The Evolution of Augmented Reality Technology

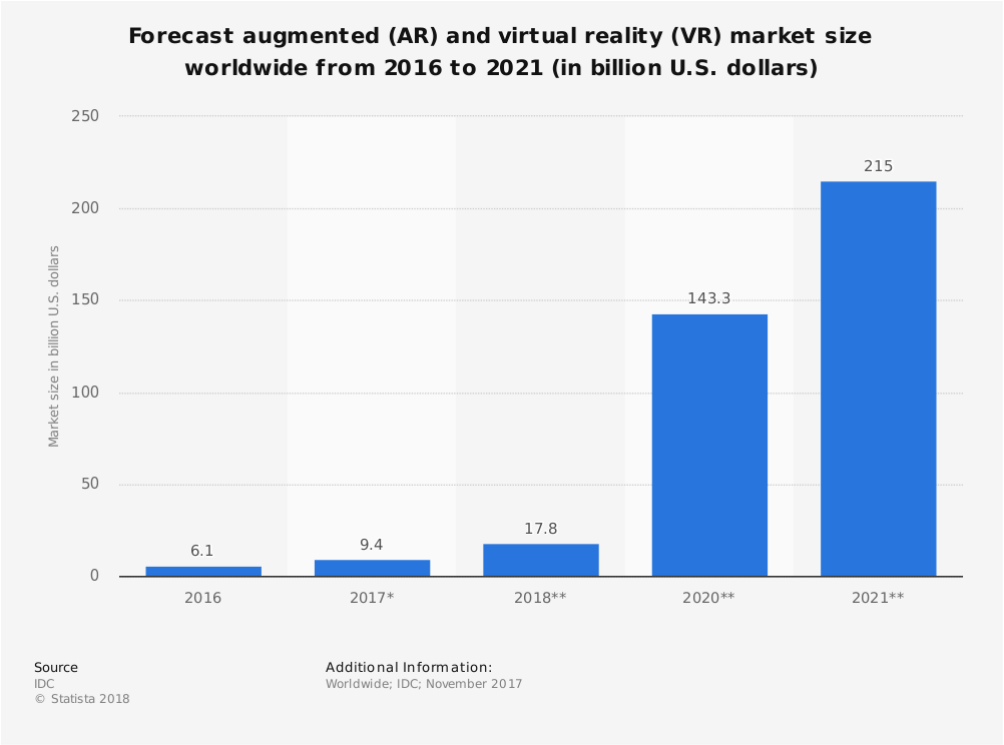
Thomas Stockhammer
Director Technical Standards
Qualcomm Technologies, Inc.
March 2018

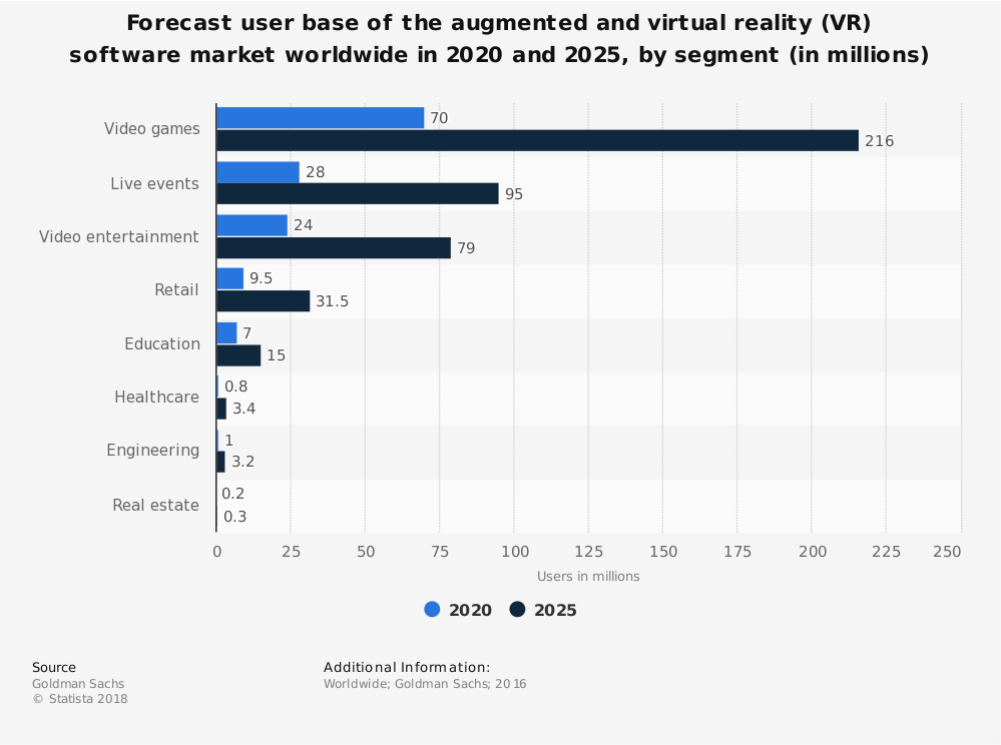
DVB World 2018

Agenda

1. Motivation and Introduction to Augmented Reality
2. Evolution of AR from today to the future
3. Challenges and Requirements for AR
4. Technical Enablers: Mobile Compute Platform
5. Initial Standardization Needs and Efforts

Motivation and Introduction to Augmented Reality





There seems to be a business case for AR

Augmented reality vs. virtual reality

Similar underlying technologies but distinct experiences

Virtual reality

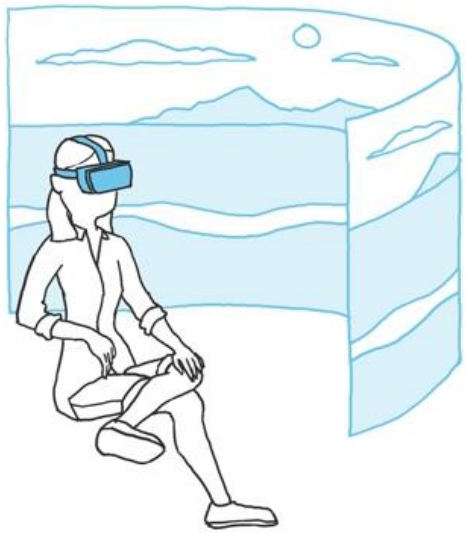
Simulates physical presence in real or imagined worlds, and enables the user to interact in that world

Augmented reality

Superimposes content over the real world such that the content appears to be part of the real-world scene

Evolution of user experience from VR to AR

Today



VR: Mostly 3-DOF, lower resolution videos & games

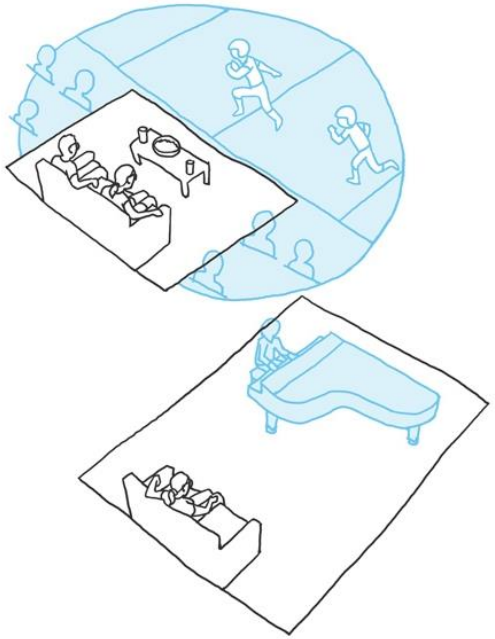
AR: Pokémon Go, Google Translate, Snapchat, and other rudimentary AR apps

Soon

VR: Ability to move around through live events, with better sense of “presence”

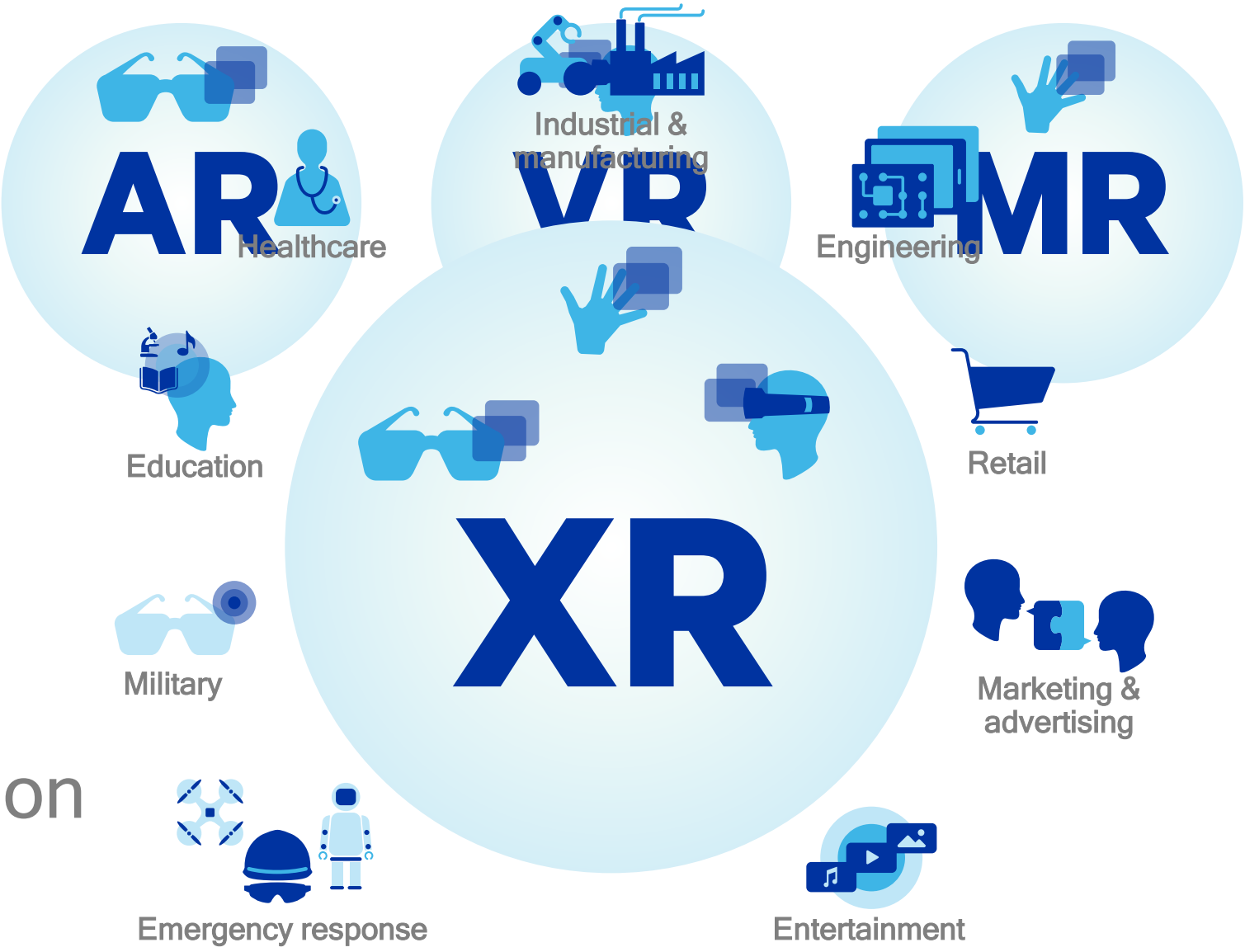
AR: Still rudimentary, yet more useful and immersive, streaming AR services, abled to be accessed on the go

The Future



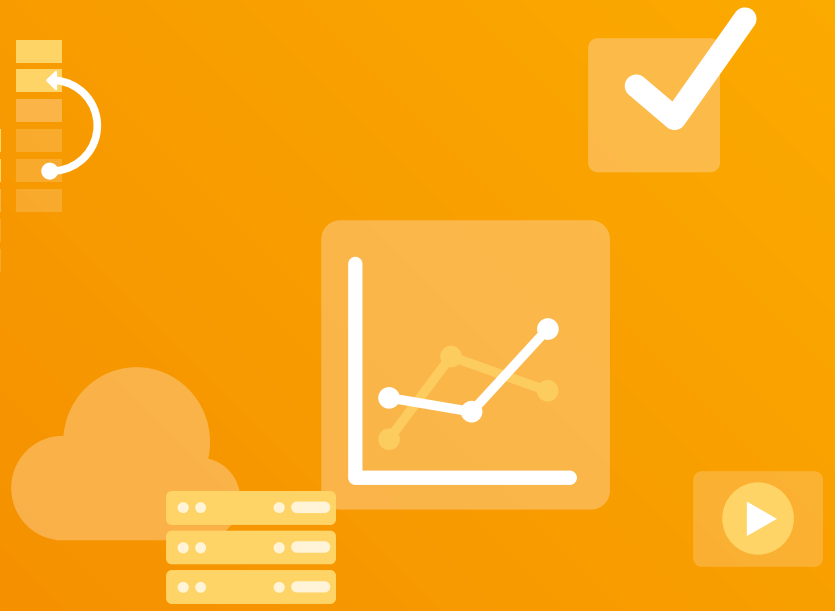
AR: Entire scenes, like entertainment events, can be accessed with your mobile AR device that are so realistic and interactive that they’ll be nearly indistinguishable from reality. VR becomes an occasionally used “mode” within AR

XR is the next mobile computing platform



We use the term AR in the presentation

Evolution of AR from today to the future



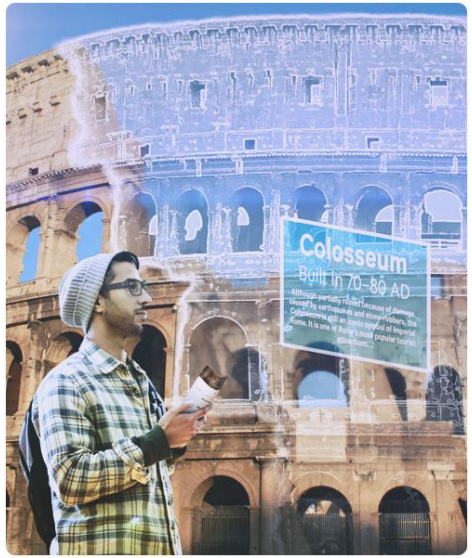
AR will serve a broad spectrum of roles in daily life

Applicable across ages, genders, and activities

Children Playing

Kids chasing virtual characters in more interactive & immersive games

Young Adults Exploring



A young man exploring Rome and seeing the Colosseum as originally built

Families Communicating

Families virtually brought together with life-like communication

Professionals Working

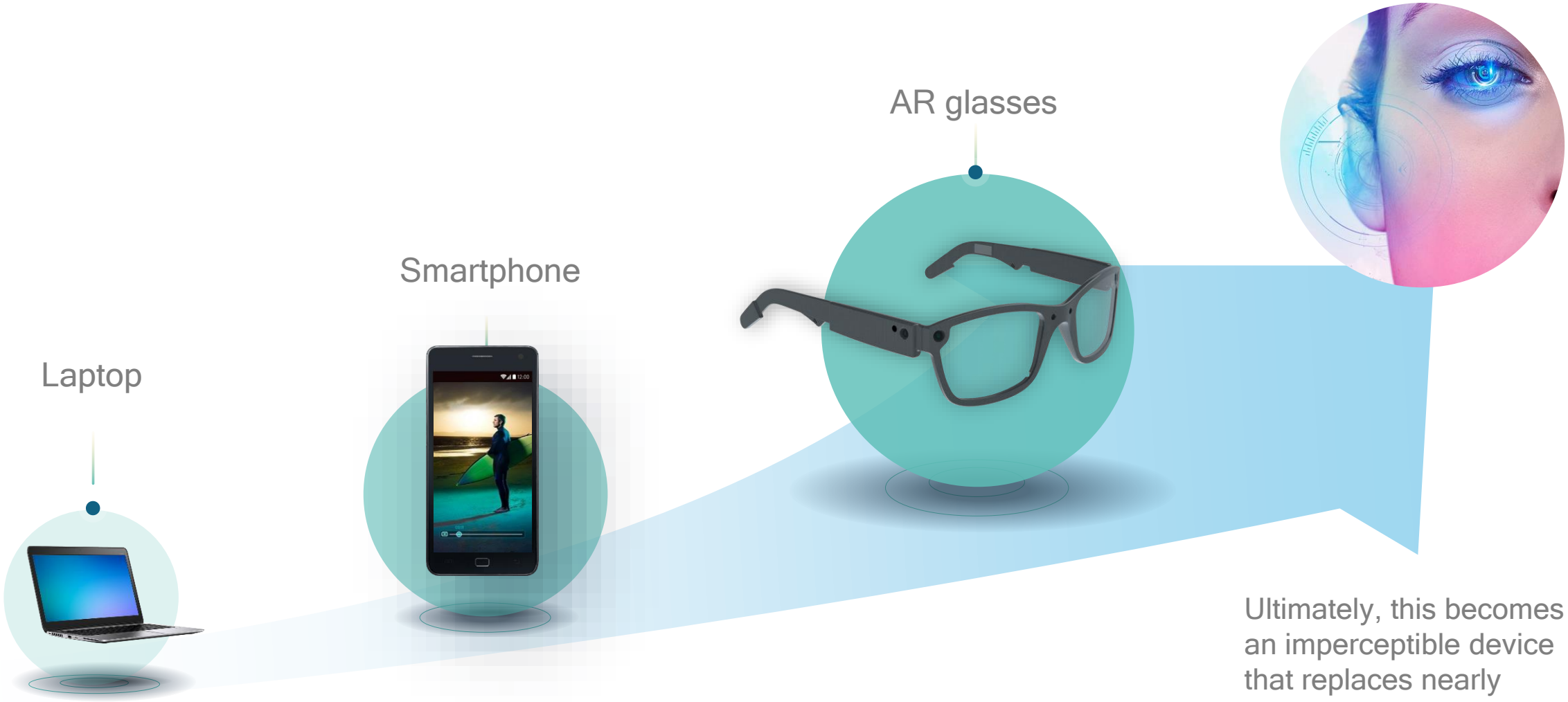
Architects collaborating on a shared design to improve efficiency

Fitness Enthusiasts Thriving

Group running with a virtual trainer to motivate them

AR is the next mobile computing platform

Nearly everything we've learned for smartphones will be used for AR



Ultimately, this becomes an imperceptible device that replaces nearly all others

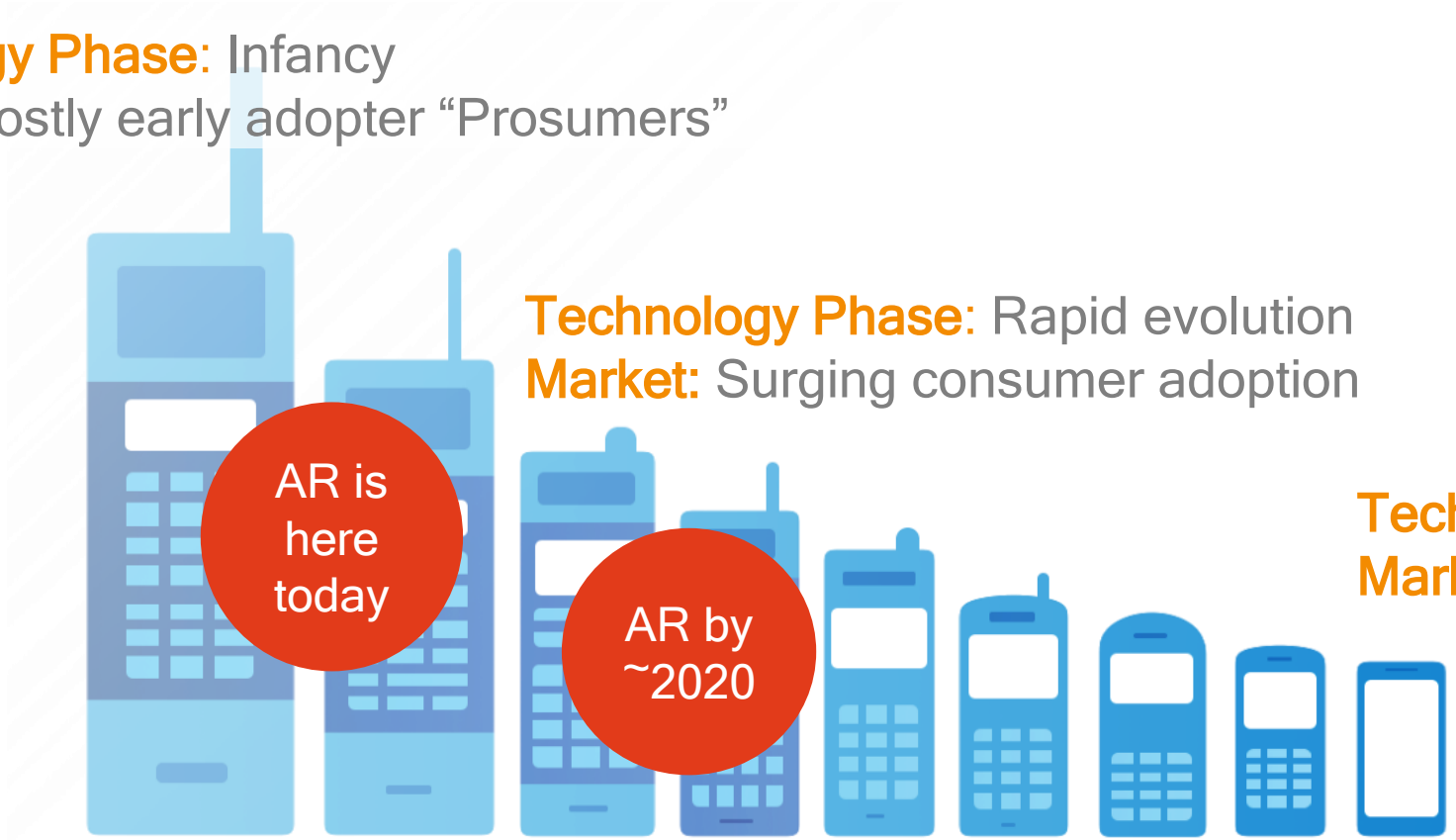
AR is here today, but it is still in its infancy

Analogy to cell phones: AR evolution will take years…opportunity will be immense

Technology Phase: Infancy
Market: Mostly early adopter "Prosumers"

Technology Phase: Rapid evolution
Market: Surging consumer adoption

Technology Phase: Maturity
Market: Worldwide, ubiquitous use



AR will follow a similar ~30 year cycle of sleeker designs, with tremendously increasing functionality

AR is here today, but it is still in its infancy

Currently shipping devices that use Qualcomm® Snapdragon™ processors

Today's Smartphones and Phablets

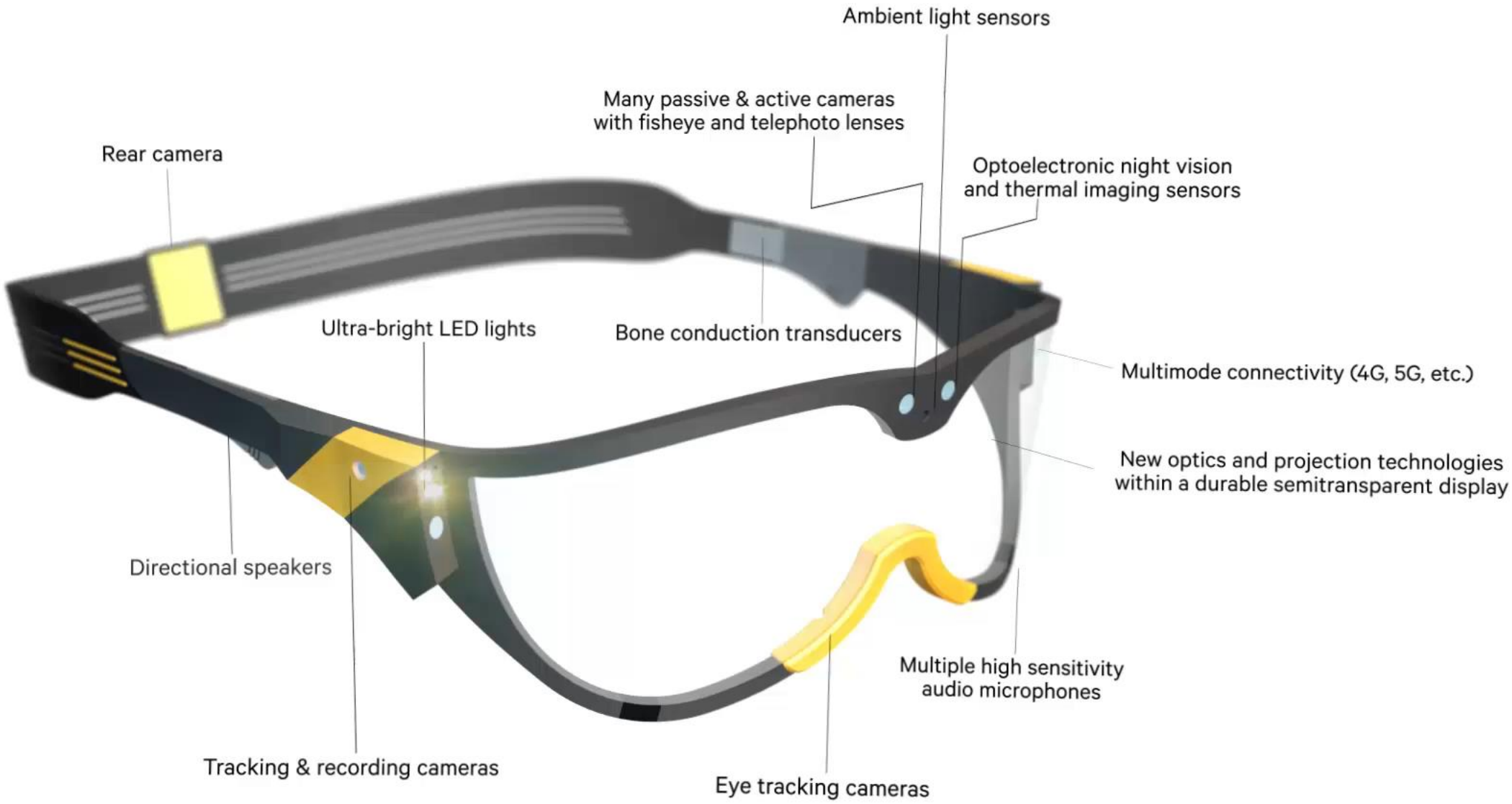
Tango

Glasses

ODG
OSTERHOUT DESIGN GROUP

Qualcomm Snapdragon is a product of Qualcomm Technologies, Inc.

Will the smartphone become an AR wearable?


Ambient light sensors
Many passive & active cameras
with fisheye and telephoto lenses
Rear camera
Optoelectronic night vision
and thermal imaging sensors
Ultra-bright LED lights
Bone conduction transducers
Multimode connectivity (4G, 5G, etc.)
New optics and projection technologies
within a durable semitransparent display
Directional speakers
Multiple high sensitivity
audio microphones
Tracking & recording cameras
Eye tracking cameras

Challenges and Requirements for AR

New technologies for future AR requirements

Providing an always-on experience that intelligently enhances our lives

Immersive

The visuals, sounds, and interactions are so realistic that they are true to life

Cognitive

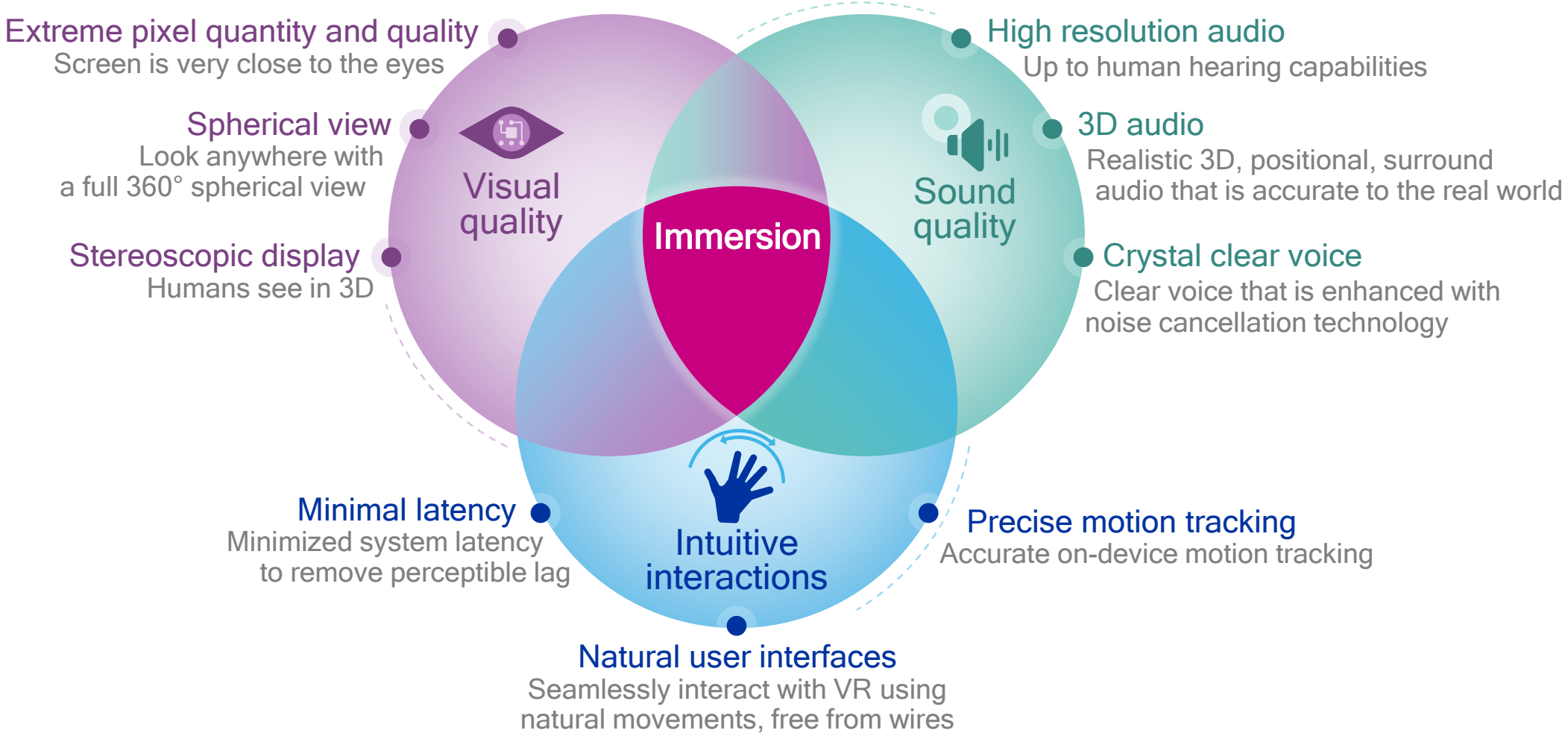
It understands the real world, learns personal preferences, and provides security & privacy

Connected

An always-on, low power wearable with fast wireless cloud connectivity anywhere

AR shares requirements similar to VR for immersion

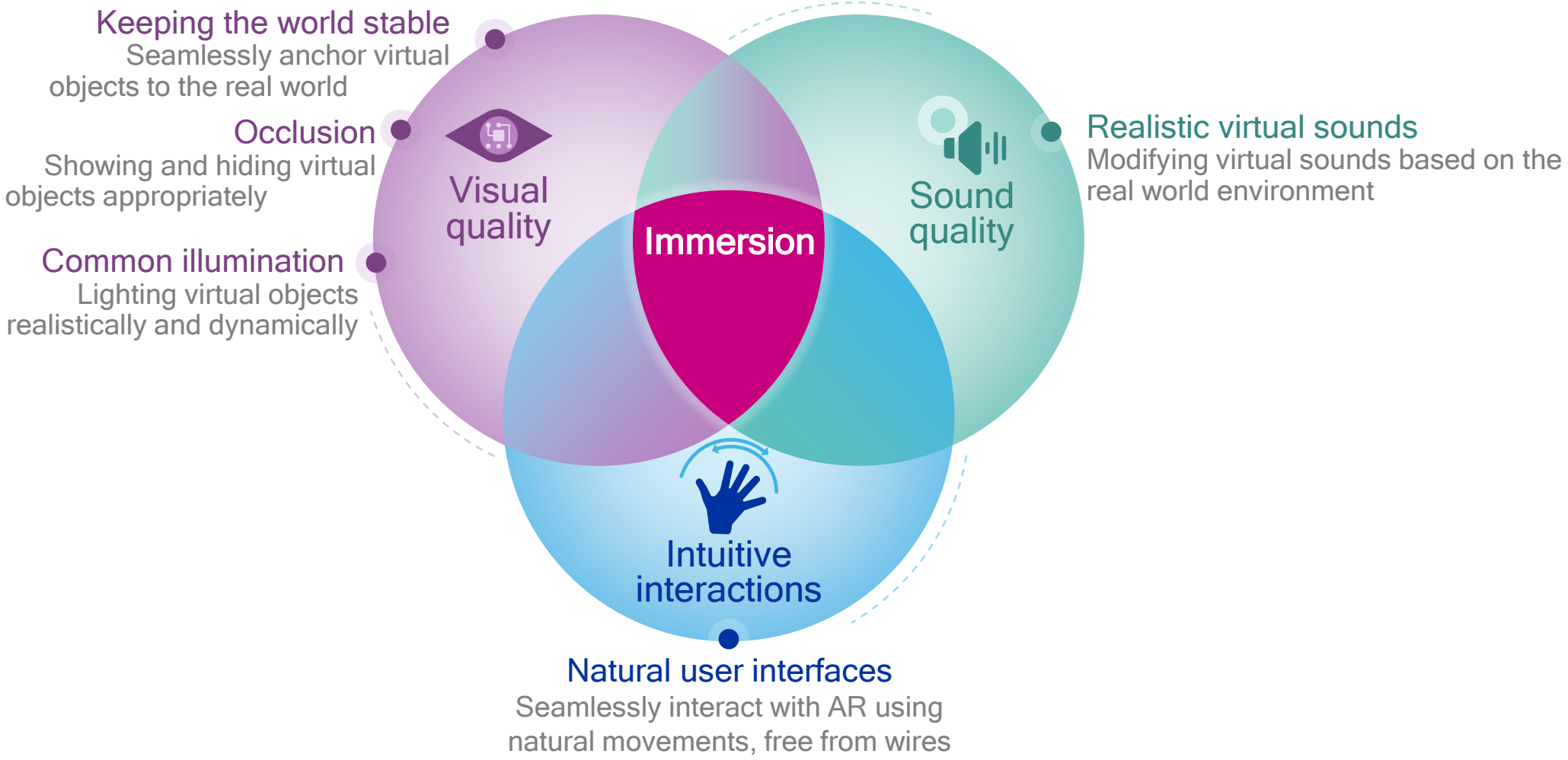
Achieving realistic AR at low power to enable a comfortable, sleek form factor



Learn more about our vision for the future of VR: www.qualcomm.com/VR

AR introduces additional requirements for immersion

Seamlessly integrating virtual objects with the real world



Keeping the world stable

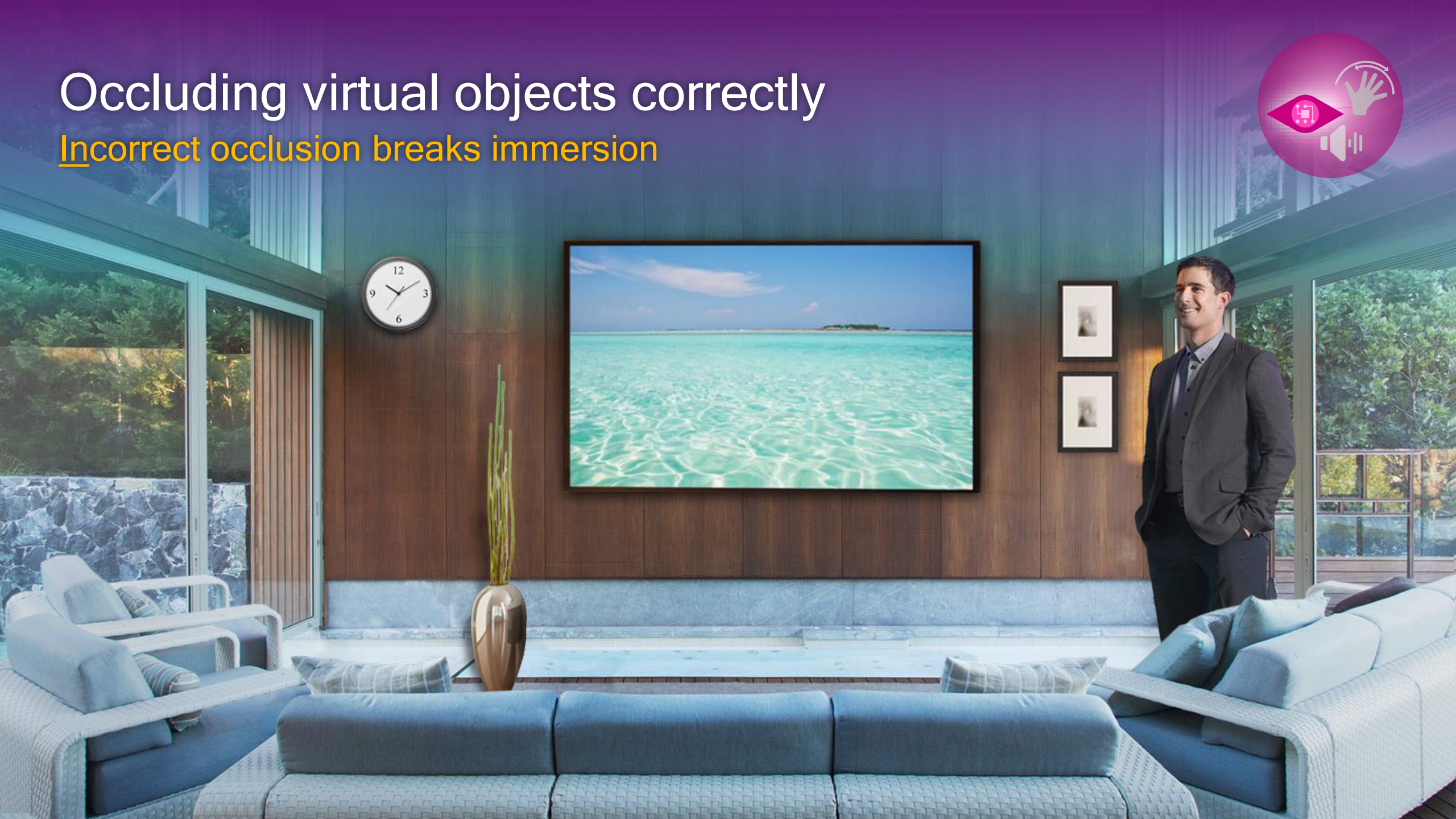
In an unstable environment, virtual objects are NOT seamlessly anchored to the real world

Keeping the world stable

In a stable environment, virtual objects are seamlessly anchored to the real world

Occluding virtual objects correctly

Incorrect occlusion breaks immersion



Occluding virtual objects correctly

Correct occlusion accounts for the depth of virtual and real objects

Occluding virtual objects correctly

Smart occlusion accounts for both object depth and user preferences

Lighting virtual objects realistically and dynamically

Incorrect lighting poorly represents the position, intensity, and orientation of all light sources

Poor environment processing

- Virtual objects look fake and out of place
- Static lighting; often incorrect for environment
- Solid objects do not look solid
- Materials look physically incorrect
- Interactivity is not smooth

Lighting virtual objects realistically and dynamically

Correct lighting considers the position, intensity, and orientation of all light sources

Proper AR environment processing

- Virtual objects look real and correctly placed
- Dynamic lighting; correct for the environment
- Solid objects look solid
- Materials look physically correct
- Interactivity is smooth

Making it possible

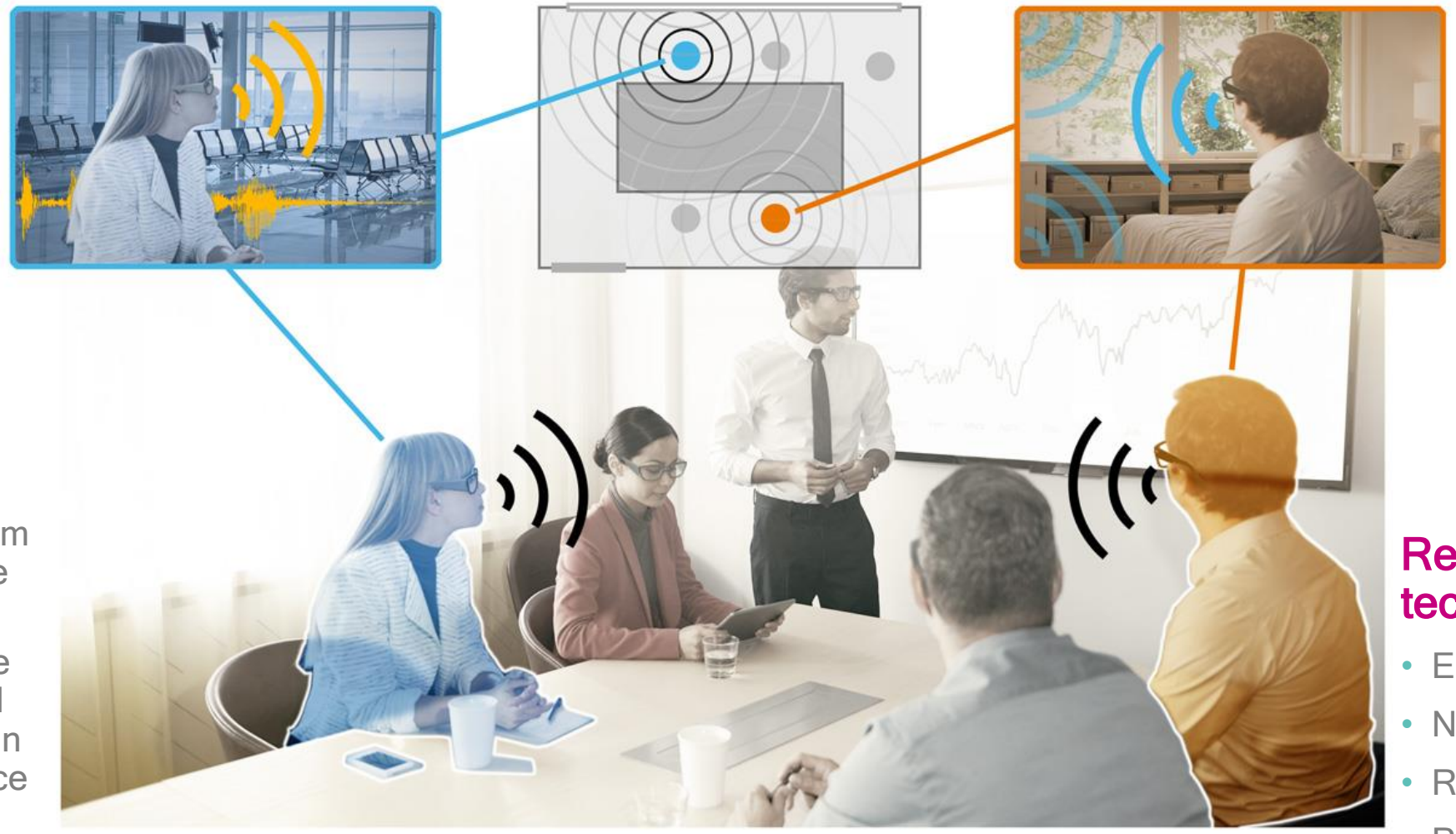
- Intelligent, fast interaction between many different sensors & rendering systems
- New computer vision and global illumination algorithms to dynamically render and overlay realistic AR objects

Creating virtual sounds based on the real world

Sound reflections spread and interact with the environment appropriately

Airport

- Limited sound reflections
- Significant ambient sound



Hotel room

- Significant sound dampening

Conference room

- Enclosed room with reflective surfaces
- Virtual people should sound like they are in the conference room

Required technologies

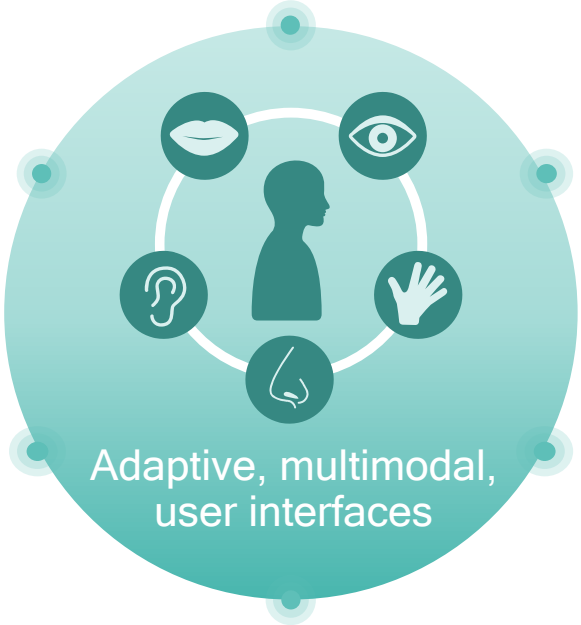
- Environment modeling
- Noise filtering
- Reverberation
- Positional audio

Interacting naturally with AR

Interactions will become more intuitive & adaptive to personal preferences

Motion & gesture recognition
Use CV along with motion sensors, and new types of connected, haptic devices to help users interact within AR

Speech recognition and learning
Use of natural language processing, intelligently personalized to user's voice and lexicon



Face recognition
Use of advanced CV to authenticate and accurately recognize facial expressions

Personalized interfaces
Learn and know a myriad of user preferences based on machine learning

Eye tracking
Use CV to much more accurately authenticate, and also track & measure point of gaze

Bringing life to objects
Efficient user interfaces for controlling interaction with IoT devices and cloud services

Cognitive technologies are key for AR adoption

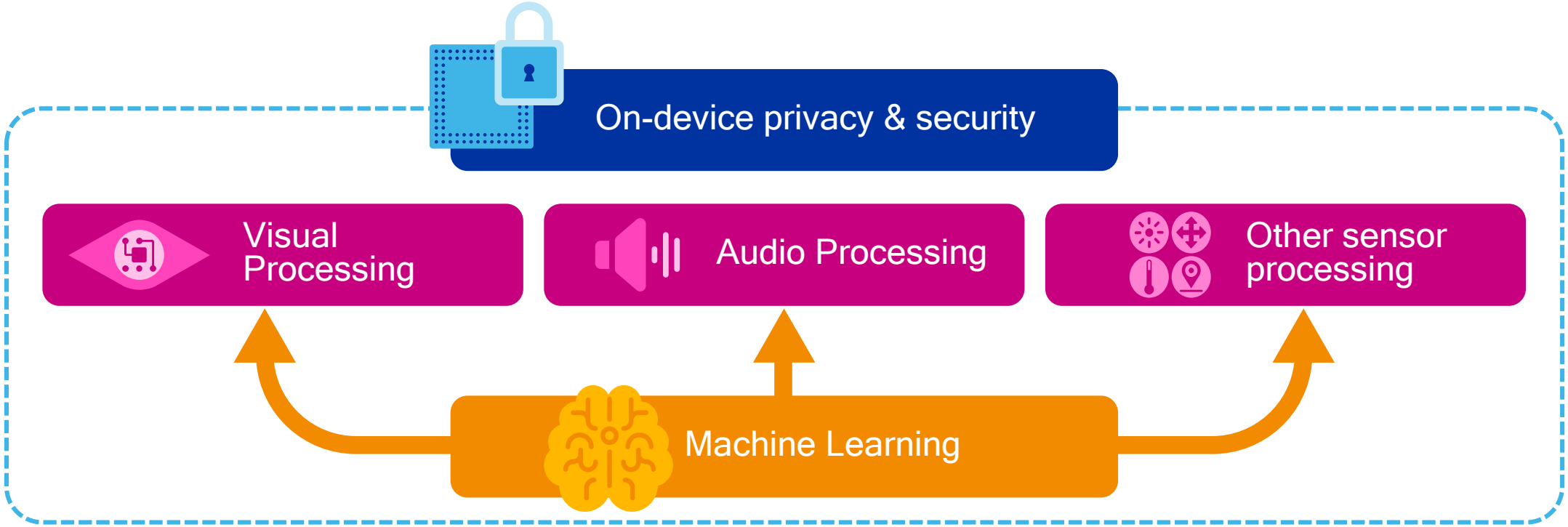
Making sense of the world while protecting our privacy and security

Machine learning

- Makes visual, audio, and other sensor processing more intelligent

Security & privacy are critical for AR

- Continuous authentication necessary for identity and access
- Utilize combined machine learning on biometrics and behavioral activity



Contextual intelligence to “sense” the world

Many types of sensors and personal information are required

Sensor data

Environment | Iris scan | Ambient light | Compass | Camera | Microphone | Temperature | Humidity | Gyroscope | Pulse

On-device data

Calendar | Messaging | Apps

Cloud and IoT data

Cloud data | Connectivity

Low power sensing, processing, and connectivity

- Efficient, heterogeneous architectures
- Sensor fusion and machine learning
- Integrated, always-on data capturing
- Low-energy wireless technologies (e.g. BT-LE, LTE IoT, LTE Direct, 802.11ah)

Visual intelligence to “see” the world around you

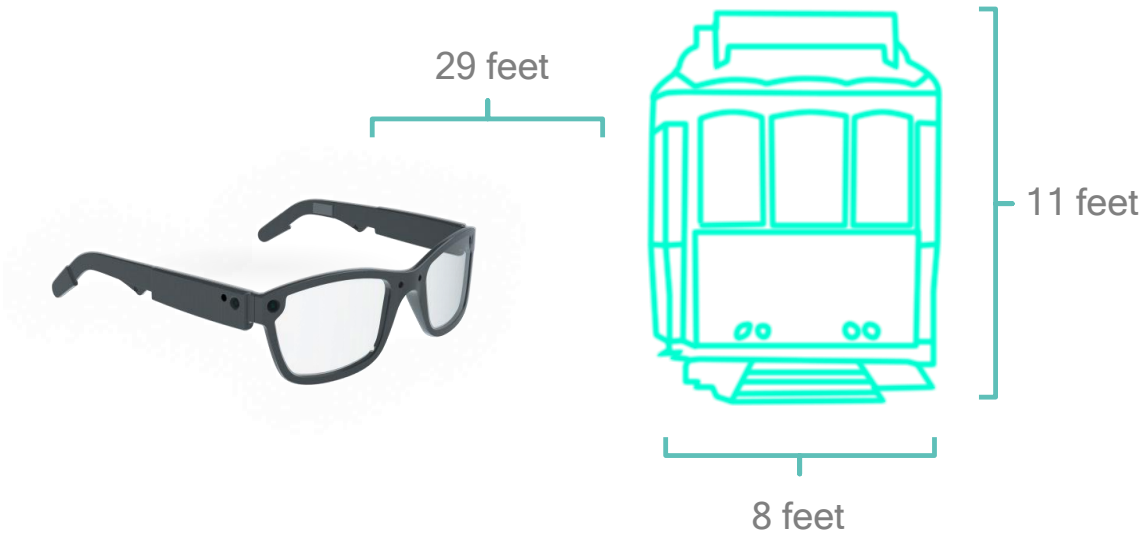
Continually monitoring the visual world to intelligently identify

3D depth capture, interpolation & reconstruction

Using passive + active cameras, along with advanced CV & machine learning algorithms

Object recognition, tracking & registration

Using CV and machine learning so that objects in the real and virtual worlds are properly aligned with respect to each other



Determine the size, direction, and distance of different objects, and (sometimes) store the whole 3D scene for various uses

Recognize, track, map & reconstruct surroundings

Underlying technologies

Object recognition, tracking and registration

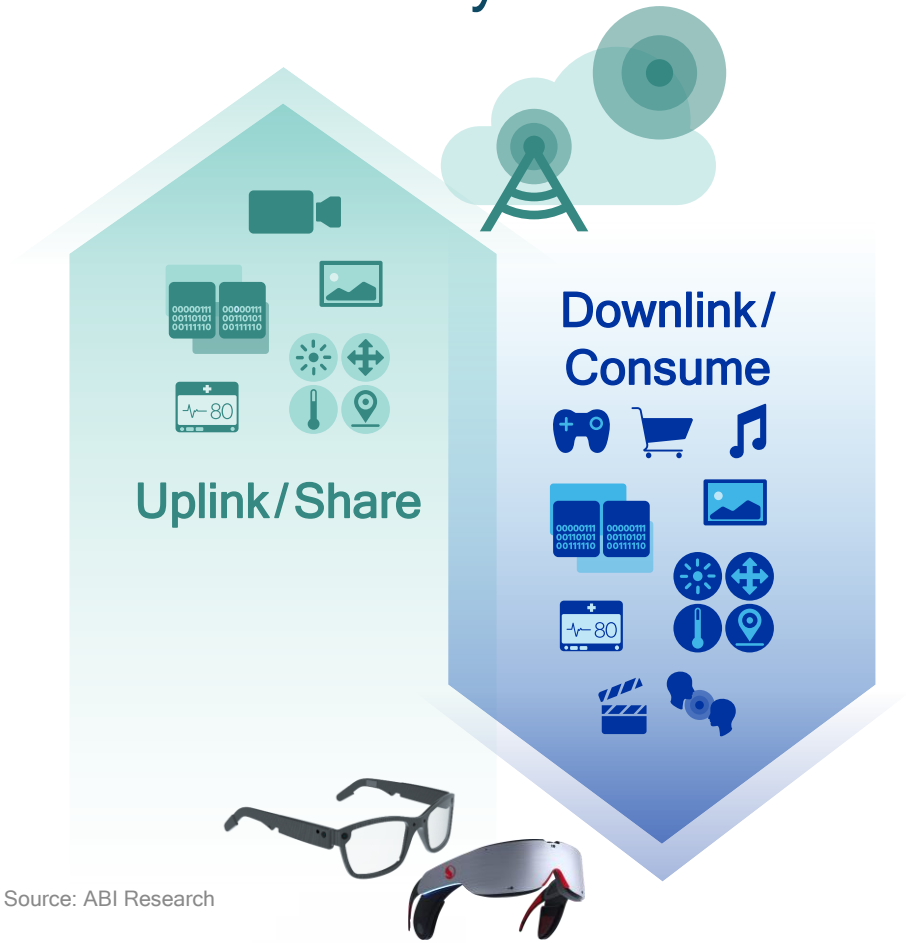
Simultaneous localization and mapping (SLAM)

Visual inertial odometry (VIO)

3D reconstruction

VR and AR require efficient increase in wireless capacity

Constant up/download on an all-day wearable



Richer visual content

- Higher resolution, higher frame rate
- Stereoscopic, High Dynamic Range (HDR), 360° spherical content, 6 DoF

2 Mbps
Video conferencing

5 to 25 Mbps
Two-way telepresence

50 to 200 Mbps
Next-gen 360° video (8K, 90+ FPS, HDR, stereoscopic)

Bandwidth

1 Mbps
Image and workflow downloading

2 to 20 Mbps
3D model and data visualization

10 to 50 Mbps
Current-gen 360° video (4K)

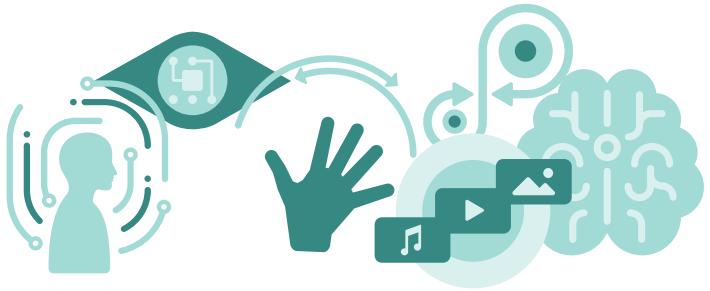
200 to 5000 Mbps
6 DoF video or free-viewpoint

Critical for immersive experiences

Source: ABI Research

Power efficiency is also essential for AR

The AR headset needs to be comfortable to wear all day



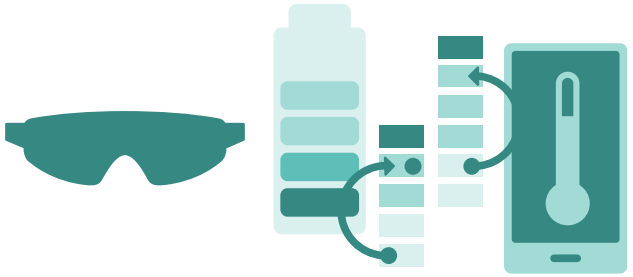
AR workloads

Compute intensive

Complex concurrencies

Always-on

Real-time

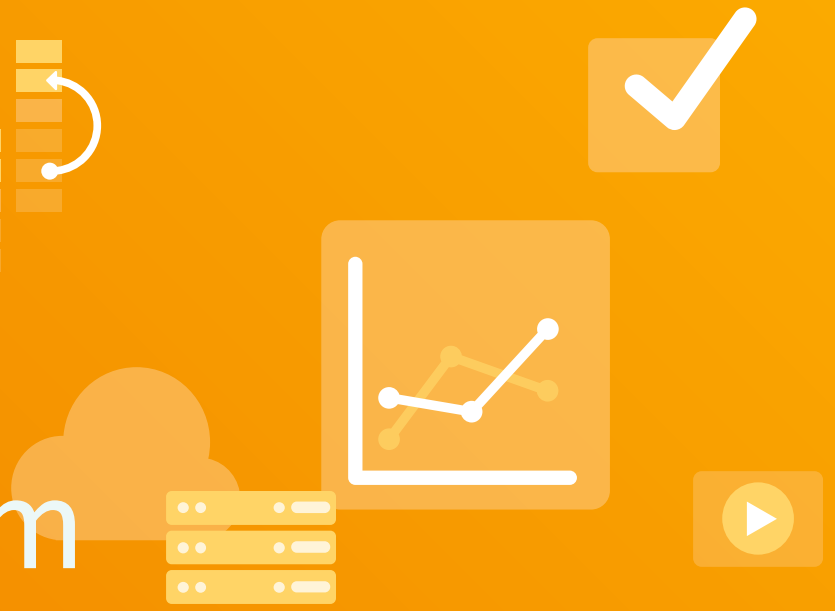


Constrained mobile wearable environment

Thermally efficient for sleek, ultra-light designs

Long battery life for all-day use

Technical Enablers: Mobile Compute Platform



We’re developing foundational technology for AR

Qualcomm Technologies’ investments and the confluence of mobile technologies

Computer Vision

- 6-DOF VIO
- SLAM & 3DR
- Object detection & recognition

Cognitive & Security

- AI for advanced cognitive processing
- Local and cloud machine learning
- Security and privacy

Heterogeneous Computing

- Lower power, higher perf. AR visual processing
- Advancements in always-on sensor fusion
- Next-gen AR audio

Next-Gen Connectivity

- Gigabit LTE and Wi-Fi
- Pioneering 5G technologies
- Connectivity convergence

Mobile XR is advancing in every generation

Qualcomm® Snapdragon™ 820

3DoF

1K x 1K @ 60 fps

Snapdragon 835

6DoF

1.5K x 1.5K @ 60 fps

Snapdragon 845

Room-Scale 6DoF

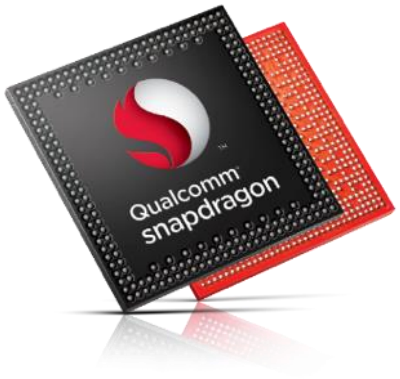
SLAM

2K x 2K @ 120fps

Qualcomm Snapdragon is a product of Qualcomm Technologies, Inc.

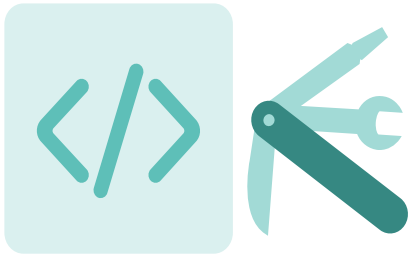
Supporting Tangible Efforts

Qualcomm® Snapdragon™ 835/845 Mobile VR Platform



Snapdragon 835/845

Purpose-built mobile processor for superior XR experiences

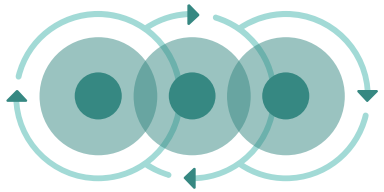


Snapdragon VR SDK

Easy developer access to Snapdragon accelerated XR libraries that simplify application development

HMD Accelerator Program

Reference designs, working with ODMs and technical support to commercialize XR HMDs quickly



Platform and Ecosystem support

Working with multiple content, technology and platform companies

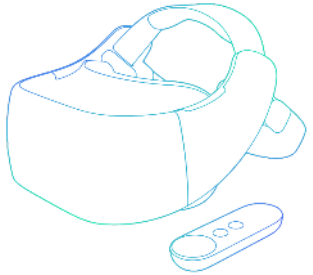
Actively working with AR/XR device manufacturers

AR/XR products based on Snapdragon Mobile VR Platform

ODG

Pico

Lenovo



htc

VR HMD	VR HMD	Tango and Daydream	Daydream	Daydream	Tango
Baofeng Matrix	Coocaa Wondergate G1	ASUS Zenfone AR	Google Pixel	Google Pixel XL	Lenovo Phab 2 Pro
VR HMD	VR HMD	Daydream	Daydream	Gear VR	Gear VR
iQiyi Adventure	WhaleyVR	ZTE Axon	Moto Z	Samsung Galaxy S8	Samsung Galaxy S8 Edge

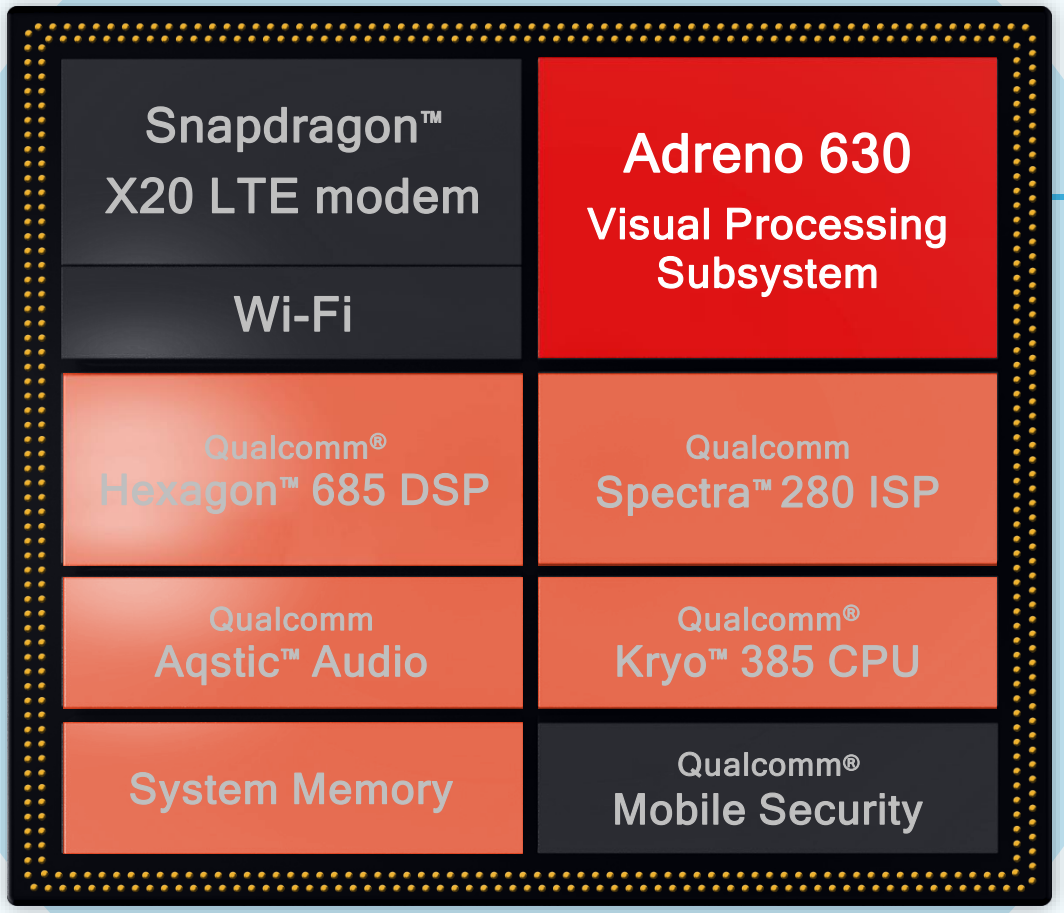
Snapdragon 845

XR/AR

Computer vision, image processing, sensor processing, graphics, video processing, location, and cloud interaction

Benefits

- Integrated and optimized
- Enhanced battery life
- Thermal efficiency
- Standardized implementation
- Mass market cost
- Variety of use cases and industry support



Entire SoC is used!

Room-Scale 6DOF with SLAM

Adreno Foveation

2K x 2K @ 120 fps XR

30% Faster Graphics and Video*

30% Power Reduction*

Improved 6DOF Hand Tracking and Controller support*

Iris Authentication

Voice UI with custom keyword

Qualcomm 3D Audio SDK for Unity

*Compared to Snapdragon 835

Qualcomm Mobile Security, Qualcomm Hexagon, Qualcomm Spectra, Qualcomm Aqstic and Qualcomm Kryo are products of Qualcomm Technologies, Inc.

Smart Computing

Intuitive Interactions

Room Scale
6DoF SLAM

Advanced Security

Intuitive Interactions

Iris Authentication

Visual Processing

Qualcomm® Adreno™ GPU Foveation

Adreno Tile Rendering

Eye-Tracking

Multiview Rendering

Fine Grain Preemption

Qualcomm Adreno is a product of Qualcomm Technologies, Inc.

Visual Processing

Qualcomm® Adreno™ GPU Foveation

Adreno Tile Rendering

Eye-Tracking

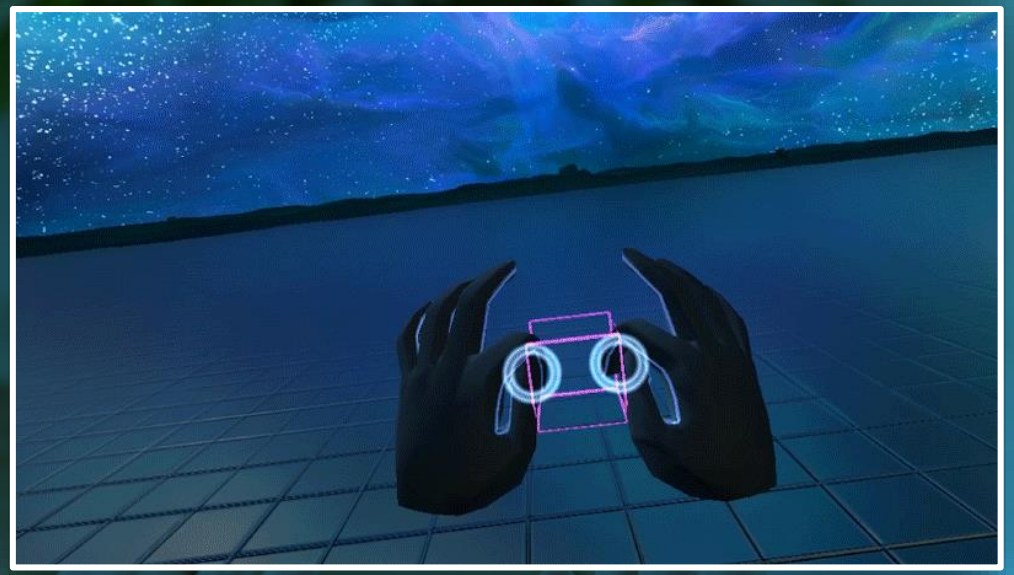
Multiview Rendering

Fine Grain Preemption

Qualcomm Adreno is a product of Qualcomm Technologies, Inc.

Natural Interactions

Hand Tracking



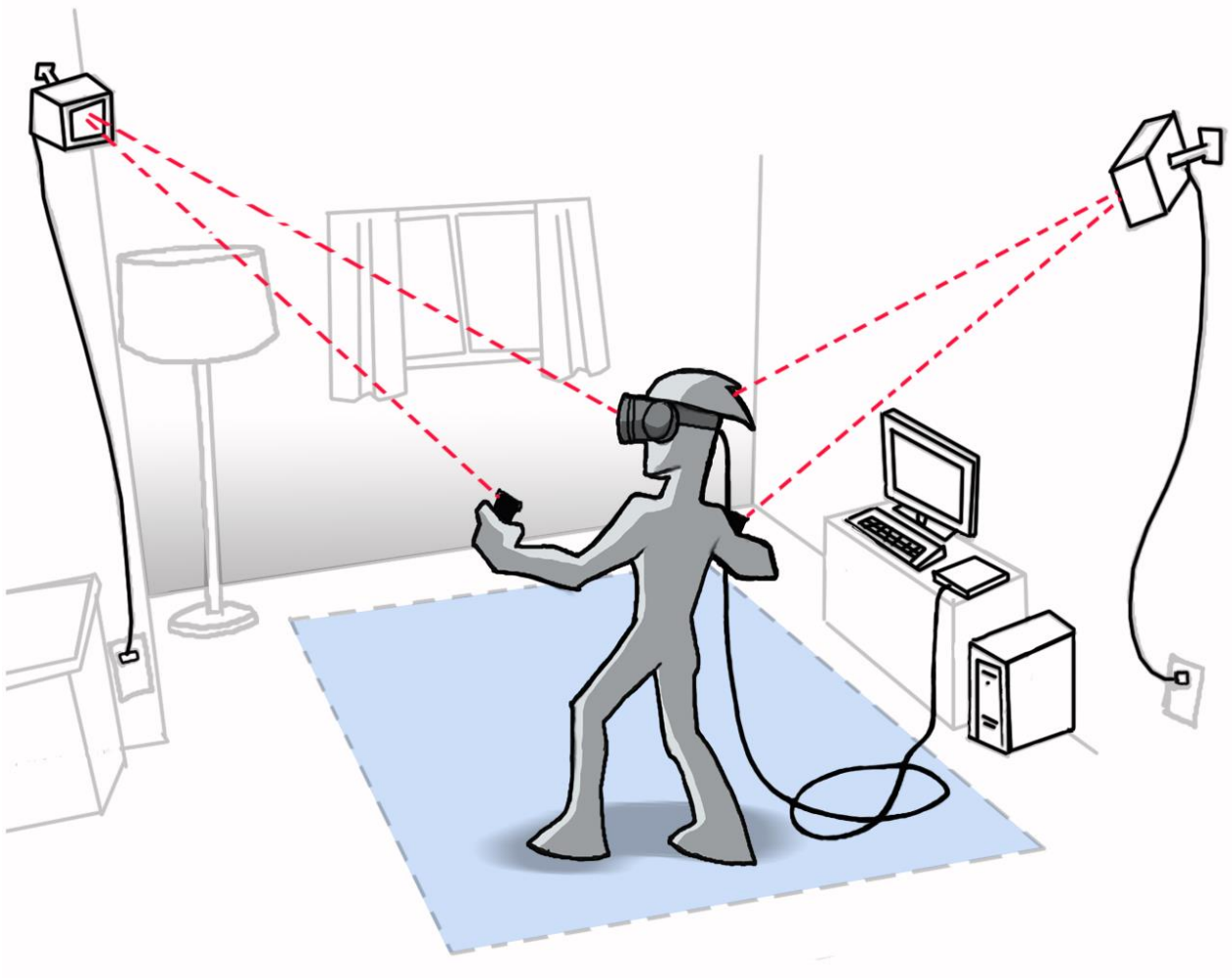
Audio Intelligence

Sounds

- Always On
- Voice commands
- 3D Audio
- Echo Cancellation
- Noise reduction

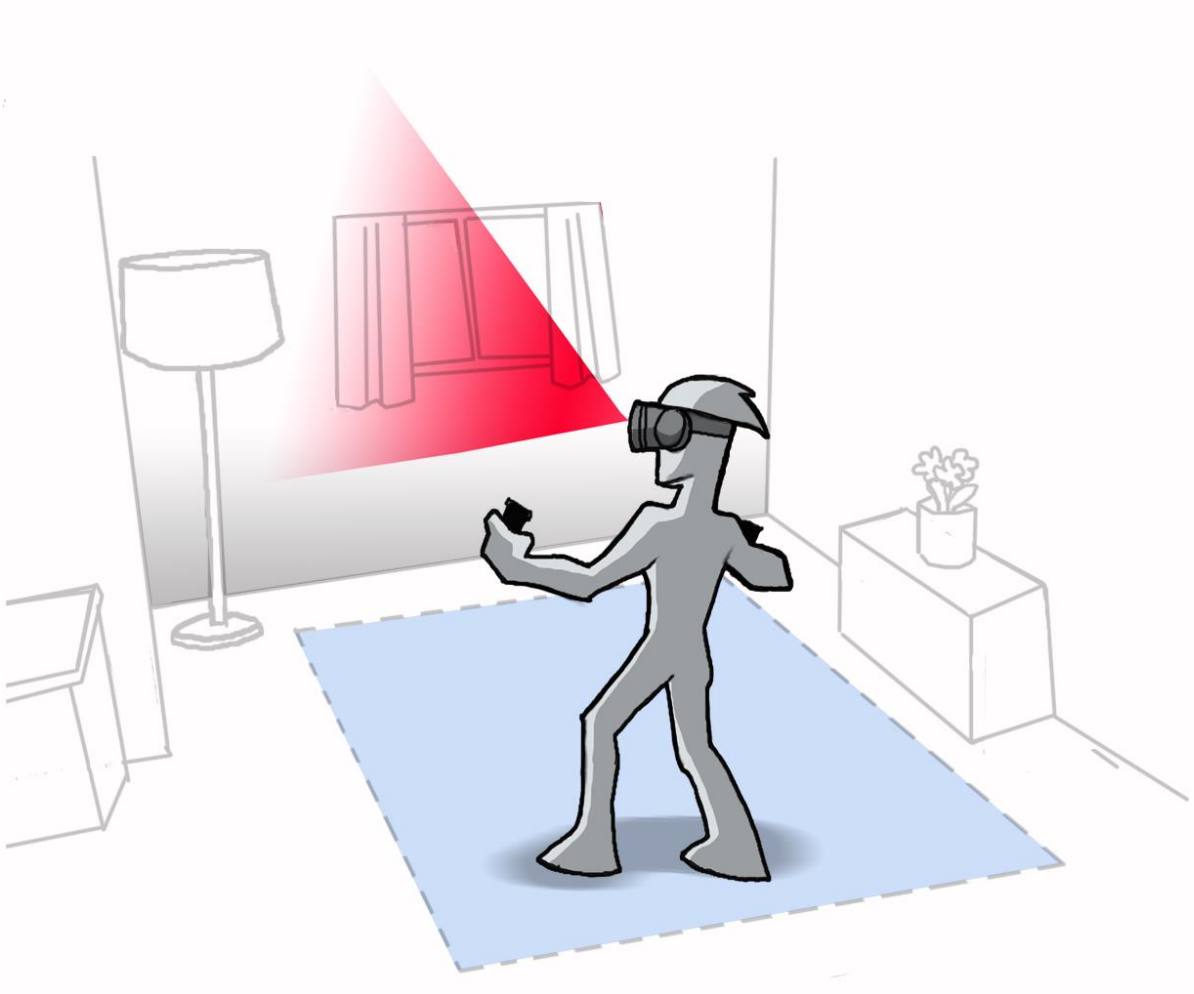
Conventional 6-DoF: “Outside-in” tracking

External sensors determine the user’s position and orientation



Mobile 6-DoF: “Inside-out” tracking

Visual inertial odometry (VIO) for rapid and accurate 6-DoF pose



Mobile 6-DoF: “Inside-out” tracking

Visual inertial odometry (VIO) for rapid and accurate 6-DoF pose

Mono or stereo camera data

Captured from tracking camera image sensor at ~30 fps

Accelerometer & gyroscope data

Sampled from external sensors at 800 / 1000 Hz

“VIO” subsystem on Qualcomm® Snapdragon™ Mobile Platform

Camera feature processing

Inertial data processing

Qualcomm® Hexagon™ DSP algorithms

- Camera and inertial sensor data fusion
- Continuous localization
- Accurate, high-rate “pose” generation & prediction

6-DoF position & orientation

(aka “6-DoF pose”)

Qualcomm Hexagon is a product of Qualcomm Technologies, Inc.

Device Architecture / Technology Enablers for XR

Optimizations needed across the SoC and system SW

HW streaming interfaces

- Bypass DRAM with engine to engine communication
  - ISP to DSP
  - Sensors to DSP

Global time stamps

- Maintain synchronization across various processing engines

Accurate motion tracking

- Fast and accurate 6-DOF
- Accurate predictive 6-DOF for a small future window

Late latching

- Using the latest pose
- Asynchronous threads for consistent frame rate

Multiview stereoscopic rendering

- Single pass render of left and right eye

High frame rate

- 90 FPS for reduced frame latency

Foveated rendering

- Reduce pixel generation workload while maintaining high image quality

Direct access to VR features

- Bypass Android latency

Fast connectivity

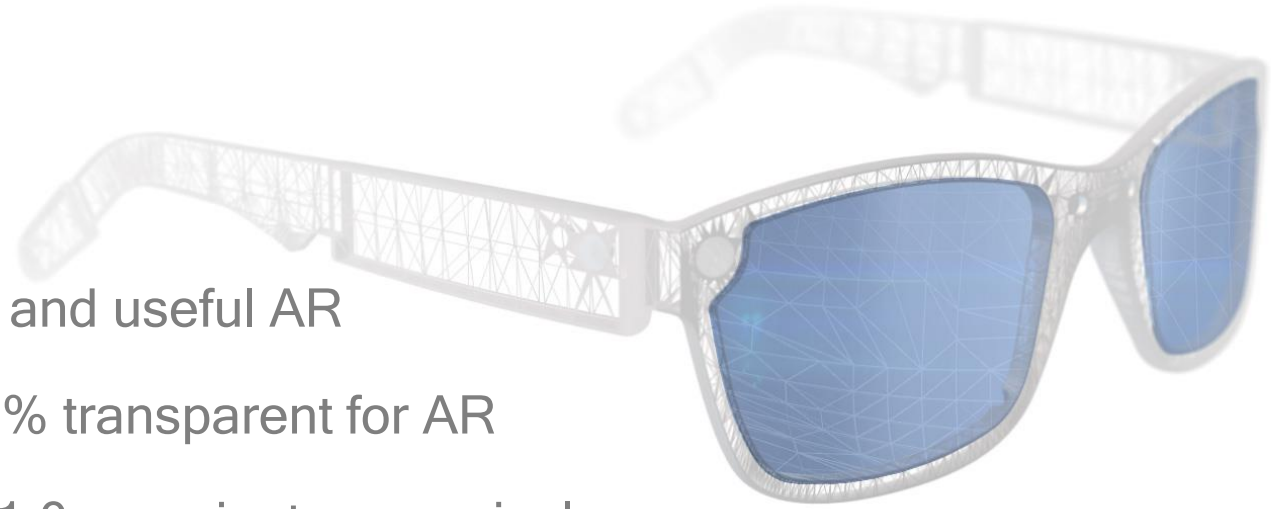
- Low latency connectivity
- High bandwidth

Single buffer

- Low latency connectivity
- High bandwidth

Future XR needs a disruption in display technology

First step towards high volume converged XR form factor are new displays



- Solve the vergence accommodation conflict
- Deliver necessary FoV both for immersive VR and useful AR
- Be completely opaque for VR, yet at least ~85% transparent for AR
- Support an angular resolution of at least 0.5 – 1.0 arc minutes per pixel
- Drive HDR, at least Rec. 2020 gamut, with ~5X improvement in nits
- Be capable of refreshing at a minimum of ~120Hz (per eye)
- Be light, mechanically flexible, very durable, and eventually cost under ~$100 at very high volumes

Standardization Efforts

Global, Optimized HW Implementation, Interoperability

We need alignment in several key areas

Creating immersive XR experiences through standards-based technologies

Professional production & user generation

Video and audio compression

Transport: Content delivery

Access network

Apps

Display and rendering

ITU

3GPP A GLOBAL INITIATIVE

SMPTE

MPEG

DASH Industry Forum

DVB

ATSC 3.0

IETF

3GPP A GLOBAL INITIATIVE

IEEE

W3C WebVR

OpenXR

Java

VESA

Consumer Technology Association

ITU

VR INDUSTRY FORUM

5G enhanced mobile broadband

is required for XR mass adoption

Extreme throughput – multi Gbps
Ultra-low latency – down to 1 ms
Uniform experience – even at cell edge

XR video will be the killer use case for 5G

~10 to 50 Mbps

Current-generation
360° 4K/30fps video

~200 to 5000 Mbps, very low latency

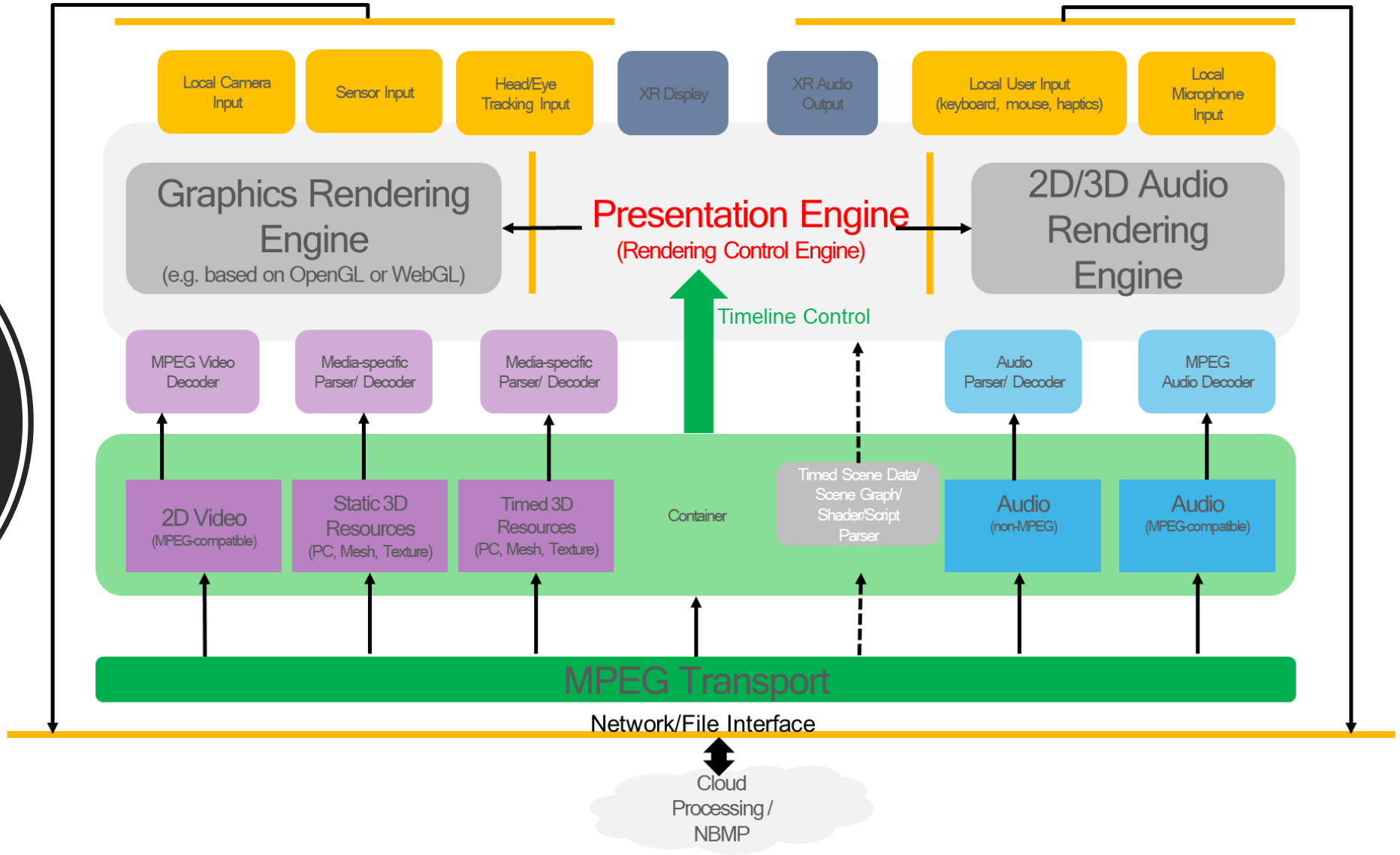
Next-decade
Interactive, real-time 3D "Free-Viewpoint"
6-DoF 8K/90-120fps HDR-next video

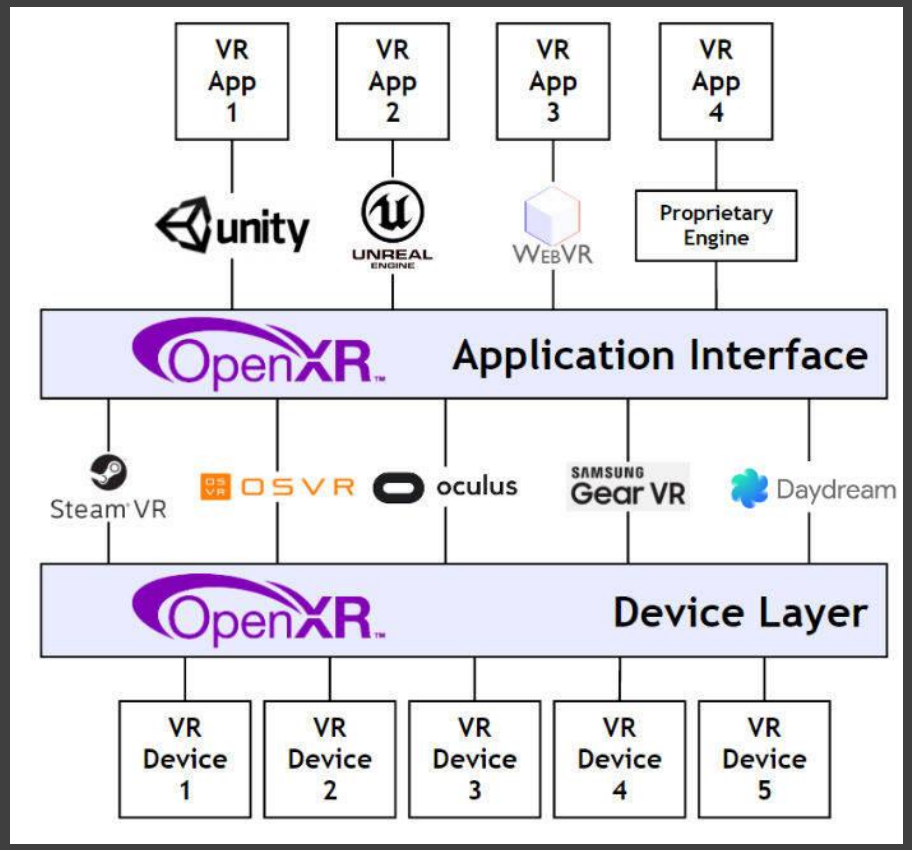
Sustained network performance

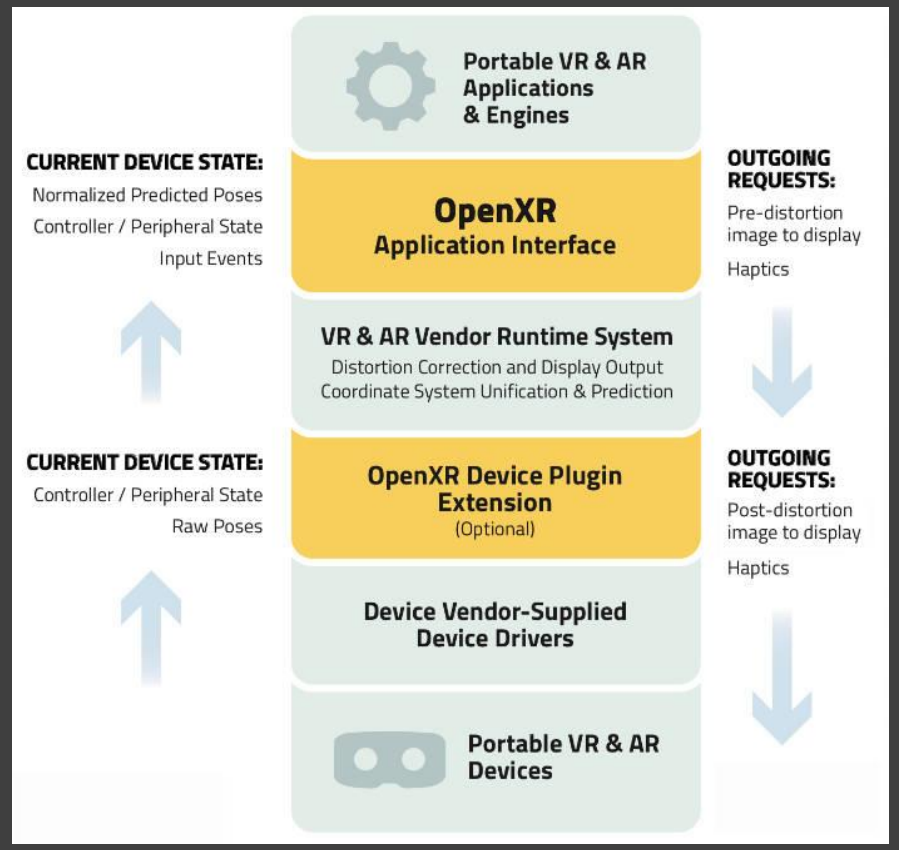
~50 to 200 Mbps, lower latency

Next-generation (2019)
3D 360° 8K/30fps viewport-aware HDR10 video

Rendering Centric View

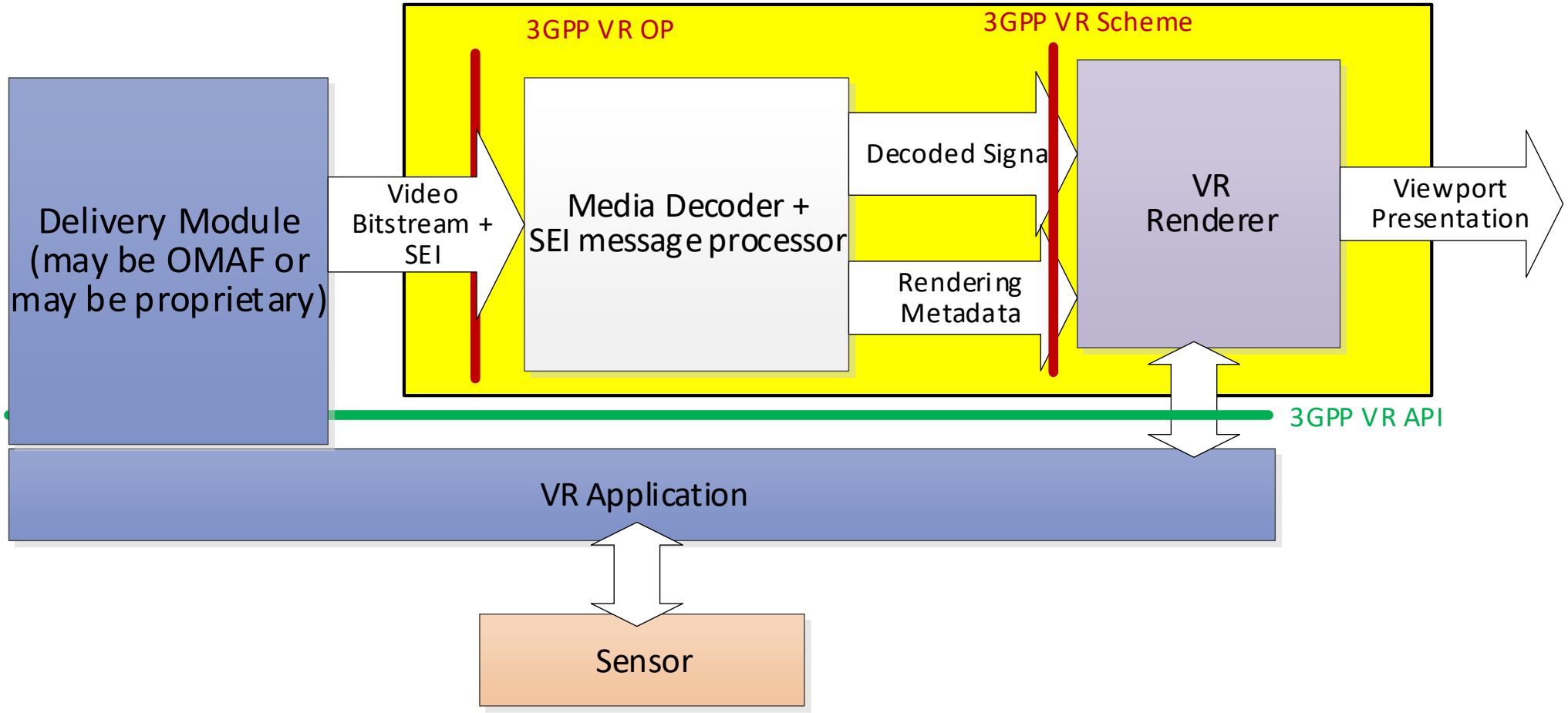




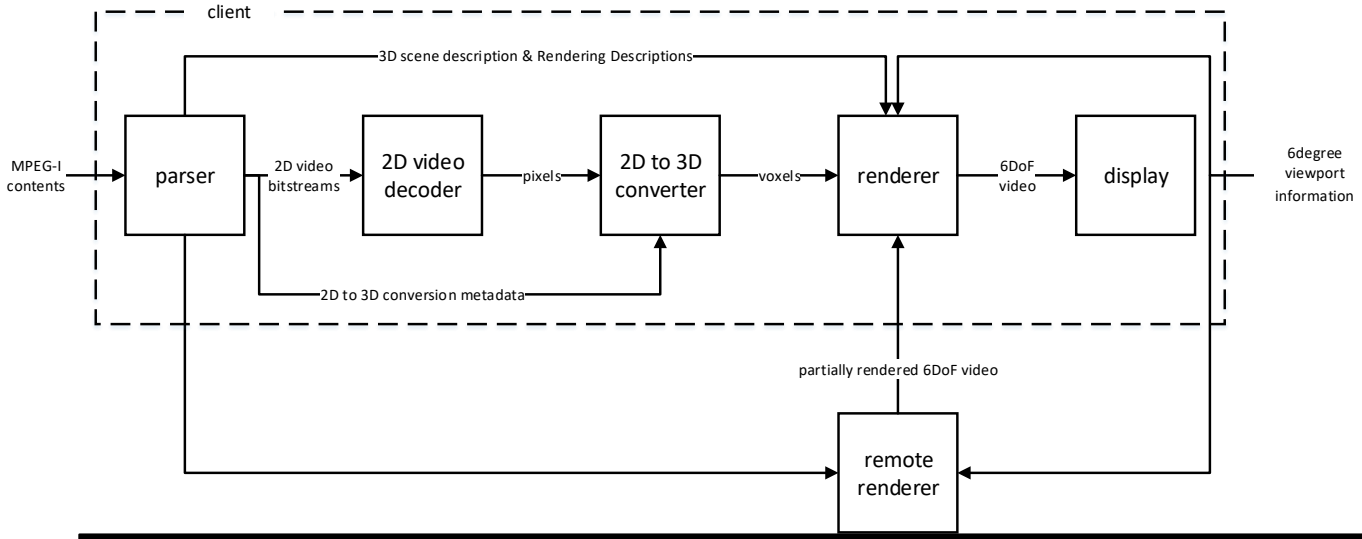


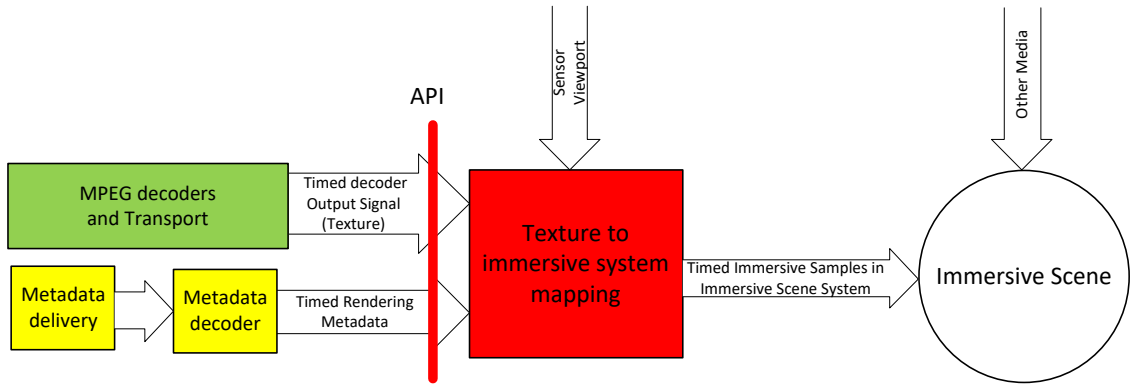
The OpenXR™ working group - previously known as the Khronos VR Initiative - is creating an open standard for VR and AR applications and devices.

The 3GPP VR approach for Rel-15 for Audio and Video

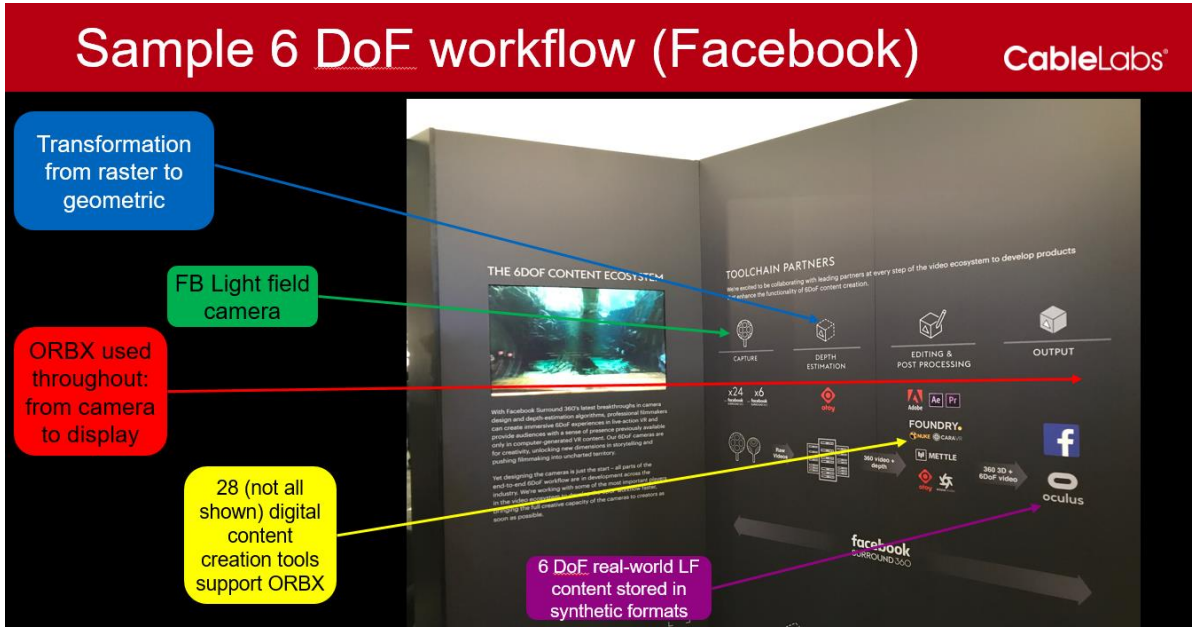


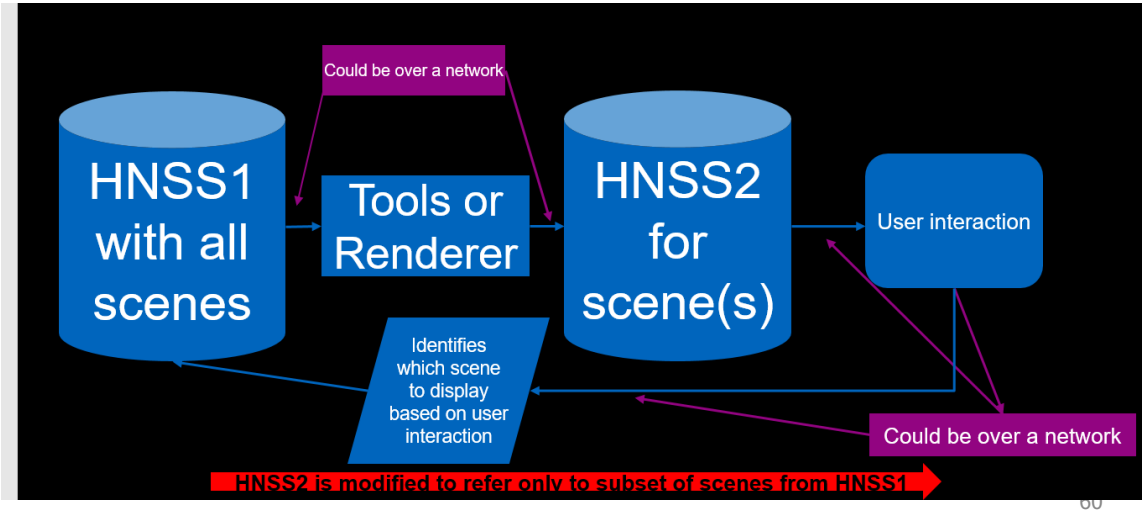
The traditional MPEG world



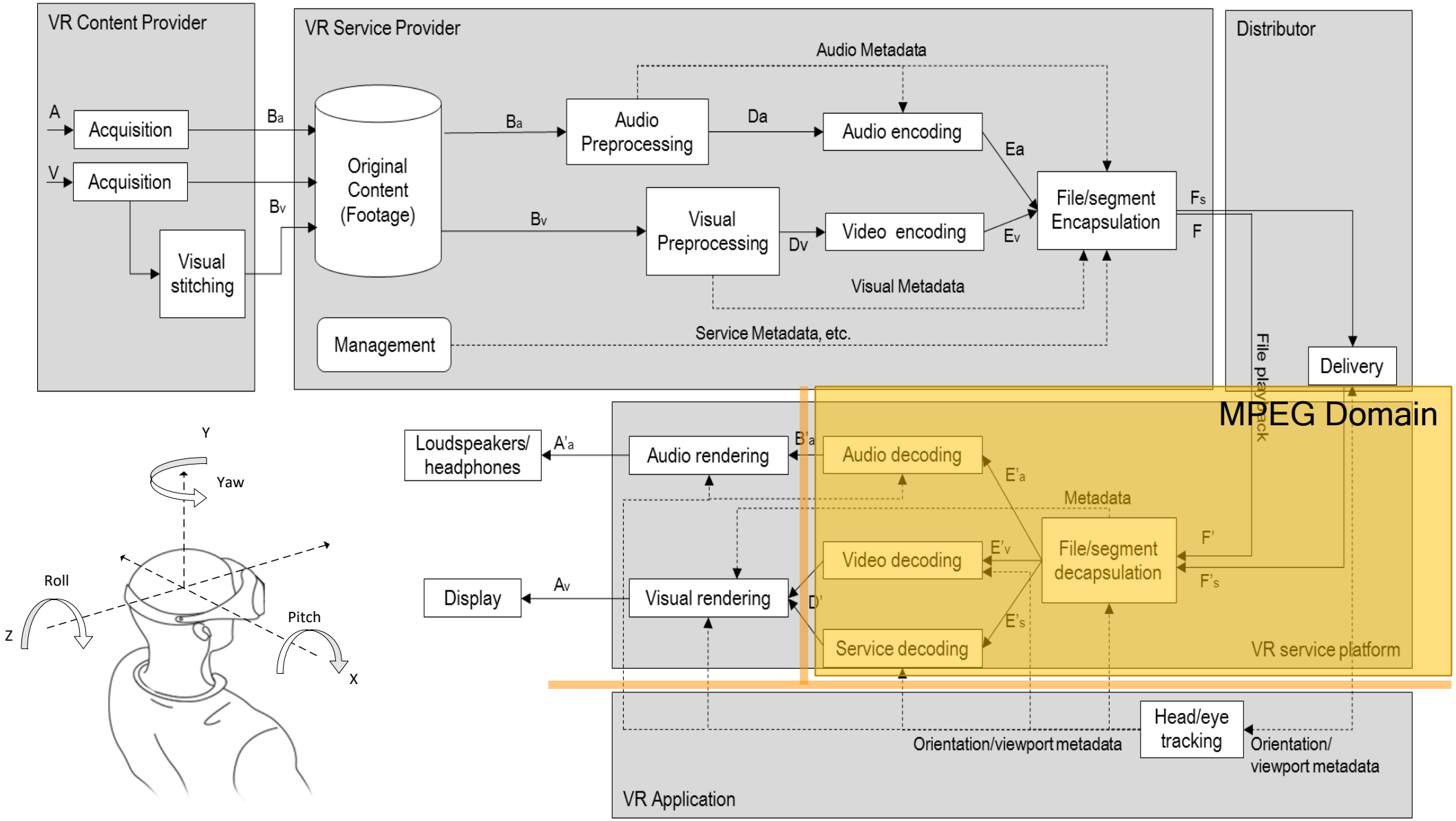


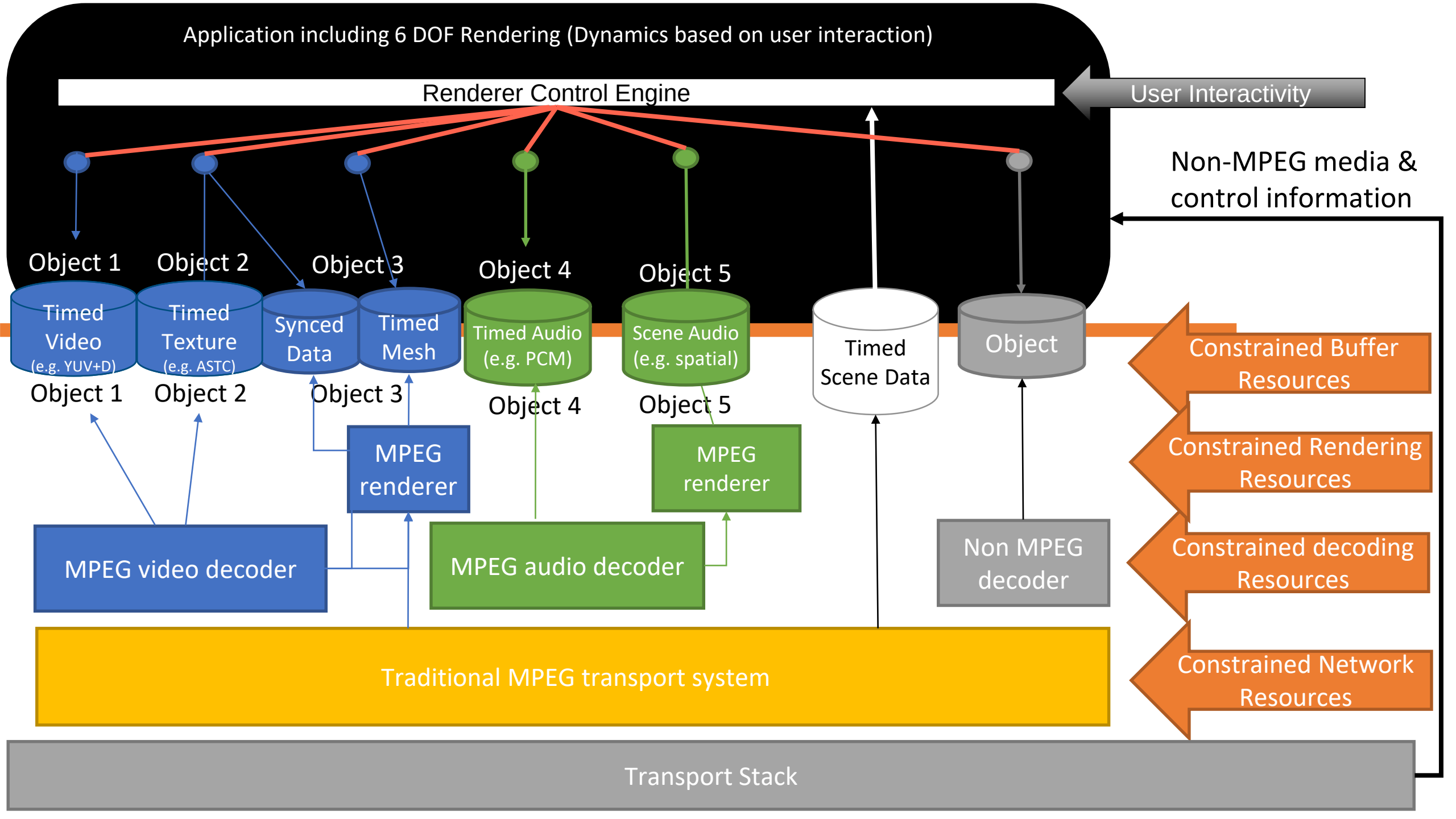
The new MPEG world?



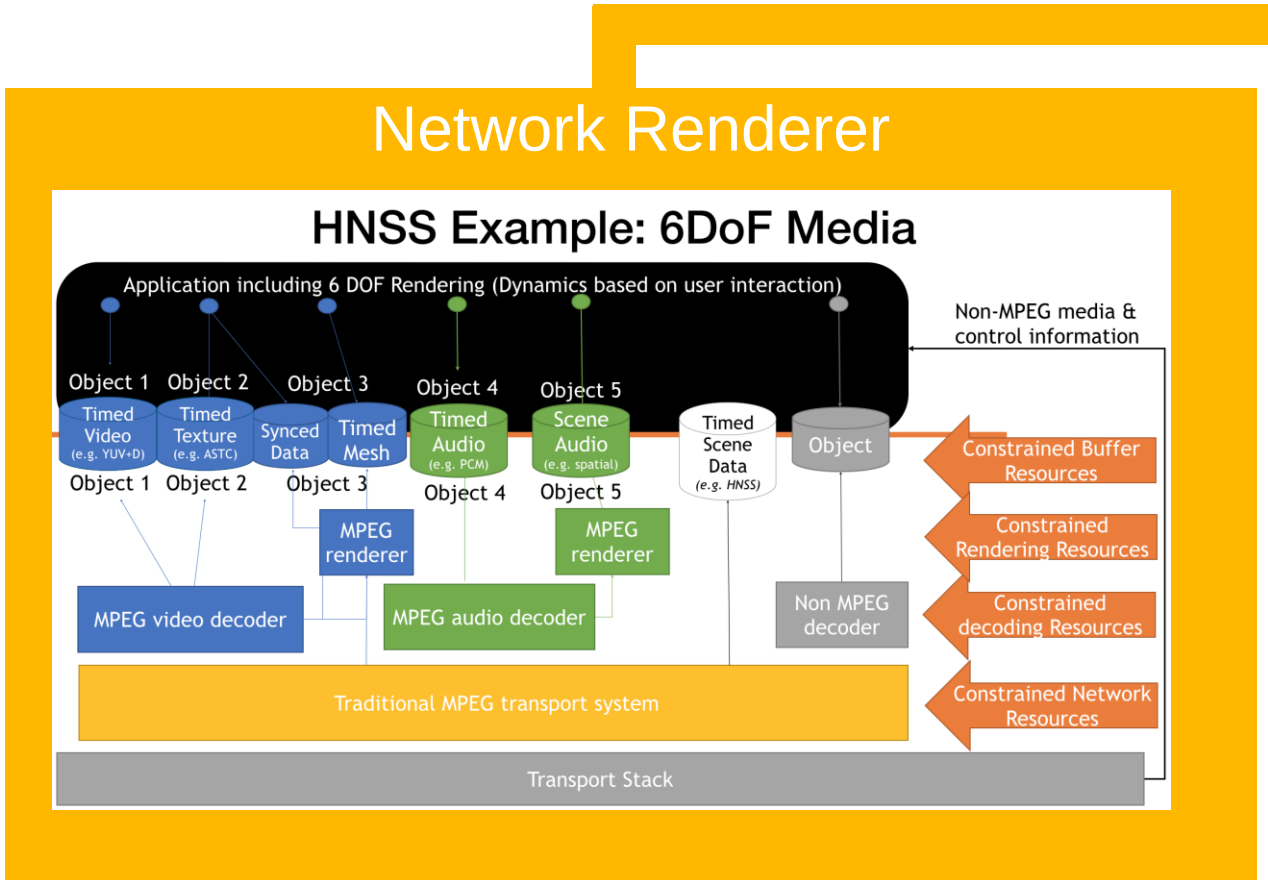


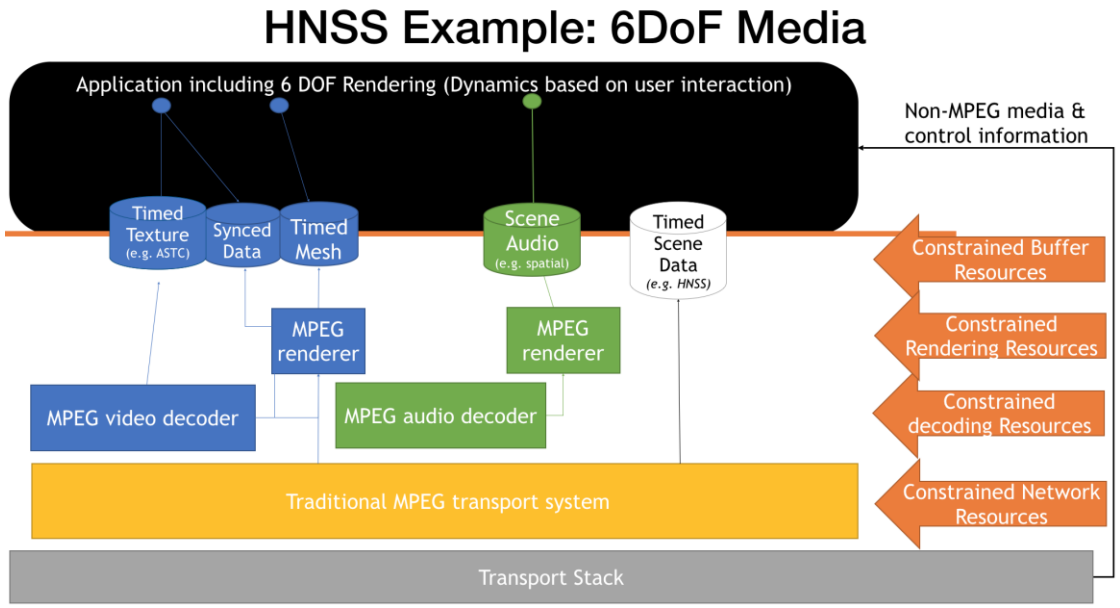
Traditional MPEG – the world in front or around moves and I give you everything (almost)




Application including 6 DOF Rendering (Dynamics based on user interaction)
Renderer Control Engine
User Interactivity
Non-MPEG media & control information
Object 1
Object 2
Object 3
Object 4
Object 5
Timed Video (e.g. YUV+D)
Timed Texture (e.g. ASTC)
Synced Data
Timed Mesh
Timed Audio (e.g. PCM)
Scene Audio (e.g. spatial)
Timed Scene Data
Object
Object 1
Object 2
Object 3
Object 4
Object 5
MPEG renderer
MPEG renderer
MPEG video decoder
MPEG audio decoder
Non MPEG decoder
Constrained Buffer Resources
Constrained Rendering Resources
Constrained decoding Resources
Constrained Network Resources
Traditional MPEG transport system
Transport Stack

Network Rendering Example: Flattening Complex 6DoF





Recap

Recap

What is the state of the art for AR capabilities?
What are the device requirements for the optimal user experience?

Latency, Resolution, Frame Rate,
Video Decode, Power, Connectivity

What is expected to hit the market over the next few years?

New Displays, Neural Network Processors,
VR Hardware Blocks, Connectivity, Immersive Streams

What does the device architecture look like?

XR Optimized Hardware Blocks

Where do you see needs for standards?

Application / Device APIs (Khronos OpenXR)
Immersive Streaming Content (MPEG-I)
Low Latency Networks (3GPP)

Thank you

Follow us on:

For more information, visit us at:

www.qualcomm.com & www.qualcomm.com/blog

All data and information contained in or disclosed by this document is confidential and proprietary information of Qualcomm Technologies, Inc. and/or its affiliated companies and all rights therein are expressly reserved. By accepting this material the recipient agrees that this material and the information contained therein is to be held in confidence and in trust and will not be used, copied, reproduced in whole or in part, nor its contents revealed in any manner to others without the express written permission of Qualcomm Technologies, Inc. Nothing in these materials is an offer to sell any of the components or devices referenced herein.

©2016 Qualcomm Technologies, Inc. and/or its affiliated companies. All Rights Reserved.

Qualcomm and Snapdragon are trademarks of Qualcomm Incorporated, registered in the United States and other countries. Other products and brand names may be trademarks or registered trademarks of their respective owners.

References in this presentation to "Qualcomm" may mean Qualcomm Incorporated, Qualcomm Technologies, Inc., and/or other subsidiaries or business units within the Qualcomm corporate structure, as applicable. Qualcomm Incorporated includes Qualcomm's licensing business, QTL, and the vast majority of its patent portfolio. Qualcomm Technologies, Inc., a wholly-owned subsidiary of Qualcomm Incorporated, operates, along with its subsidiaries, substantially all of Qualcomm's engineering, research and development functions, and substantially all of its product and services businesses, including its semiconductor business, QCT.