi-Threading, etc.

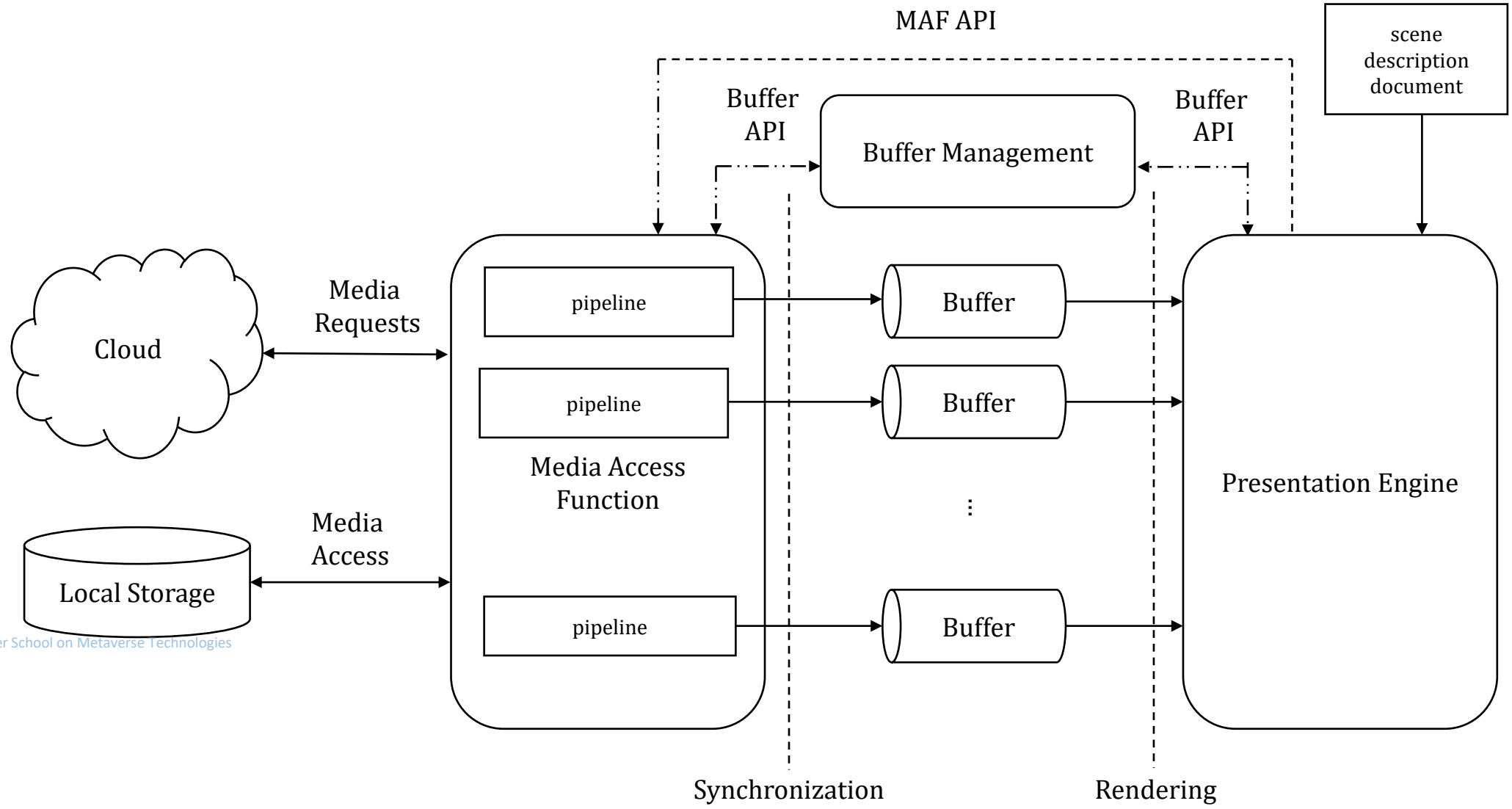
- Combining glTF and MPEG-I is addressed in **MPEG-I Scene Description ISO/IEC 23090-14**
  - Entry point document to the 3D experience
  - From shared experiences to 6DoF content and XR
  - Extensions to glTF for networked and real-time media

MPEG Immersive Media Architecture

Interfaces

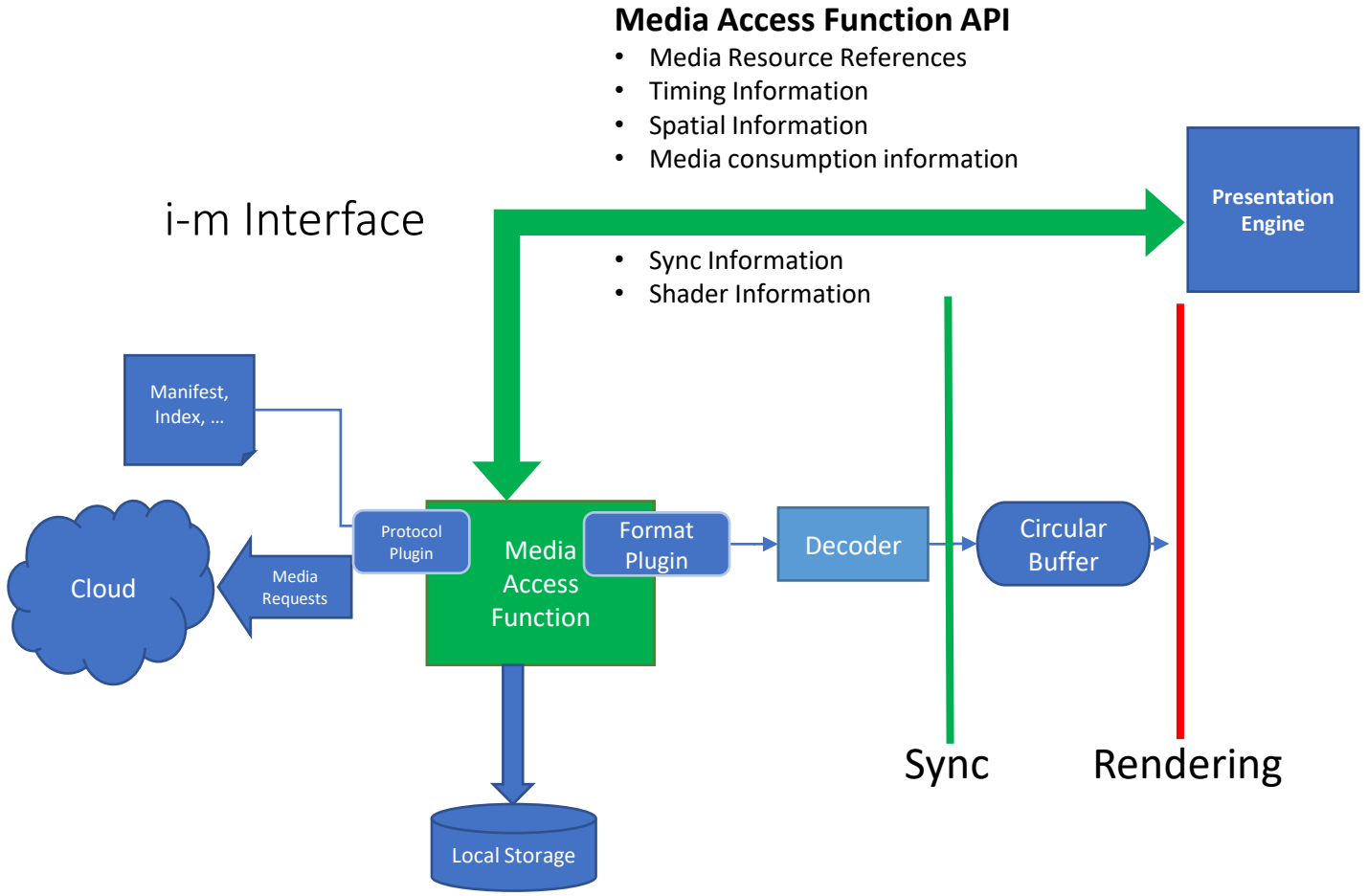

Local Camera Input
Local Micro-phone Input
i-l
i-m
i-i
2D/3D Media
Presentation Engine
User Input
Sensor Input
i-s
Scene Graph
Scene Graph Update
t

Scene Description Architecture



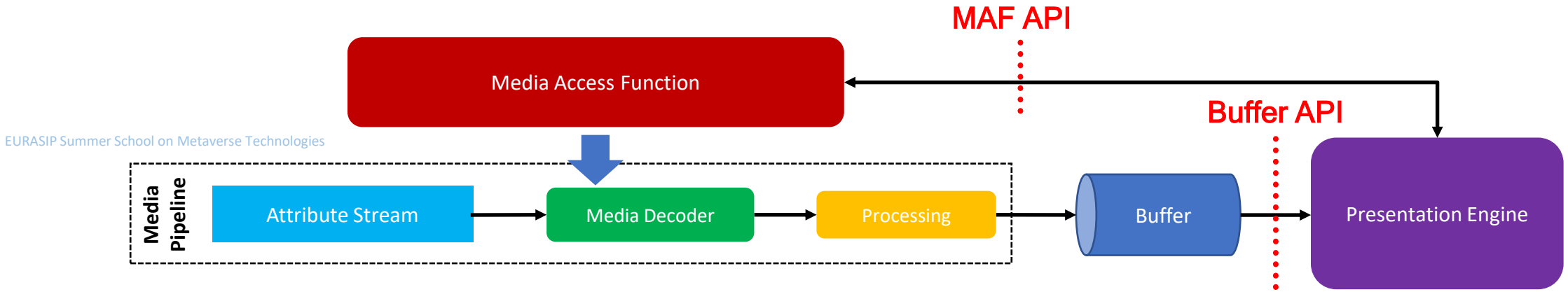
Media Access Function (MAF) API

- Support for wide range of formats through Plugins
- Endpoint for the Media Access Function API
- Optimized Media Fetching
  - Random spatial and temporal access
  - Partial delivery matching Presentation Engine needs
  - Integration with Cloud and Edge media processing



Media Pipelines

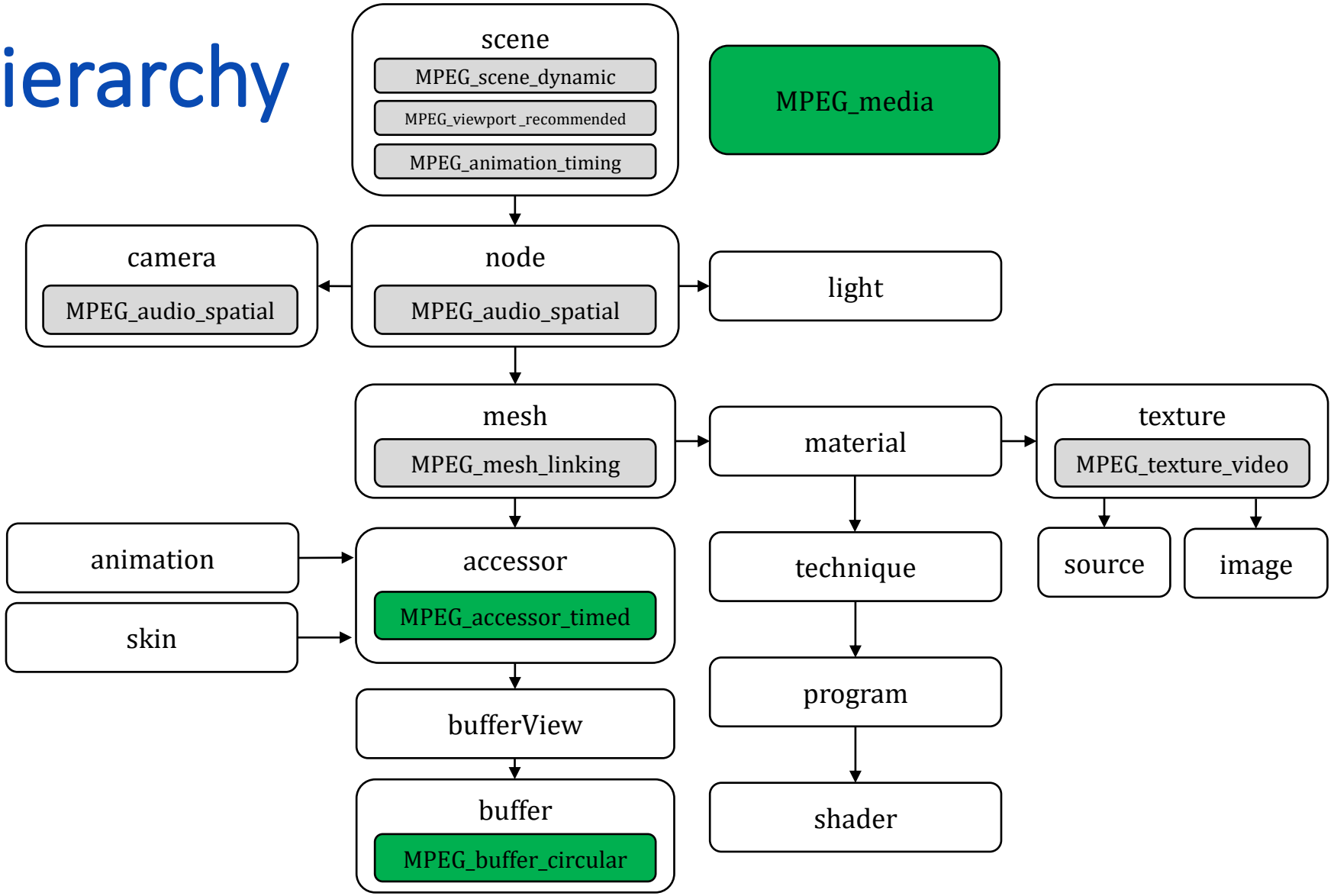
- The MAF instantiates and manages Media Pipelines
  - A media pipeline typically handles content of an attribute/component of an object/mesh
  - It produces content in the format indicated by the glTF file
  - The formatted frame is then pushed into the circular buffer
- Media Pipelines are highly optimized and customized for the type and format of media that is being fetched
- Media Pipeline maintains sync information (time and space) and passes that information as buffer metadata



MPEG-I Scene Description

Node Hierarchy

Core Extensions



External Media References

MPEG_media extension

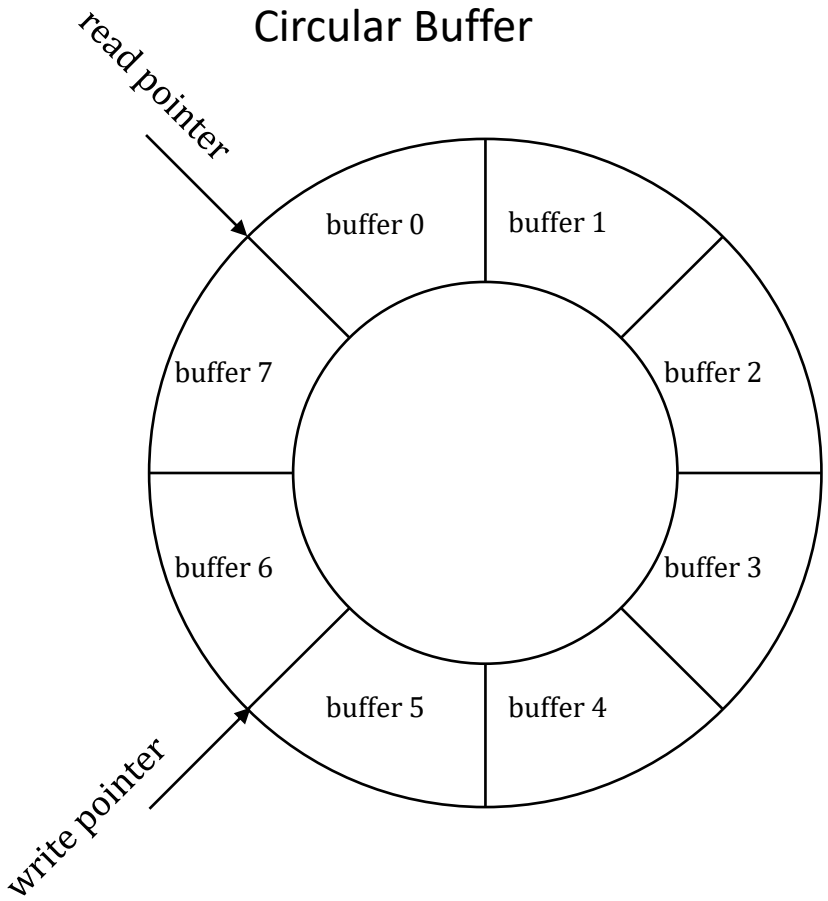
- Top-level extension to glTF 2.0
- Allows referencing all types of media
  - Timed and non-timed
  - Compressed and non-compressed
  - MPEG and non-MPEG
- It supports different types of delivery
  - DASH & CMAF
  - WebRTC
  - HLS & CMAF
  - Local Storage (ISO BMFF, MP4)
- Orthogonal Functions: encryption,etc.
- This extension decouples Media Access Function from Presentation Engine in the Scene Graph

```
{
 "extensions": {
   "MPEG_media": {
     media: [
      {
       "name": "source 0",
       "renderingRate": 25.0,
       "timeOffset": 0.0,
       "autoplay": "true",
       "loop": "true",
       "alternatives": [
         {
          "mimeType": "application/dash+xml",
          "uri":"https://www.foo.com/manifest.mpd",
          "tracks": [
            {
              "track": "#track=1"
            }
           ]
         }
       ]
      }
     ]
   }
  }
}
```

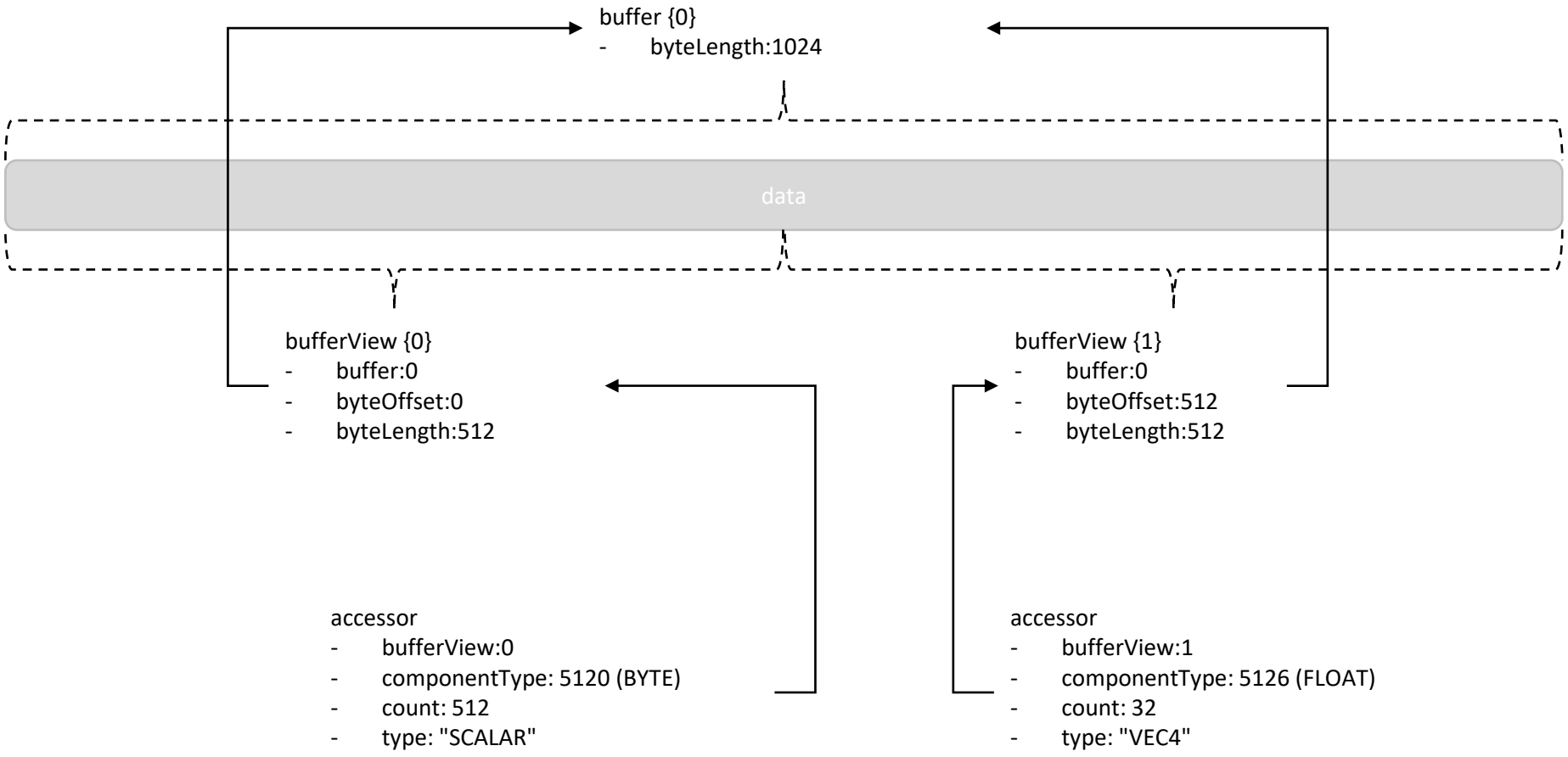
Describing how to access data

MPEG_accessor_timed and MPEG_buffer_circular extensions

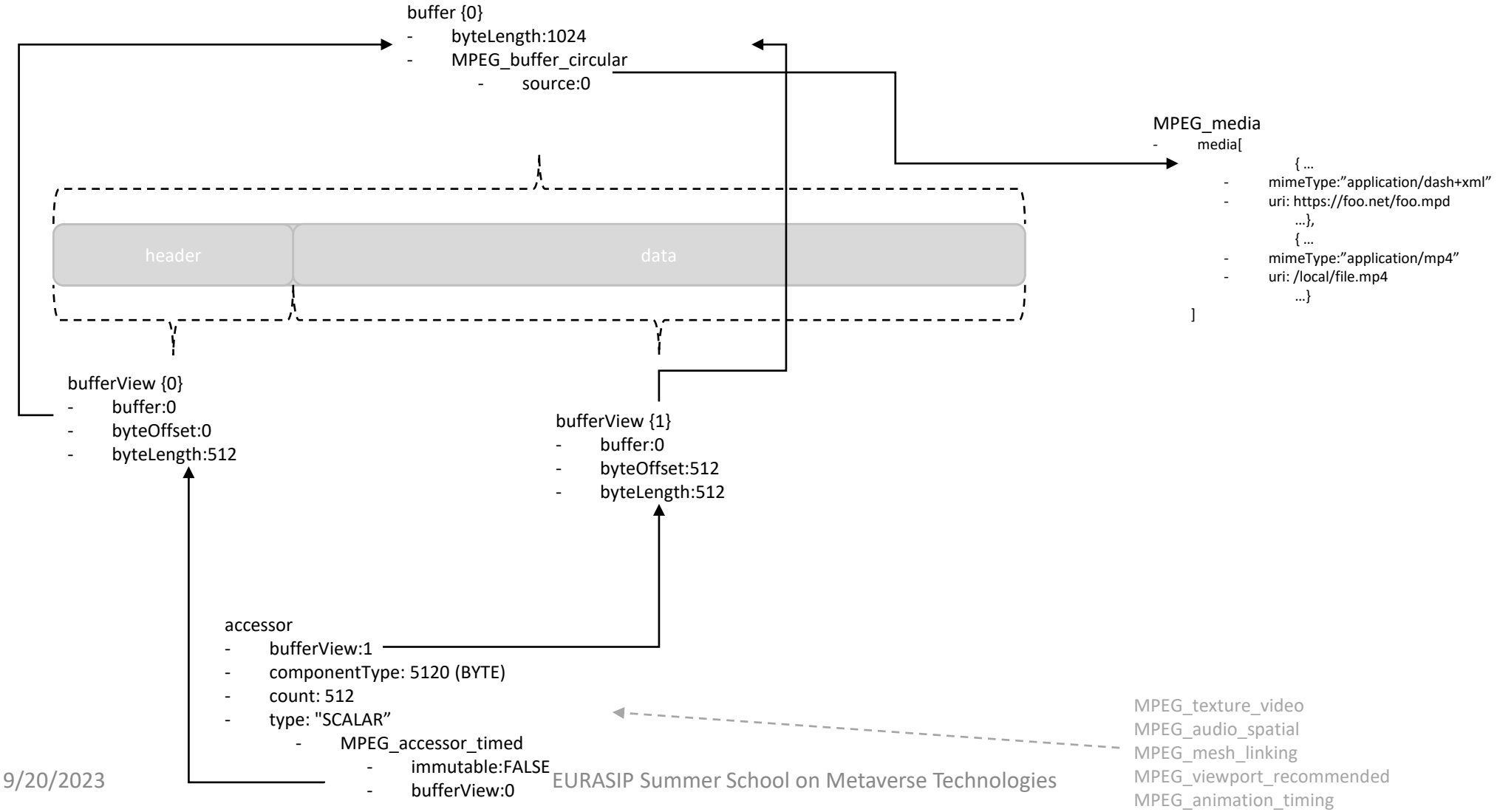
- glTF accesses data through accessors
  - They define the components of the data and their data types (e.g. a VEC3 of floats)
  - Semantics are provided by the referencing attribute/property (e.g. position)
  - The accessor points into a bufferView, which defines how the data is packed in the referenced buffer
  - No support for timed data
- MPEG_accessor_timed
  - Extension to accessor
  - Used to access all types of dynamic and timed media (audio, visual, volumetric, …)
  - Backwards compatible: in case of no support, fallback to static data
- MPEG_buffer_circular
  - Extension to buffer
  - Dynamic variable-size swap chain buffer for exchange of media data for rendering
  - Acts as the interface between the Presentation Engine and MAF. All requested data through MAF API is delivered through a Buffer or Circular Buffer
  - Header is used to propagate metadata such as timestamps
  - Circular Buffer references MPEG_media



Static Buffer Usage



Dynamic Buffer Usage



Timed accessor header infromation

- Mutable Information in Buffer View and Accessor
- Accessor information that may change over time
  - componentType
  - bufferView
  - type
  - normalized
  - byteOffset
  - count
  - max
  - min
- bufferView information that may change over time
  - bufferViewByteOffset
  - bufferViewByteLength
  - bufferViewByteStride

Video Textures

MPEG_texture_video extension

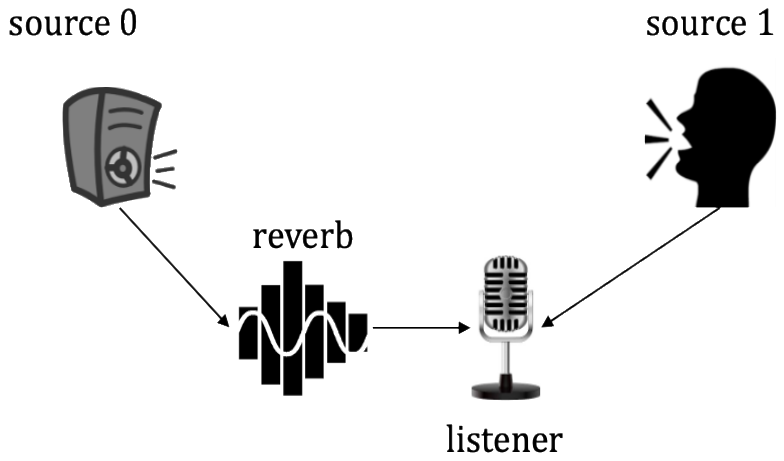
- Materials in the scene may make use of textures
- A texture in glTF 2.0 only supports references to images of format JPEG or PNG
- MPEG_texture_video adds support for dynamic textures such as atlases
  - Point into a timed accessor instead of an image
  - Keep the image pointer as fallback
  - To support dynamic atlases, texture coordinates themselves are dynamic and fed through a timed accessor as well

```
{
  "textures": [
    {
      "sampler": 0,
      "source": 1,
      "extensions": {
        "MPEG_texture_video": {
          "accessor": 2,
          "width": 2048,
          "height": 2048,
          “format": "RGB"
        }
      }
    }
  ]
}
```

Spatial Audio

MPEG_audio_spatial extension

- glTF has no support for audio
- The MPEG_audio_spatial extension:
  - Audio Sources can be coupled to visual nodes to share the same transformations
  - Supports 3 types of nodes:
    - Audio Source: emits audio signals. Simple mono and HOA sources are supported
    - Audio Effect: a reverb zone effect is currently supported
    - Audio Listener: provides the position of the listener
  - The Audio Listener may be linked to the scene camera to allow for an immersive spatial experience. The listener will move together with the camera.
  - Actual rendering is not defined.
    - It is up to the Audio Rendering Engine to convert the signals that are received at the audio listener into a format that matches the actual speaker setup.
    - For example, binauralization is done for users wearing an HMD.



Other extensions

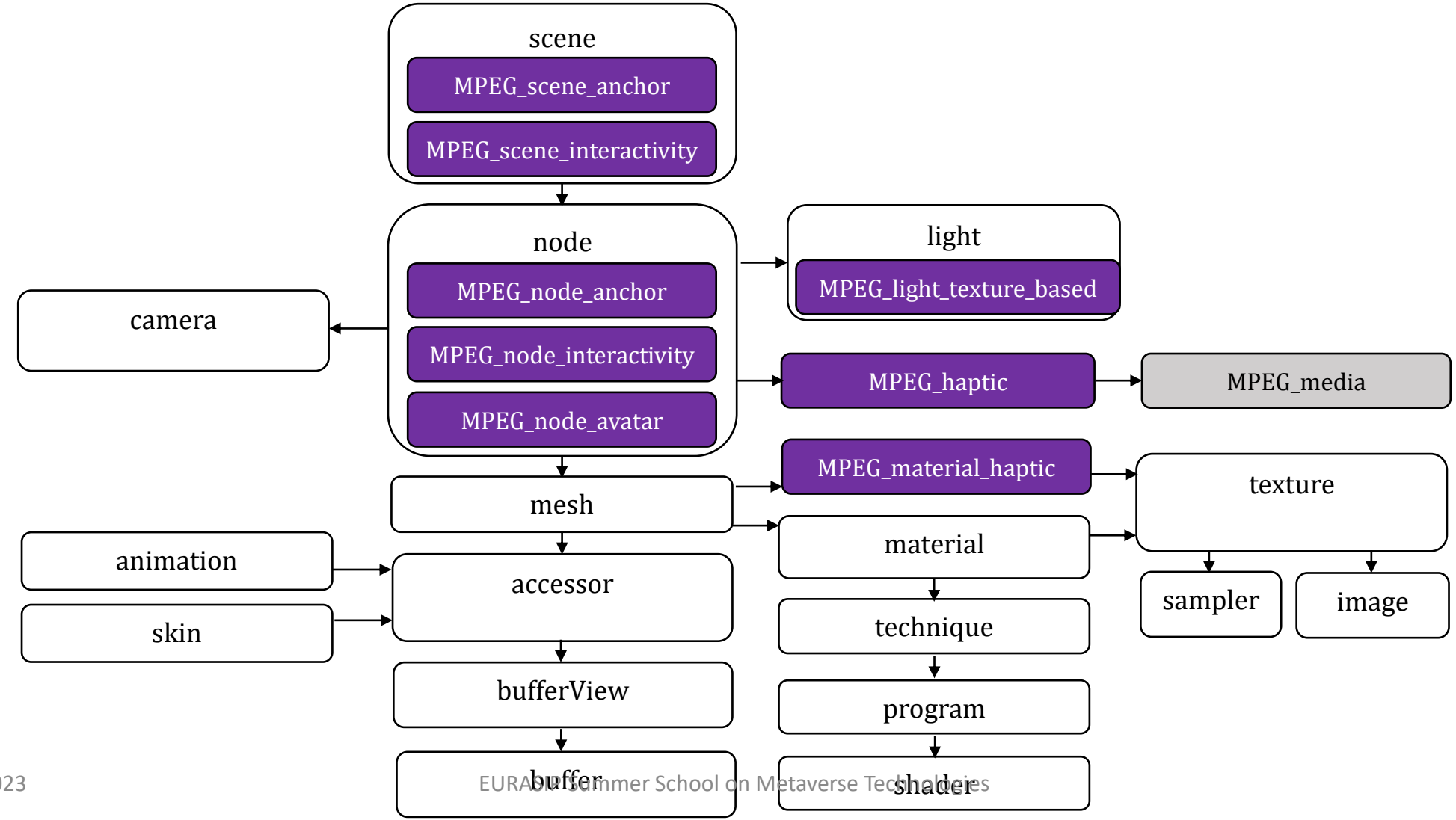
- MPEG_scene_dynamic
  - Provides the possibility to indicate that the scene description document will be updated
  - Updates are provided through JSON patch protocol
  - Patch sample is an atomic update operation (all patch operations part of one transaction)
  - Consistency/Validity of scene after application of a patch is the responsibility of the author
- MPEG_viewport_recommended
  - provides dynamically changing information which includes translation and rotation of the node which includes the camera object, as well as the intrinsic camera parameter of the camera object.
- MPEG_animation_timing
  - provides alignment between MPEG media timelines and animation timeline defined by glTF 2.0

March 1, 2023: Khronos adds MPEG-I Scene Description Extensions to glTF2.0

https://github.com/KhronosGroup/glTF/blob/main/extensions/README.md

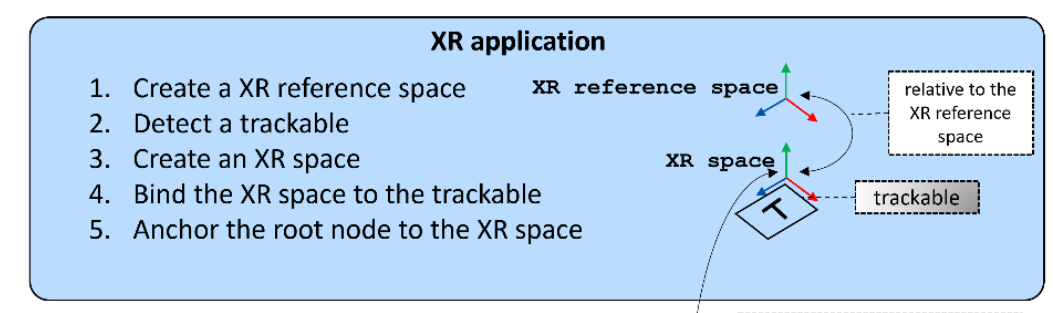
MPEG-I Scene Description Phase 2 Extensions

New Phase 2 Extensions



AR Anchoring

- Extension to allow anchoring a scene or a root node to a trackable
- Trackables can be reference trackables or application-defined trackables



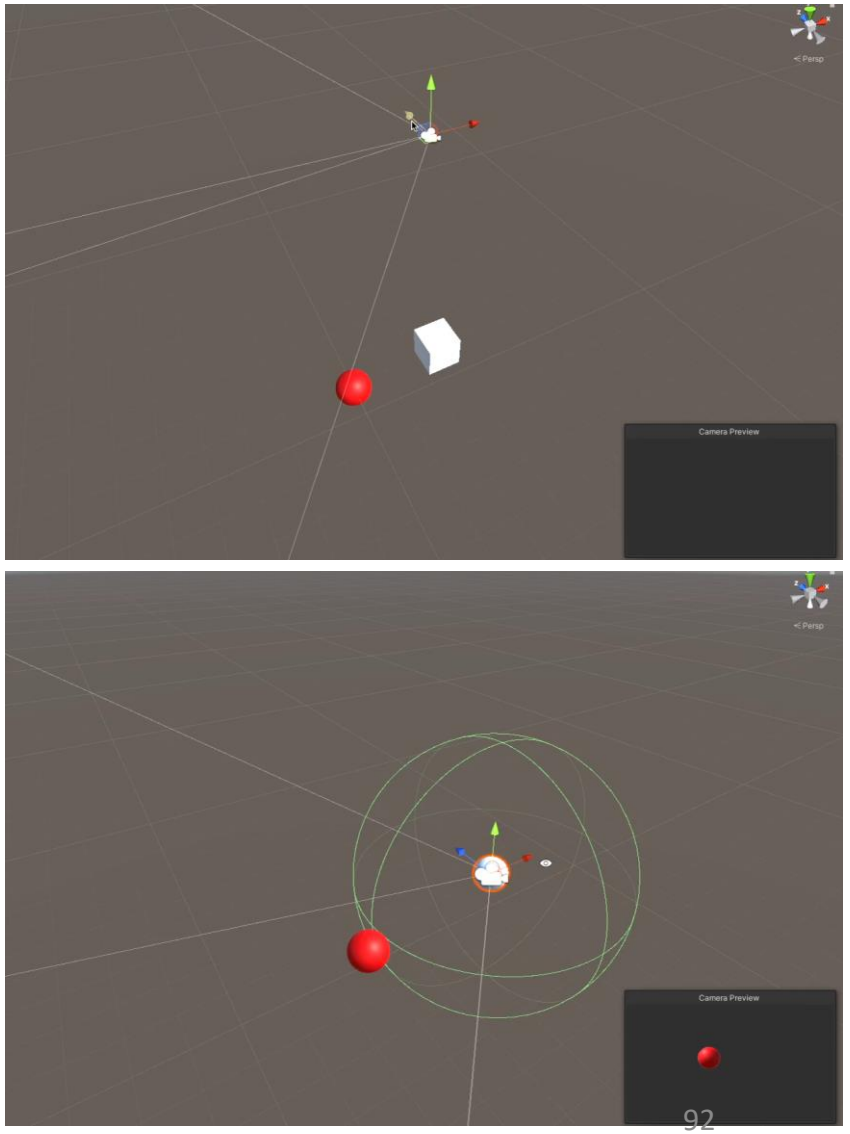
Structure	local transform	global transform
World space	RS	RS
Trackable	TS	RS*TS
node#1 (root)	R	RS*TS*R
+- node#2	A	RS*TS*R*A
+- node#4	C	RS*TS*R*A*C
+- node#5	D	RS*TS*R*A*D
+- node#3	B	RS*TS*R*B

Pseudo glTF2.0

```
"scenes" : [
        {
         "nodes" : [0]
        }],
    "nodes" : [ . . .
        {
            "name" : "NODE#1",
            "children" : [1,2],
            "matrix" : "#R"
        }, {
            "name" : "NODE#2",
            "children" : [3,4],
            "matrix" : "#A"
        }, {
            "name" : "NODE#3",
            "matrix" : "#B"
        }, {
            "name" : "NODE#4",
            "matrix" : "#C"
        }, {
            "name" : "NODE#5",
            "matrix" : "#D"
        }
    ]
```

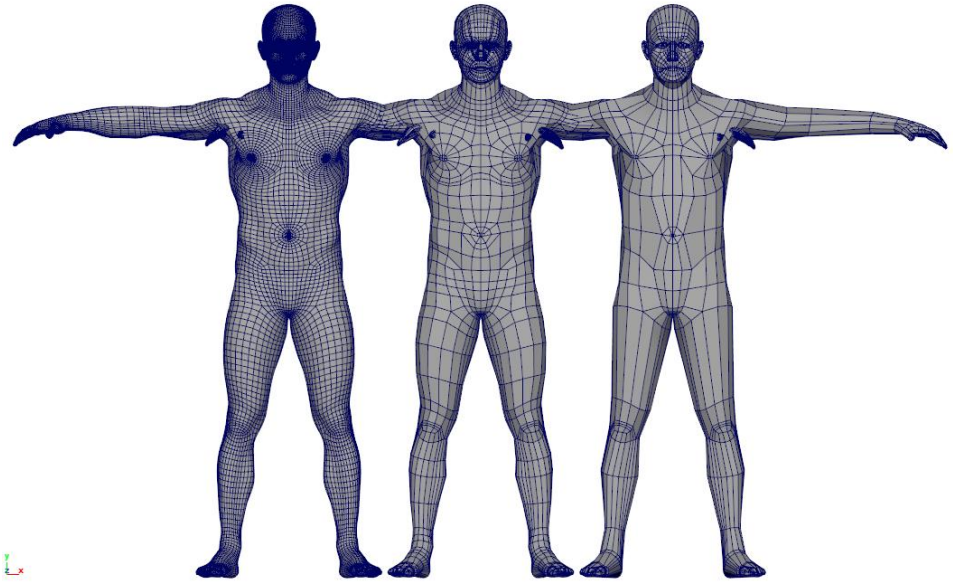
Interactivity

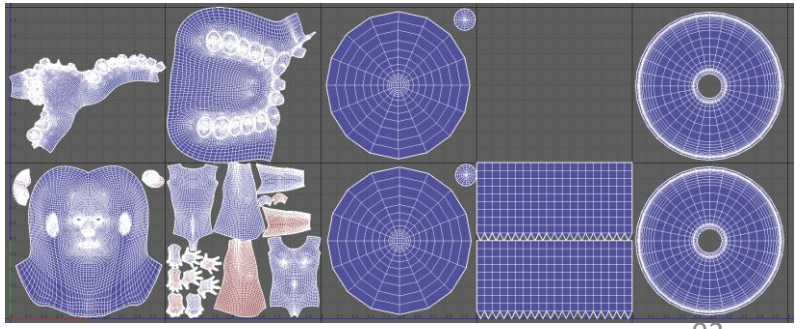
- Simple Trigger-Action interactivity
- Triggers: collision, proximity, user input, visibility
- Actions: activate, transform, block, animation, media, manipulate, set material, set haptic



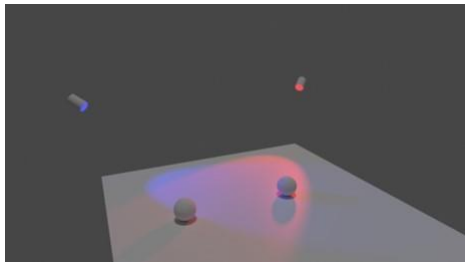
Avatars

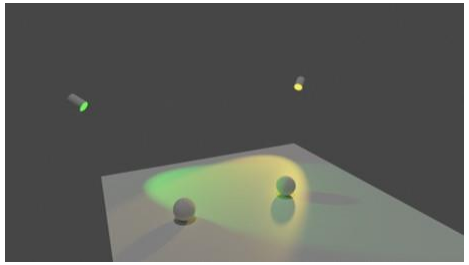
- Reference avatar with UV maps and Blendshapes
- Extension to signal nodes that carry Avatars and their breakdown
- Interactivity triggers can be associated with certain segments of the Avatar, e.g. hand
- Avatars may result from real-time animation streams



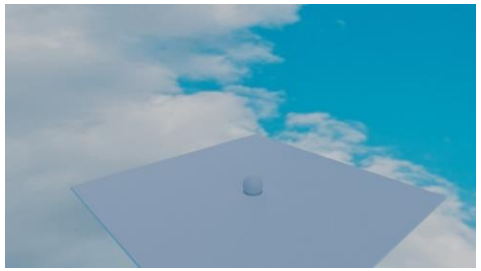


Dynamic Lights

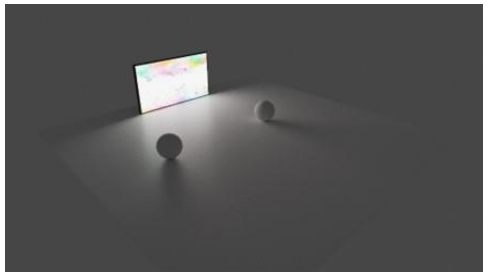


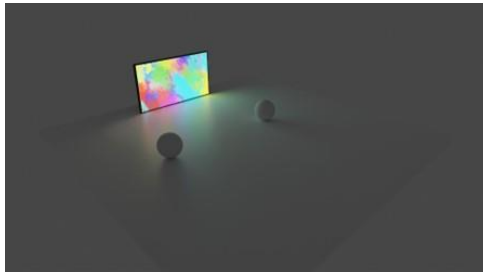


time-evolving
punctual light



time-evolving
environment light



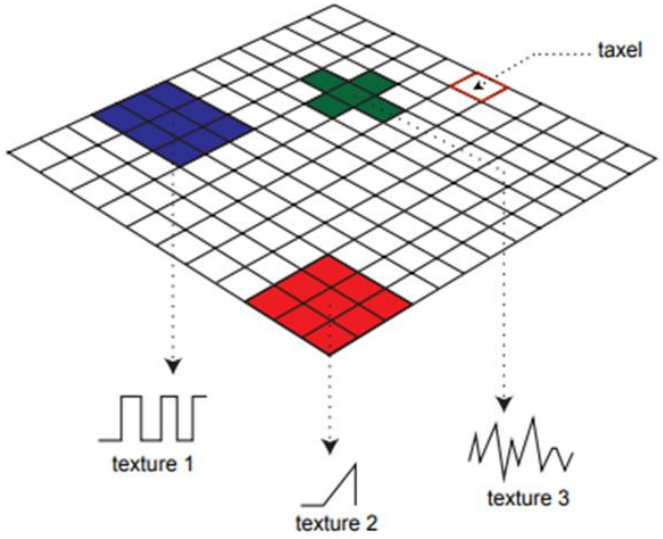


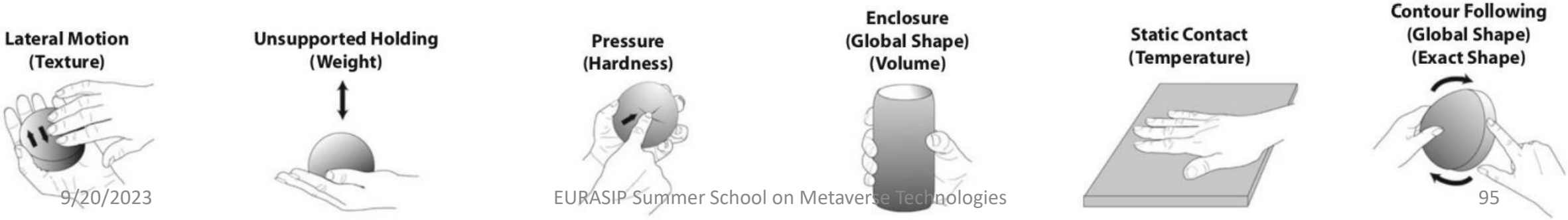
time-evolving
area light

Coherent real/virtual
lighting for AR applications

Haptics

- Extension to define Haptics materials
- Integrates with Interactivity
- Supports: stiffness, friction, vibrotactile, temperature, vibration, and custom haptic maps





References

- MPEG Vendor Extensions
  - https://github.com/KhronosGroup/glTF/tree/main/extensions/2.0/Vendor
- GitHub Repo for collecting issues
  - https://github.com/MPEGGroup/glTF
- MPEG-I Scene Description Whitepaper
  - https://www.mpeg.org/wp-content/uploads/mpeg_meetings/140_Mainz/w22138.zip
- Khronos Meetup:
  - https://www.khronos.org/events/gltf-meetup-July2023
- Open Source Reference Tools:
  - https://www.5g-mag.com/reference-tools

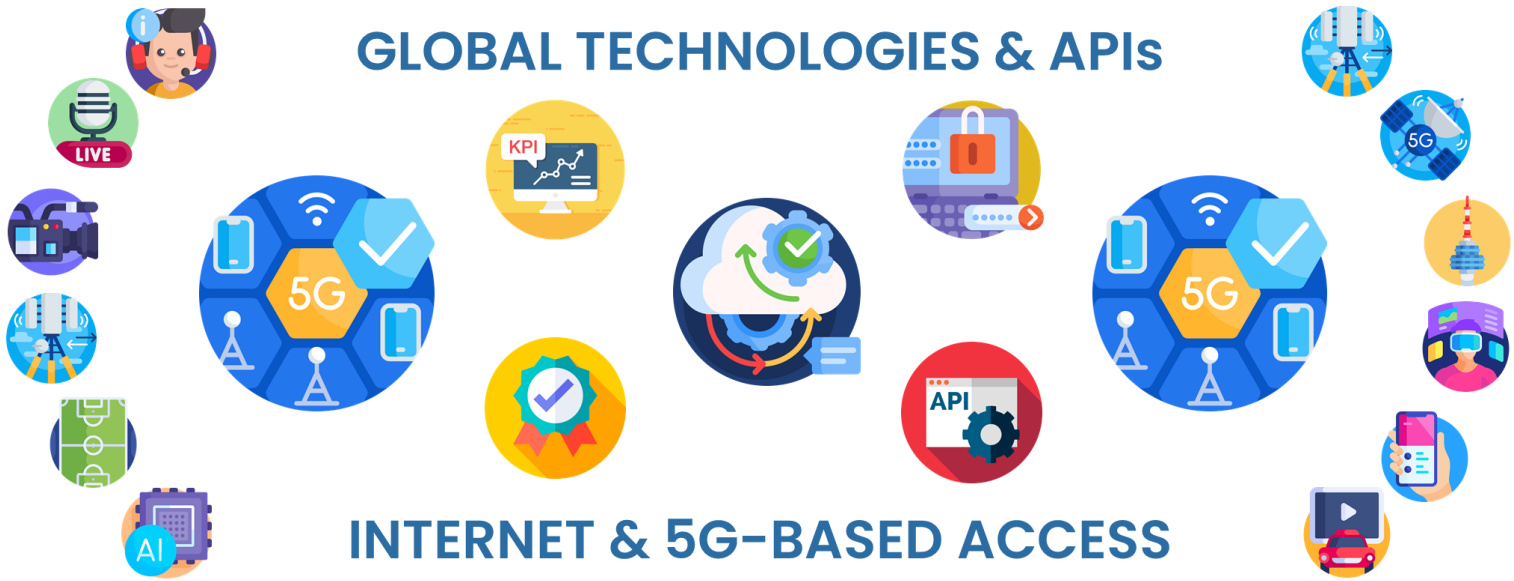
5G-MAG: Fostering media connectivity ecosystems

5G-MAG Association

An update on 5G-MAG activities - September 2023

www.5g-mag.com

5G-MAG is a non-for-profit cross-industry association



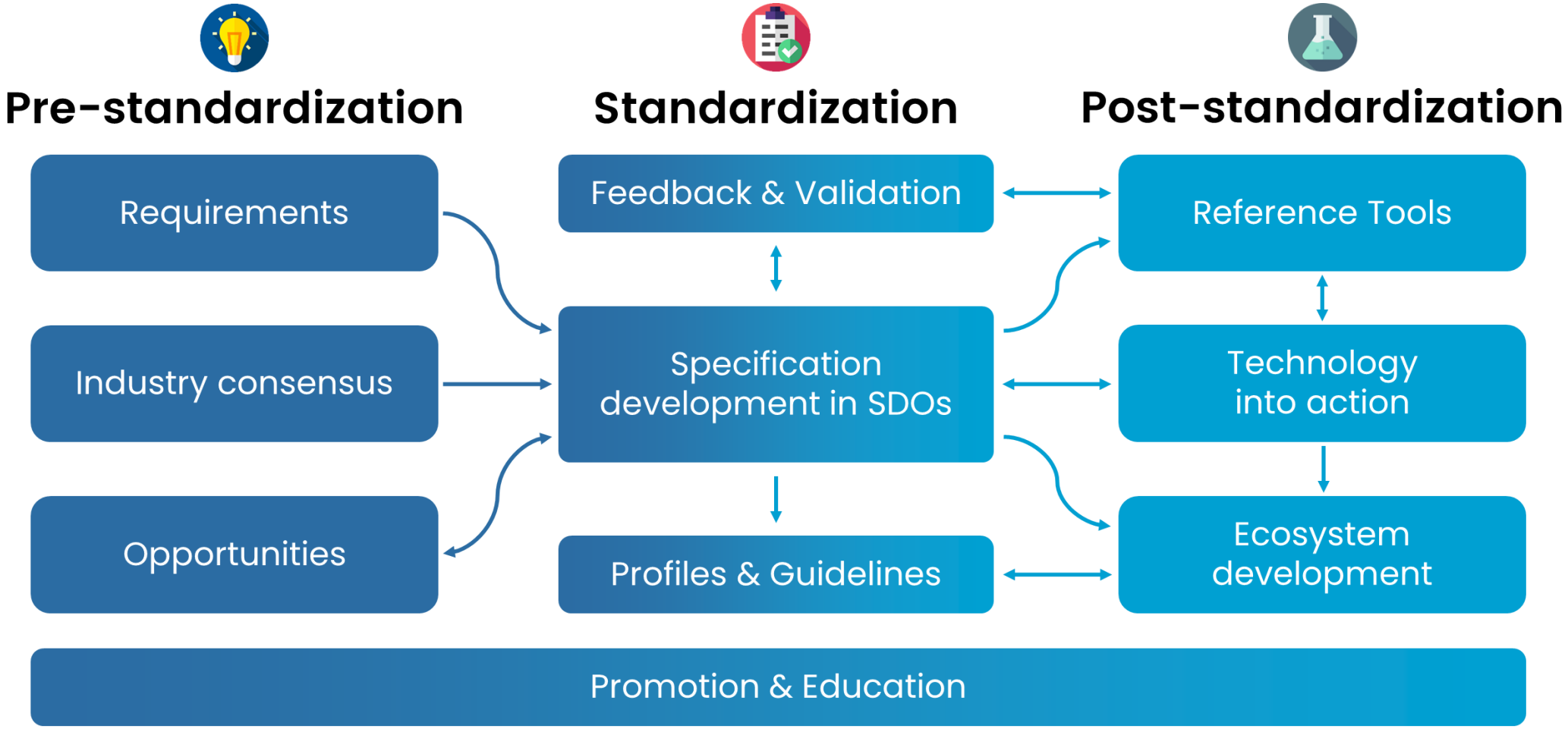
We leverage global **Internet**, **5G-based access** & **APIs** for **multimedia** applications

We shape **technology** through **engagement** and **collaboration** with stakeholders

We build **open-source reference tools** for testing and products

We promote technology **uptake** through **pragmatic** work

What we do

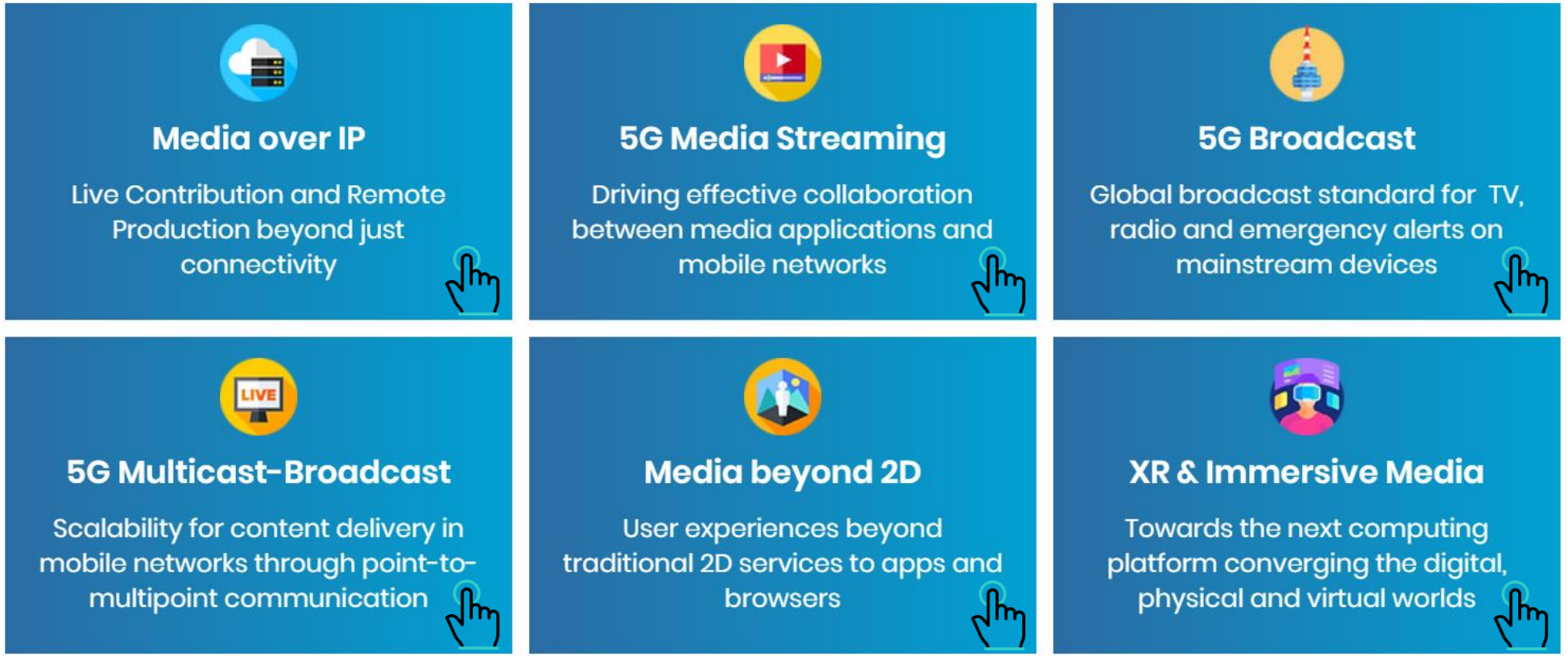


Activity Hub	Tech	Developer Space	Publications	Academy
hub.5g-mag.com	tech.5g-mag.com	developer.5g-mag.com	pub.5g-mag.com	academy.5g-mag.com

Activity Hub
hub.5g-mag.com

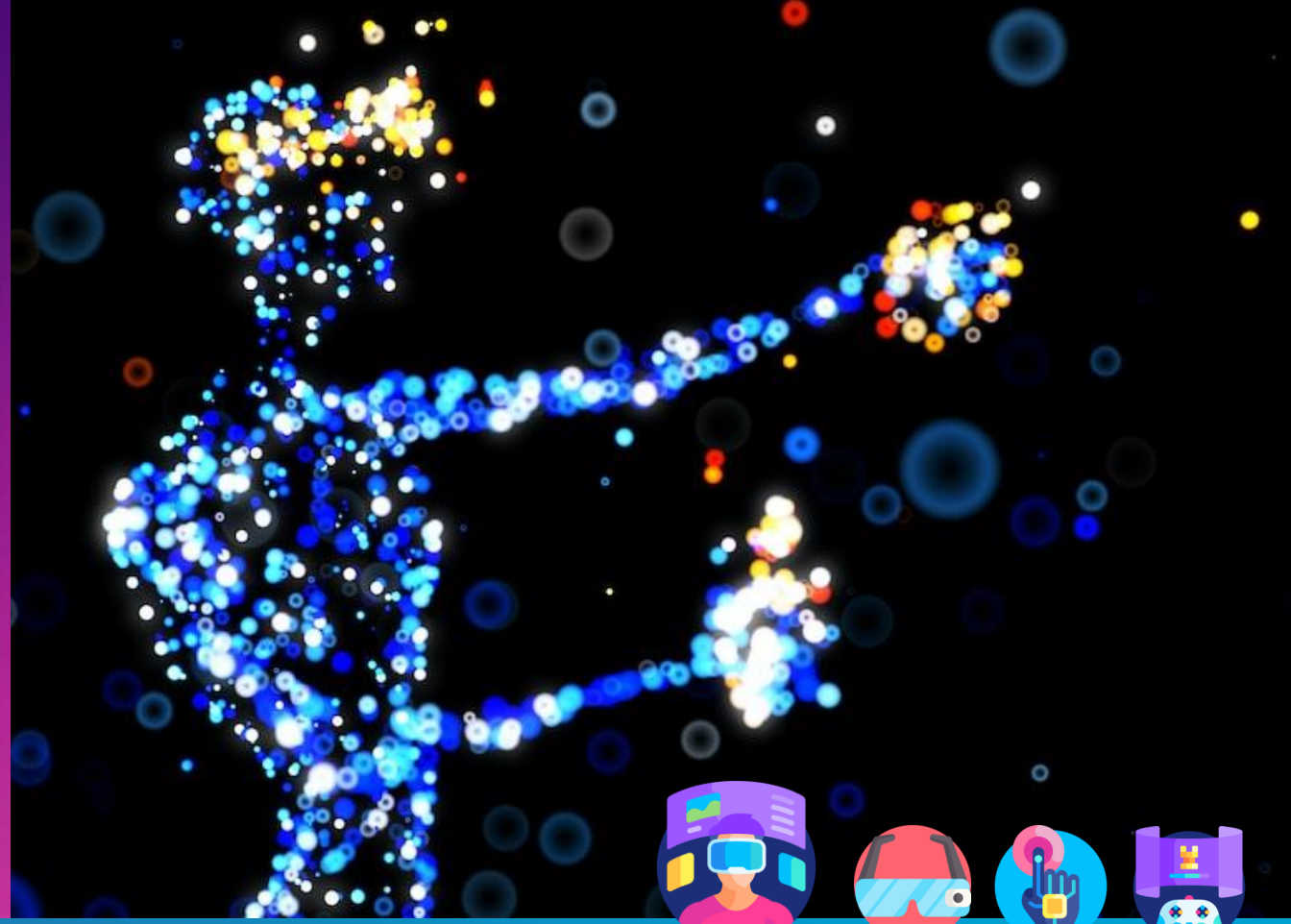
AREAS OF WORK

WORK ITEMS



Our current Areas of Work and Topics

Member- and contribution-driven work
hub.5g-mag.com



XR and Immersive Media

Learn more about the work at www.5g-mag.com/immersive

Key takeaways about the scope of the work

XR TOWARDS THE NEXT COMPUTING PLATFORM
Extended Reality (XR) embracing Augmented Reality (AR), Virtual Reality (VR), Mixed Reality (MR), and their underlying technologies

UBIQUITY AND CONVERGENCE FOR ENHANCED USER EXPERIENCES
Working towards the convergence of the digital, physical and virtual worlds building richer XR user experiences

TAILORING TECHNOLOGY AND NETWORKS TO MULTIMEDIA REQUIREMENTS
Handling of uplink XR traffic, managing traffic latency and power consumptions, enhanced QoS based on traffic payloads, mobility management,...

5G MAG REFERENCE < TOOLS />

Developer Space

developer.5g-mag.com

Community of Developers

Open community sponsored by 5G-MAG

Reference Implementations

for validation, testing, experimentation

Feedback to standards

Learning by doing and improving specifications

IPR-friendly License Model

Protecting IPR towards demonstrations & products

Open-source Reference Tools for Multimedia Applications

5G Media Streaming

Architectures, implementation details and repositories

Multimedia delivery protocols

Architectures, implementation details and repositories

LTE-based 5G Broadcast

Architectures, implementation details and repositories

5G Multicast-Broadcast Services

Architectures, implementation details and repositories

XR and Immersive Media

Architectures, implementation details and repositories

Machine Learning & Artificial Intelligence

Architectures, implementation details and repositories

Open-source Reference Tools for Multimedia Applications

Open Source as a Promotion Tool

Content Playback

- Unity and Unreal Engine 5 are widely used for the creation of 3D experiences
- Internal project ongoing to develop a Metaverse Player in Unity
- Player is able to load at runtime a 3D scene and render it to create an immersive experience
- Open sourcing the project will make this format accessible to the developer community and raise awareness about it

Content Creation

- Blender is an open source and widely used 3D authoring tool with native support for glTF
- Extend Blender for authoring Metaverse 3D scenes
- Open-source project to close the loop on content creation/consumption
- Enable developers to create content and ship players that can consume it
  - Cutting down on the effort to create immersive experiences for the Metaverse
  - It will also create a large base of Metaverse content

Participate

slack

- **Discussions** around development of features and resolving issues
- Dedicated channels for each project
- Join: http://tinyurl.com/join5gmagslack

Calls

- Public **Calls**
  - Fridays – with three-week cadence: 13:00 - 14:30 CEST
- Internal
  - Fridays - every other week: 13:00 – 14:30 CEST

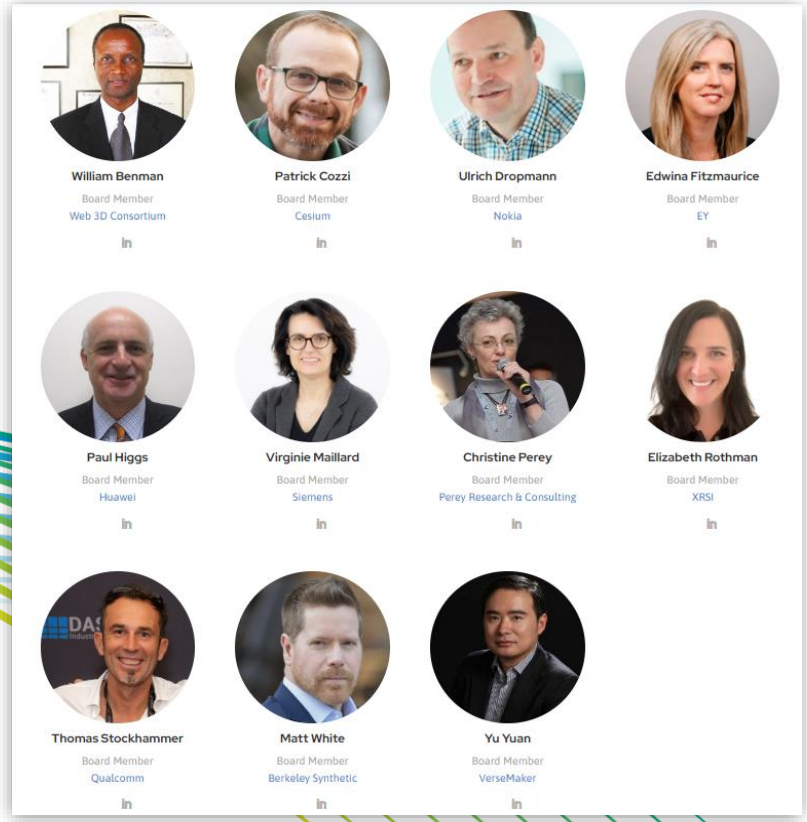
Groups

- **Announcements** of upcoming meetings, new release candidates and new releases
- http://tinyurl.com/join5gmaggroup

Overview Metaverse Standards Forum

Neil Trevett
NVIDIA | Vice President Developer Ecosystems
Metaverse Standards Forum and Khronos | President

Thomas Stockhammer
Qualcomm Incorporated | Senior Director Technical Standards
Metaverse Standards Forum Standards Register WG Chair, Board Member



The Vision

A Venue for
Cooperation between
Standards Organizations and Companies to
Foster the Development of
Interoperability Standards for an
Open and Inclusive Metaverse

Metaverse
STANDARDS FORUM™

The Metaverse Will be Built on Interoperability

Combining multiple disruptive technologies to work together (AI, GPU, XR, Web3)

Building bridges between applications to scale beyond a series of disconnected silos

Depends on Interoperability

Evolving a platform that is open and inclusive for all – an immersive evolution of the web

Pervasive metaverse interoperability will need a constellation of open standards …
… involving 100s of standards organizations

End 2021

KHRONOS GROUP

Khronos finds increasing interest in its standards for the metaverse…

… but discovers that there is nowhere to coordinate with other standards organizations

June 2022

Khronos funds launching the Forum in bootstrap mode to determine industry interest

A venue for cooperation between standards organizations and the wider industry

Straightforward participation agreements with Khronos to enable standardization cooperation and communication

End 2022

Metaverse STANDARDS FORUM™

The Forum grows to over 2400 Member organizations

Multiple Domain Working Groups working to improve interoperability one project at a time

Today

The Forum incorporates with unanimous agreement from its membership

Independent, self-funded, non-profit industry consortium

The Forum's mission is to create a wavefront of business opportunities through fostering interoperability 'brick-by-brick' on the road to the metaverse

The Metaverse Brings Together Diverse Technologies

The Metaverse combines the connectivity of the Web with the immersiveness of Spatial Computing

Combining multiple disruptive technologies

Advances in GPU-driven real-time photorealistic graphics and simulation
Scenes, avatars and objects

Decentralized Trust and Storage
ID and Reputation
Economic transactions
Persistence

Artificial Intelligence (AI) a.k.a. Machine Learning
Natural user interfaces
Semantic scene understanding
User generated 3D content

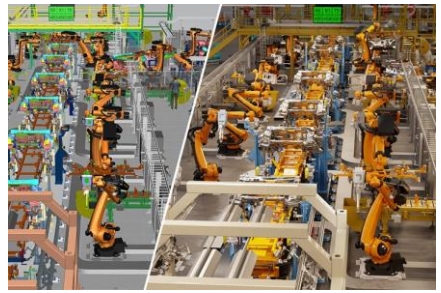
XR – Virtual Reality (VR) and Augmented Reality (AR)
VR for generated environments
AR to overlay the real world

Networking
Edge computing
5G, 6G, 10G

GPU-accelerated photorealistic rendering and simulation
E.g., Epic MetaHumans

Social gaming with end user-generated content
E.g., Roblox

Affordable and accessible XR Devices
E.g., Meta Quest Pro



Digital twins for modeling, monitoring and simulation
e.g., NVIDIA Omniverse

June 2022 - 37 Founding Organizations

Today - Over 2400 Members and Counting

Wide diversity of organizations, including...

SDOs
Khronos, W3C, Open Geospatial Consortium, IEEE, OMI, ASWF, Spatial Web Foundation, VRM Consortium, XRSI, OMG, Open AR Cloud, OMA3 …

Platforms
Meta, Microsoft, Sony, Google, Baidu, Huawei, General Motors, RedHat, Siemens, Tencent, Mozilla, Paramount …

Tools and Engines
Epic, Unity, Adobe, Autodesk, Otoy, Maxon, Cesium, ESRI, Blackshark.ai, Croquet, Lamina1, Niantic, Ready Player Me, DGG, Manticore …

XR
HTC, Magic Leap, Nreal, Panasonic, Tobii, zSpace …

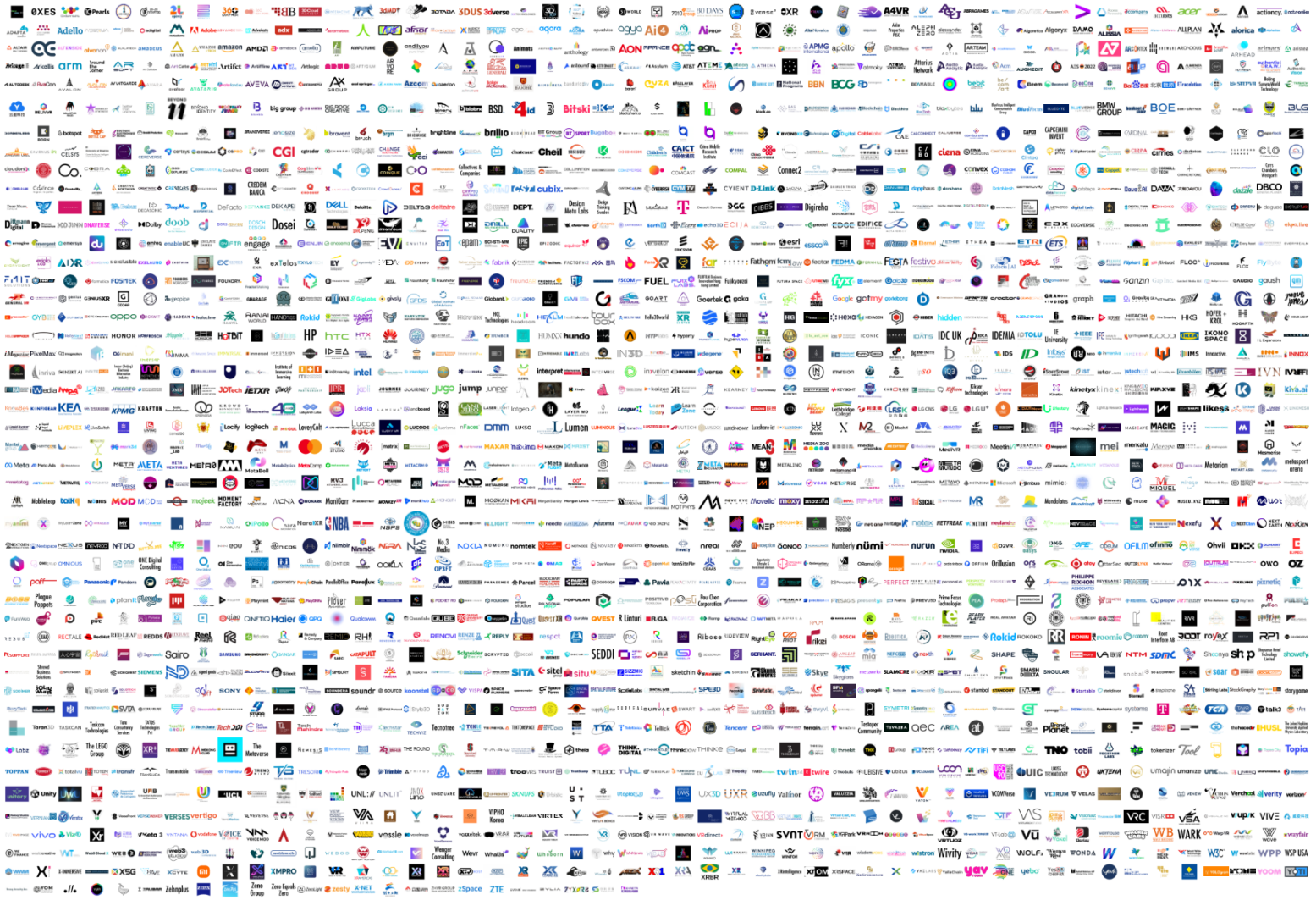
Hardware
NVIDIA, Intel, AMD, HP, Acer, Dell, Qualcomm, Samsung, Sony, MediaTek, Oppo, Lenovo, ZTE, LG …

Wireless and Networking
China Telecom, Deutsche Telekom, T-Mobile, Verizon, NTT, AT&T, Telefónica, Juniper, Comcast …

3D Commerce
Alibaba, Alvanon, Avataar, CLO, Browzwear, IKEA, VNTANA, Metaverse Fashion Council, Target, Wayfair …

Universities and Institutes
Stanford, John Hopkins, Yale (XRP), Queens University Belfast, University Salford, New York Institute Technology, APMG …

Advocacy
XRSI, AREA, XR Association, VRAR Association, XR Guild, Web3 Marketing Association, International Virtual Reality Healthcare, Swiss Institute for Disruptive Innovation, IOT Consortium, Metaverse Japan, RIAA …



Forum Domain Group Pipeline

Metaverse Standards Register
Publicly available database mapping the landscape of metaverse-relevant standardization activities

glTF/USD 3D Asset Interoperability (visuals, behaviors)
Cooperation between USD and glTF to increase synergy
and reduce duplication of effort, gaps, fragmentation and industry confusion

Real/Virtual World Integration (Digital twins, IOT)
Constructs to describe and integrate the physical world and created representations

Asset Management (web3, protection, digital rights)
Digital rights, protection, portability, access, availability

Network Requirements and Capabilities to Support Metaverse Applications
Industry requirements for seamlessly transitioning traffic on multiple wireline and wireless technologies for deploying metaverse applications at scale

Interoperable Avatars
Cross-platform avatars and characters for film, gaming, fashion and social platforms

Privacy, Cybersecurity & Identity
Recommendations for responsible innovation that mitigates human and societal harm from objective and subjective privacy risks – including cybersecurity and identity risk management

End-User Technical Troubleshooting
Enabling end-users to ensure reliable metaverse experiences

3D Web Interoperability
Enable the broadest possible interoperability of Metaverse Content using the Web

Digital Fashion/Wearables
Clothing (including layering), shoes, hats, accessories

Ownership and Identity

Accessibility

Best Practices for Living and Working in the Metaverse

Academia & Research

Ethical principles for the metaverse and their implementation

Industrial Metaverse

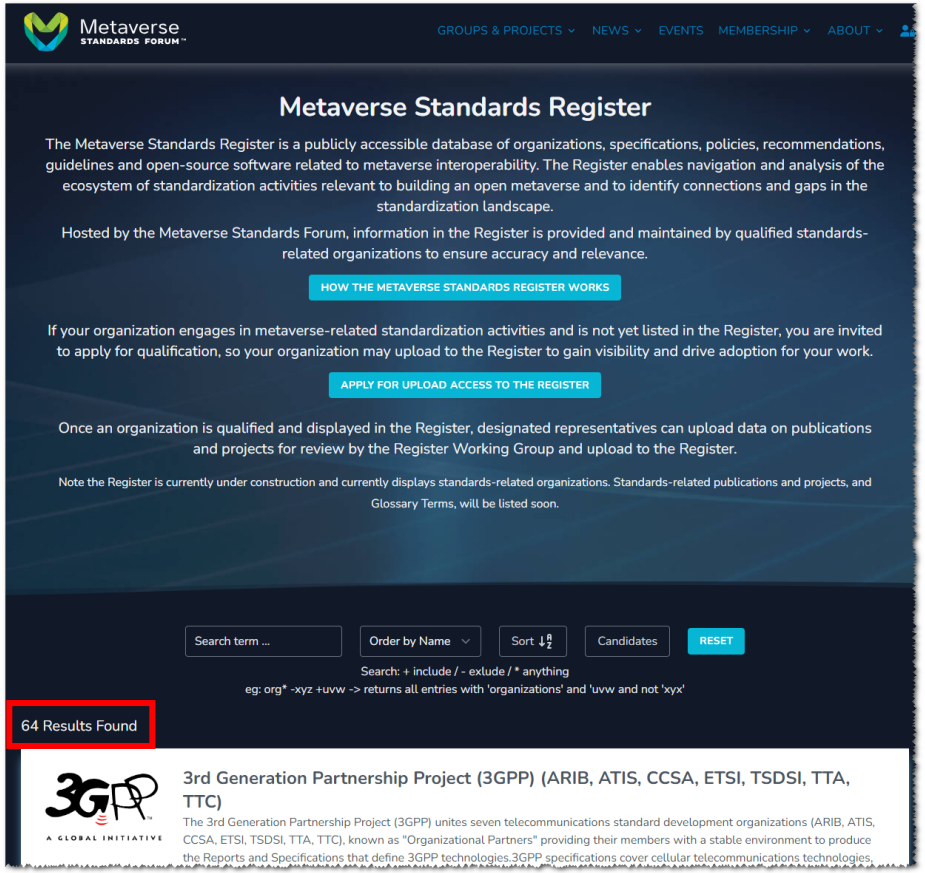
Metaverse STANDARDS FORUM™

https://metaverse-standards.org/domain-groups/

Key
Working Groups
Exploratory Groups
Exploratory Group Proposals

Metaverse Standards Register Launch

- Phase one launch of the Metaverse Standards Register
  - Searchable, sortable database of Metaverse-related standards organizations
  - How the Metaverse Standards Register Works
- Some organizations are pre-populated
  - To kick start engagement
- If your standards organization is missing or pre-populated, please enter your details ASAP!
  - Organization Application Form
  - Feedback to standards_registry-feedback@lists.metaverse-standards.org
- Register Working Group is now working on…
  - SDO outreach to populate information on specifications
  - Launch Phase 2 Register with specification database
  - Create searchable Glossary and Use Case databases
- Register Launch Blog is being drafted
  - To drive awareness and participation



USD glTF Interoperability Potential Testbed Project

- Goals
  - Confirm asset behaviours and attributes satisfy use cases
  - Test publishing and transmission pipeline
  - Exercise interoperable behaviours in multiple runtimes
- All tool engine and platform vendors invited to participate
  - Whether or not they are Forum Members
- Cooperative shared open-source and assets

AUTODESK
blackshark.ai
CESIUM
Adobe
MAXON A NEMETSCHEK COMPANY
otoy

Tools create assets using open standards

FORTNITE
ROBLOX
UNREAL ENGINE
Unity
three.js
babylon.js
PLAYCANVAS
GODOT Game engine
O3DE OPEN 3D ENGINE
NVIDIA OMNIVERSE

Run time engines ingest and process assets

USD

USD-based tools
Author assets and publish into glTF

DISTILL

glTF

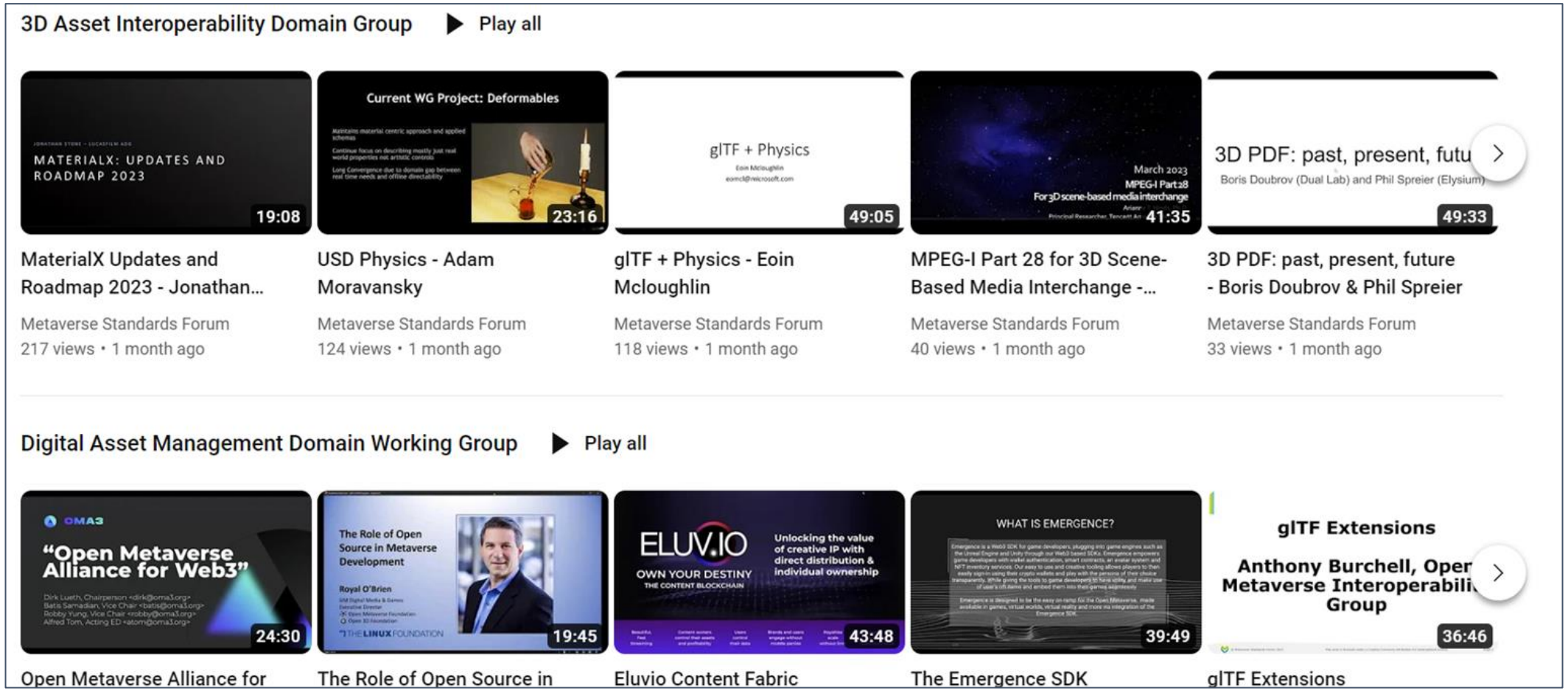
Web-based Configurator
Material variants

TRANSMIT

Runtime Demos
Open door, start engine
Drive course with physics simulation

Video Presentation Library

Invited speakers at Forum Domain Groups are posted to a growing public video library
https://www.youtube.com/@metaversestandardsforum



Call for Participation in Unique Cooperative Opportunity

Broad global participation in the Forum enables a unique opportunity for metaverse standards cooperation, coordination and leadership for Forum members to accelerate *their* organizations objectives
https://metaverse-standards.org/

Comprehensive, international gathering of industry requirements and expertise in Forum Working Groups

Any Forum member can propose, lead, contribute to, participate in, or monitor Domain Working Groups

Metaverse STANDARDS FORUM™

Wide visibility and adoption of Forum initiatives

Invitation to Collaborate

Summary & Next Steps

Join the community of open standards, innovation and development

The Metaverse is an open platform for everyone – contribute and experience

The Metaverse is are global and so are technologies, standards and reference implementations – contribute and use.

Qualcomm contributes, supports and drives open systems through technologies, standards and reference tools

Questions?

Connect with us

qualcomm.com/5g

qualcomm.com/news/onq

@QCOMResearch

.youtube.com/qualcomm

slideshare.net/qualcommwirelessevolution

Thank you

Snapdragon

Follow us on:

For more information, visit us at:

snapdragon.com & snapdragoninsiders.com