3GPP SA4 Rel-18 e-Workshop
ITT4RT and possible next steps for RL18

Source:
KPN N.V., Nokia Corporation

Aim of presentation:

Give input of current status of ITT4RT and how to extend the current normative work in RL18 (i.e. ITT4XR, Immersive Teleconferencing and Telepresence for Extended Reality). Most importantly the presentation is to trigger discussion and input from interested parties to collaborate on a new RL18 work item starting beginning of 2022.

3GPP ITT4RT vision / Timeline

- RL17 – ITT4RT – Current Status
  - 3DOF – 360-degree
  - Single direction communication streams
  - OMAF and Overlays
- RL18 – ITT4XR – Next Steps
  - 3DOF+ and Limited AR
  - Spatial metadata
  - Bi-directional communication streams
  - Volumetric encoding
  - IVAS / Full Spatial audio
- RL19 – ITT4-6G – Future
  - 6DOF – True holographic communication
  - Complex AR Feature (i.e. remote processing and rendering)

Status ITT4RT (TS 26.114)

- Immersive 360-Degree Video Support (incl., fisheye, still background, viewport depend delivery, camera calibration for network-based stitching)
- Immersive Voice / Audio Support (currently EVS)
- Overlay Support
- Media Transport (RTP / RTCP)

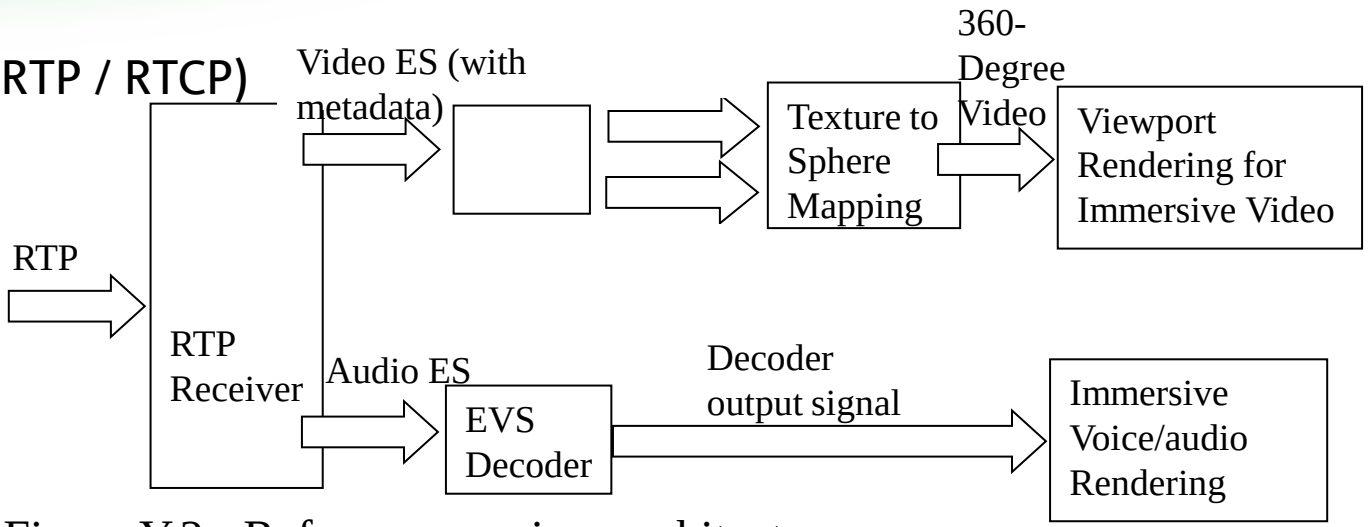


Figure Y.2 - Reference receiver architecture

New XR Communication Req. – 5GXR

Many requirements from 5GXR study (TR 26.928) not addressed in ITT4RT

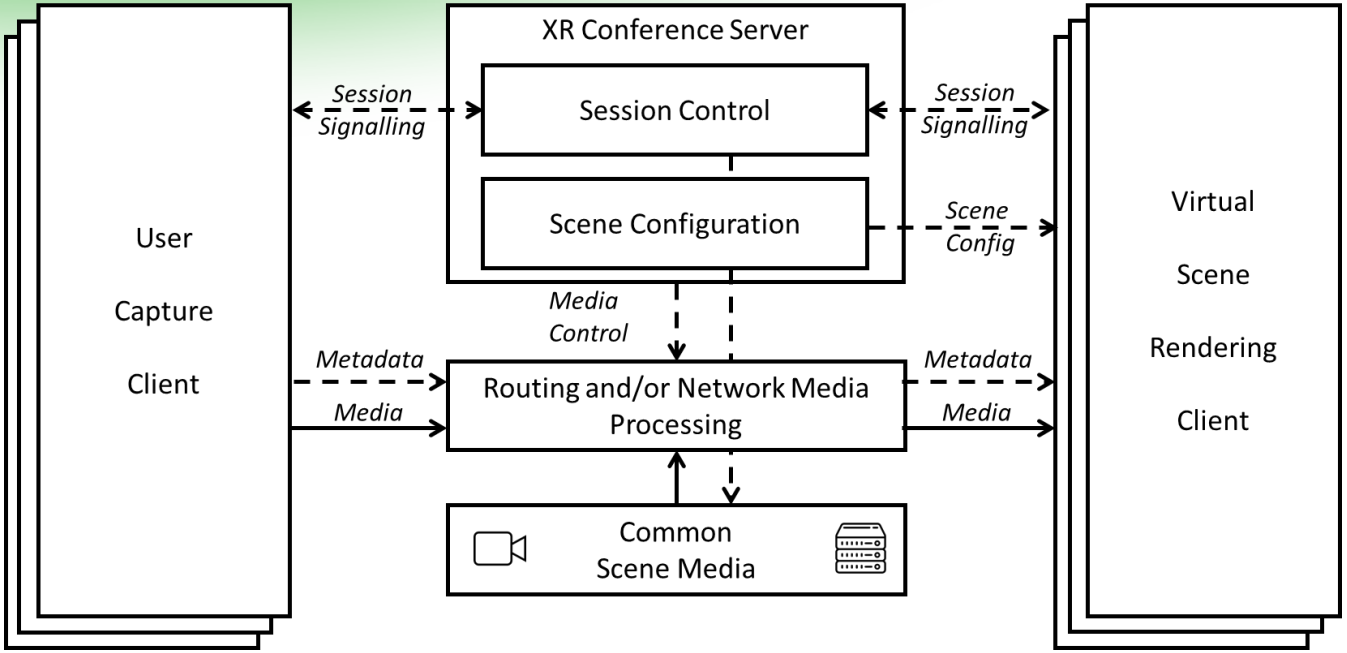
- 5.7.3 Potential Normative Work (XR Conference)
  1. Position and scaling of people
  2. Background (spatial) audio / picture / video
  3. media processing in the network/edge (e.g., cloud rendering, foreground/background segmentation of the user capture, replace HMD of the user with a photo-realistic representation of their face, calibration and synchronization of different camera capture data, etc.)
  4. Support for RGB+Depth video data (in signalling profiles, transmission and associated metadata)
  5. DOF metadata framework and a 6DOF capable renderer for immersive voice and audio:
     - including the Spatial Audio orientation of a speaker
     - 360 image metadata for associated audio
  6. appropriate codec to address XR conversational audio aspects (including directionality of users) and related metadata (e.g. IVAS)
  7. Synchronized and share state of the shared XR scene or space
- 5.8.3 Potential Normative Work (Spatial Audio Multiparty Call)
  - Visual coding and transmission of avatars
  - Coding of cut-out heads, alpha channel coding
  - Audio coding and transmission of mono objects and 3D audio for streams from all participants.

New XR Communication Req. – 5GSTAR
New requirements from 5GSTAR study (TR 26.998)

- New requirements from 5GSTAR study (TR 26.998)
  - A.4 Use Case 19: AR Conferencing
    - Standardized formats for 3D volumetric representation of participants on AR glasses.
    - Cloud APIs for processing and rendering of 3D volumetric streams.
    - Conversational methods for call initiation.
    - Spatial audio formats and associated metadata.
    - Metadata for Spatial characteristics of the AR environment (e.g. positioning of users).
  - …

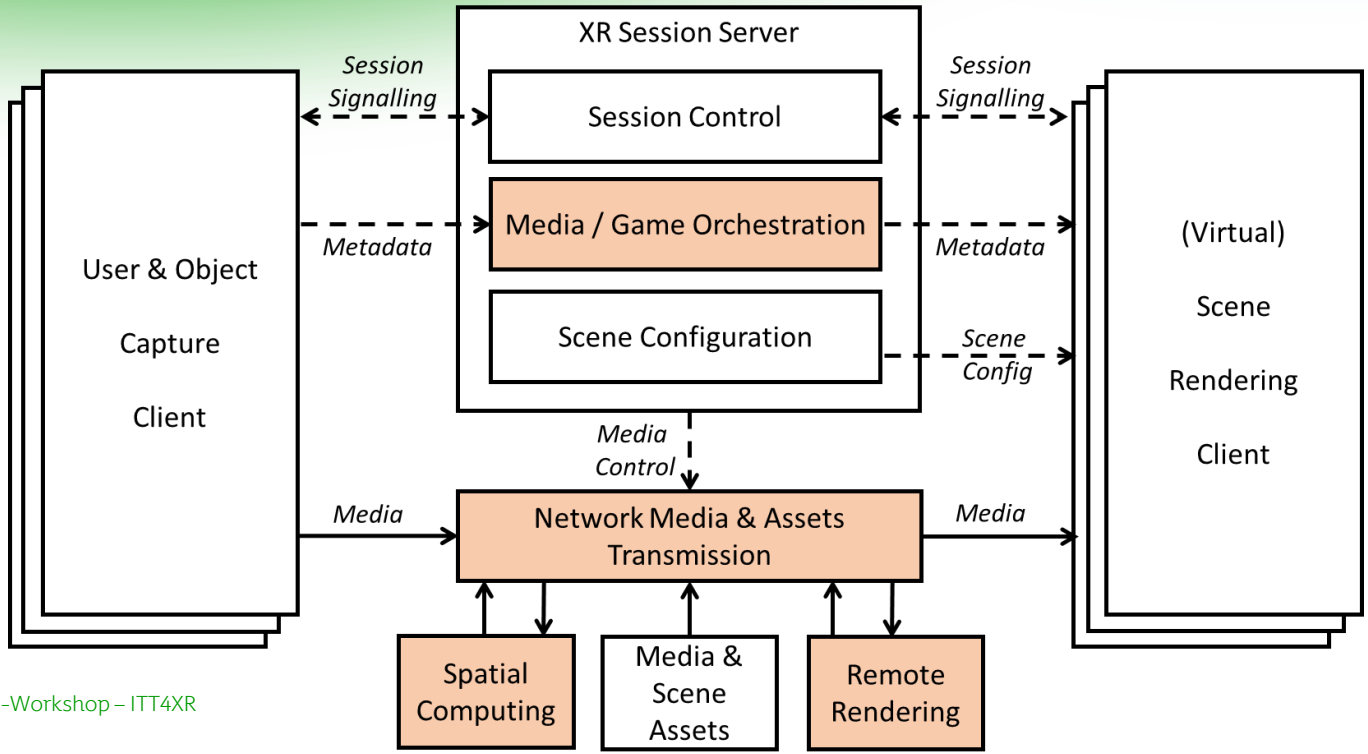
Simplified Architecture (current focus VR)

TR 26.928: Figure 6.2.8-1 General architecture for XR conversational and conference services



Possible Extension (simplified Architecture)

IBC Paper 2021 IMMERSIVE EXPERIENCES AND XR: A GAME ENGINE OR MULTIMEDIA STREAMING PROBLEM?



Timeline new WID – ITT4XR

- November 2021 (SA4#116-e)
  - **1st draft of WID**
  - Gather core use case
  - Define Scope and Requirements
- February 2022 (SA4#117)
  - **Finalize WID and SA4 approval**
  - **Finish ITT4RT**
- March 2022 (SA#95)
  - **SA approval of the new ITT4XR WID**
- April 2022 (SA4#118-e)
  - **Start ITT4XR**