Immersive Media Meet(s) 5G workshop; 15 April 2019

MPEG-I
"I" is for Immersive

Two-part Presentation

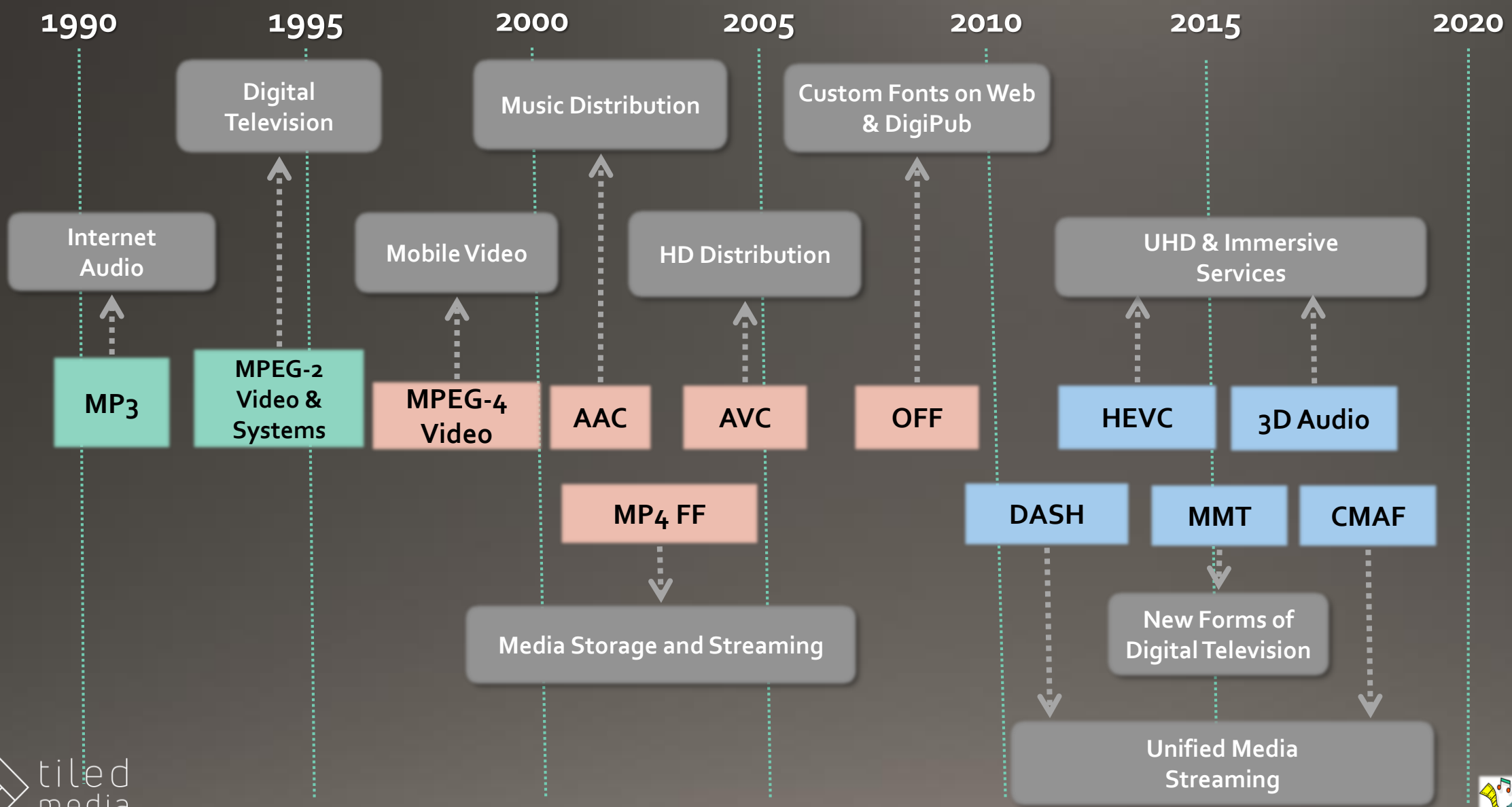
- Overview of MPEG work for Immersive Media
  - Rob Koenen, Co-Founder and CBO Tiledmedia
  - Co-Founder and first President of VRIF

- Zooming in on Point Cloud Compression
  - Danillo Graziosi
  - Manager of Next-Generation Codec at US Research Center, Sony Corporation of America

tiled media

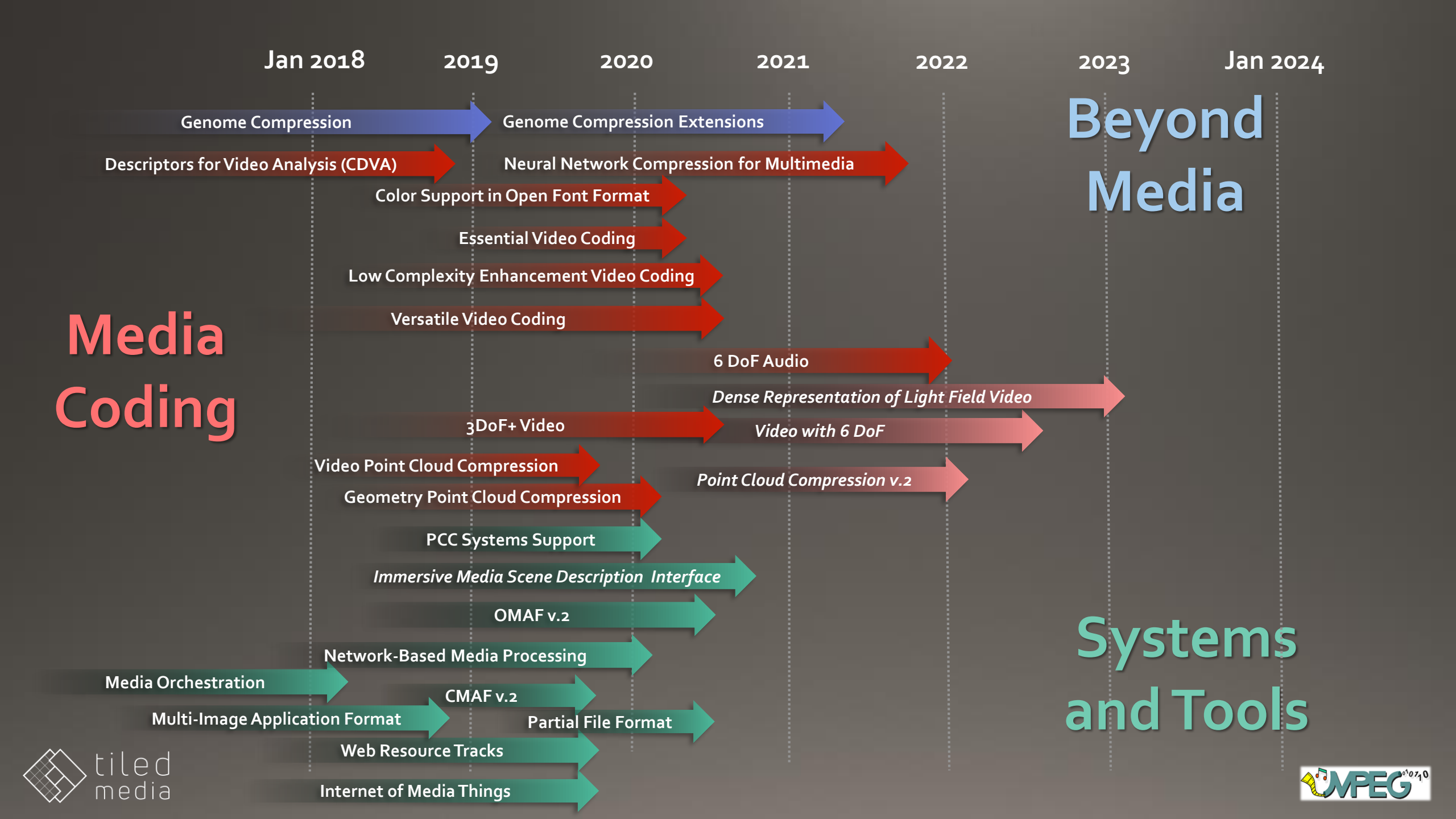
MPEG

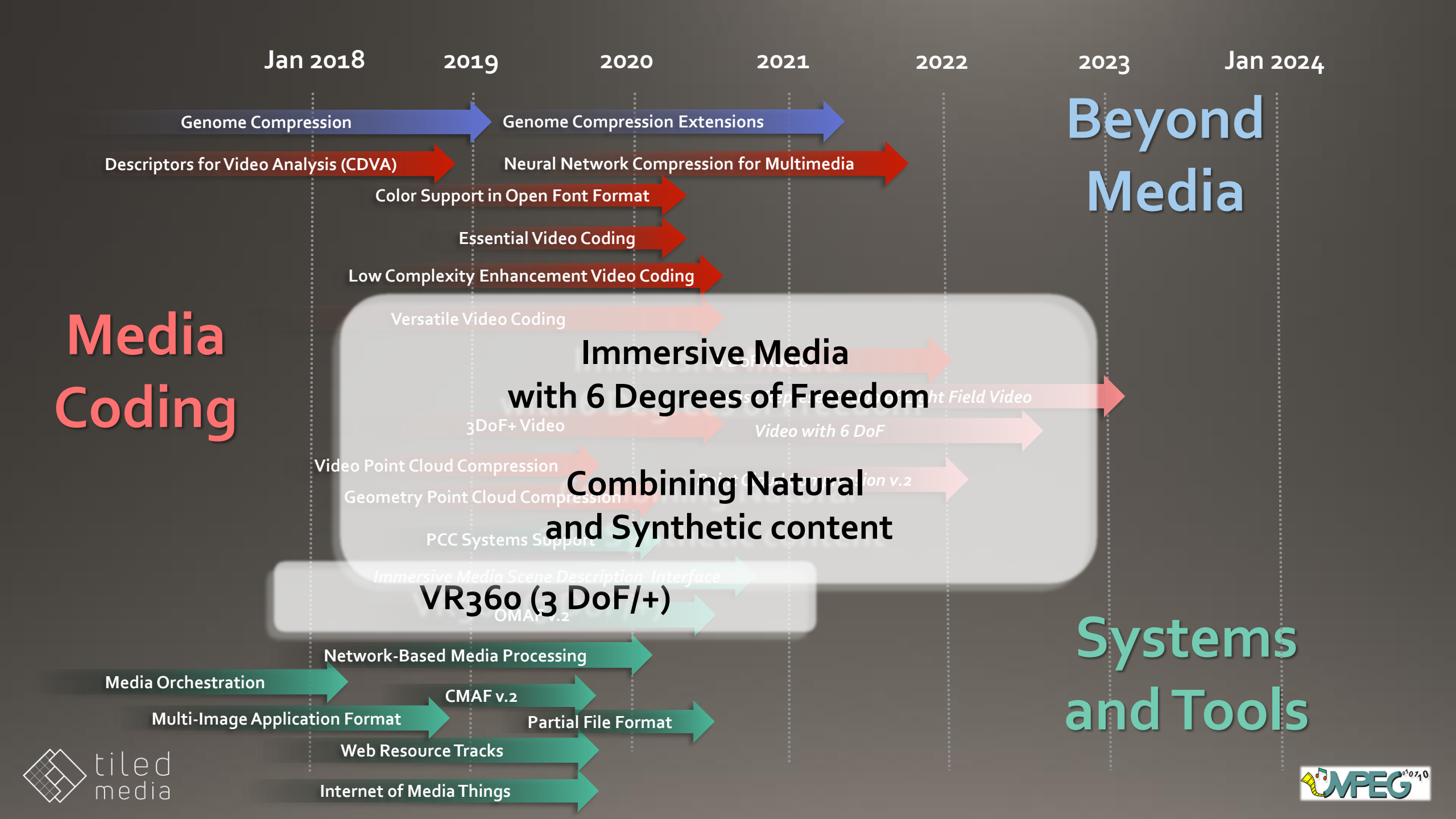
Major MPEG Standards (For Reference)

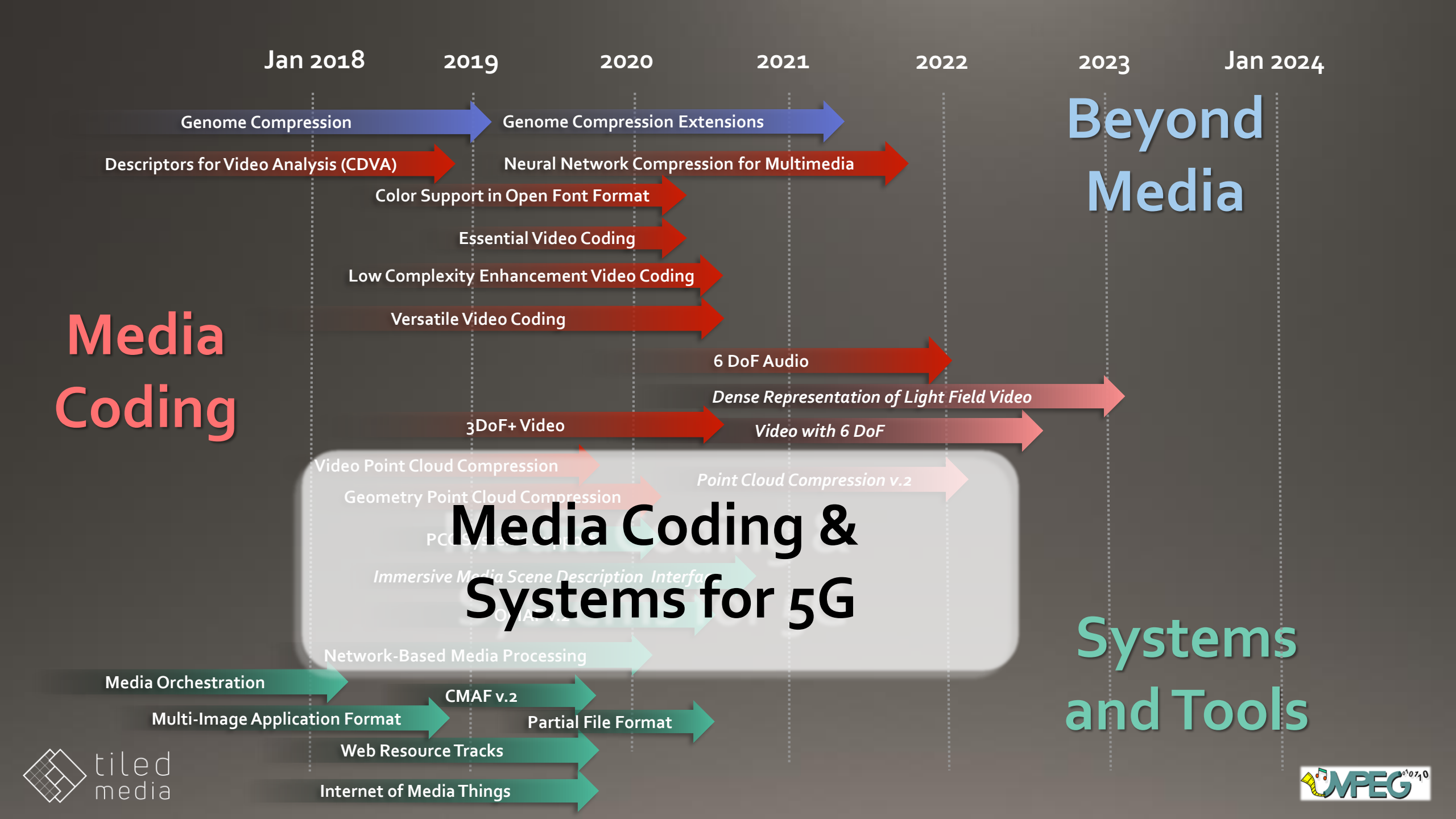


tiled media

MPEG


Jan 2018
2019
2020
2021
2022
2023
Jan 2024
Beyond Media
Media Coding
Systems and Tools
Genome Compression
Genome Compression Extensions
Descriptors for Video Analysis (CDVA)
Neural Network Compression for Multimedia
Color Support in Open Font Format
Essential Video Coding
Low Complexity Enhancement Video Coding
Versatile Video Coding
6 DoF Audio
Dense Representation of Light Field Video
3DoF+ Video
Video with 6 DoF
Video Point Cloud Compression
Point Cloud Compression v.2
Geometry Point Cloud Compression
PCC Systems Support
Immersive Media Scene Description Interface
OMAF v.2
Network-Based Media Processing
Media Orchestration
CMAF v.2
Multi-Image Application Format
Partial File Format
Web Resource Tracks
Internet of Media Things
tiled media
MPEG


Jan 2018
2019
2020
2021
2022
2023
Jan 2024
Beyond Media
Genome Compression
Genome Compression Extensions
Descriptors for Video Analysis (CDVA)
Neural Network Compression for Multimedia
Color Support in Open Font Format
Essential Video Coding
Low Complexity Enhancement Video Coding
Media Coding
Versatile Video Coding
Immersive Media with 6 Degrees of Freedom
3DoF+ Video
Video with 6 DoF
Video Point Cloud Compression
Geometry Point Cloud Compression
Combining Natural and Synthetic content
PCC Systems Support
Immersive Media Scene Description Interface
VR360 (3 DoF/+)
OMAF v.2
Network-Based Media Processing
Media Orchestration
CMAF v.2
Multi-Image Application Format
Partial File Format
Web Resource Tracks
Internet of Media Things
Systems and Tools
tiled media
MPEG


Jan 2018
2019
2020
2021
2022
2023
Jan 2024
Genome Compression
Genome Compression Extensions
Descriptors for Video Analysis (CDVA)
Neural Network Compression for Multimedia
Color Support in Open Font Format
Essential Video Coding
Low Complexity Enhancement Video Coding
Versatile Video Coding
6 DoF Audio
Dense Representation of Light Field Video
3DoF+ Video
Video with 6 DoF
Video Point Cloud Compression
Point Cloud Compression v.2
Geometry Point Cloud Compression
Immersive Media Scene Description Interfa
Network-Based Media Processing
Media Orchestration
CMAF v.2
Multi-Image Application Format
Partial File Format
Web Resource Tracks
Internet of Media Things
Beyond Media
Media Coding
Media Coding & Systems for 5G
Systems and Tools
tiled media
MPEG

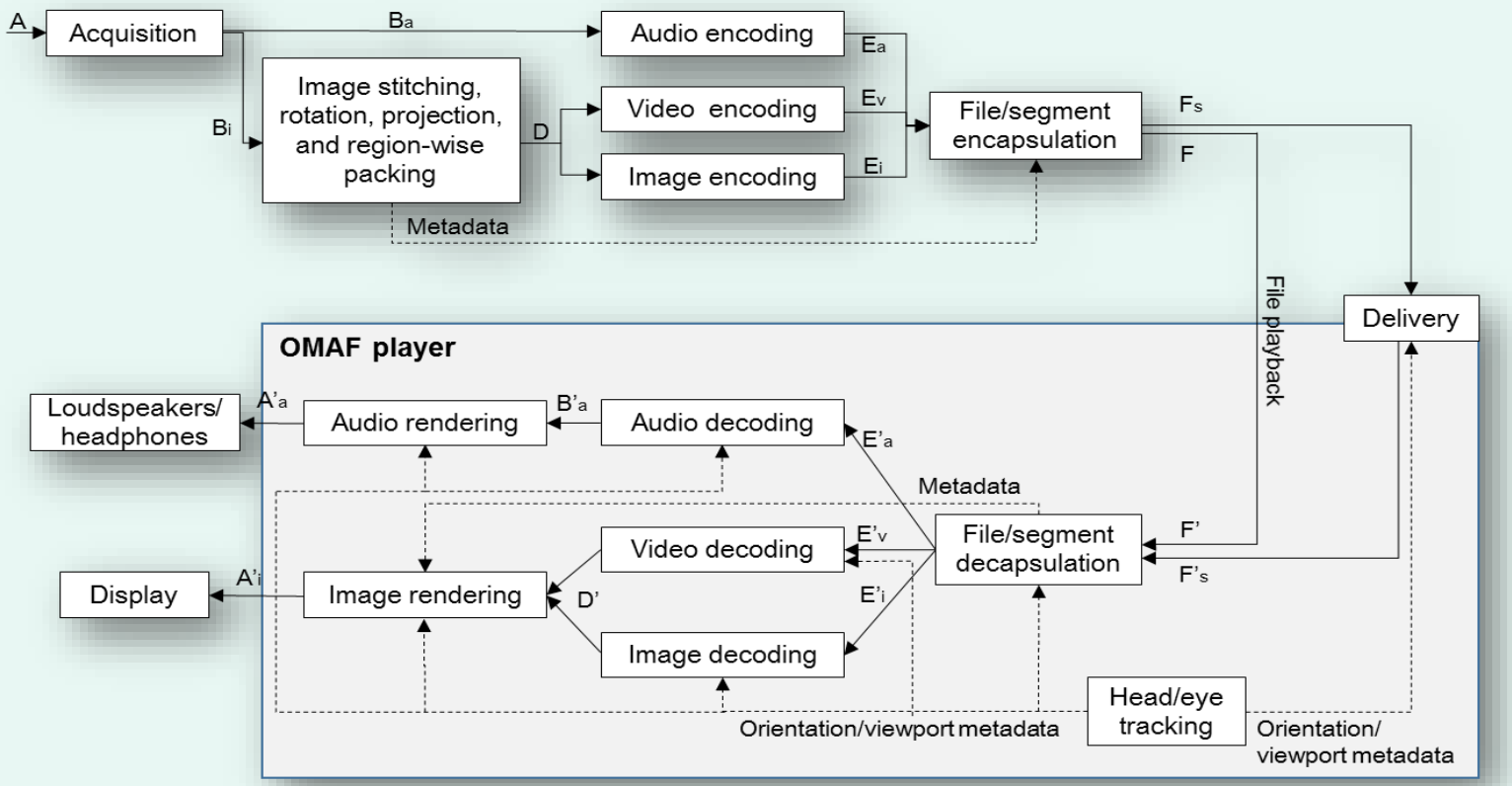
MPEG-I: ISO/IEC 23090

Coded Representation of Immersive Media

11 Parts (and Counting …)

1. Architectures for Immersive Media (Technical Report)
2. Omnidirectional Media Application Format
3. Versatile Video Coding
4. 6 Degrees of Freedom Audio (name t.b.d.)
5. Video-Based Point Cloud Coding (V-PCC)
6. Immersive Services and Applications
7. Immersive Media Metadata
8. Network-Based Media Processing
9. Geometry Point Cloud Coding (G-PCC)
10. Carriage of Point Cloud Data
11. Implementation Guidelines for Network-based Media Processing

OMAF - Omnidirectional MediA Format



tiled media

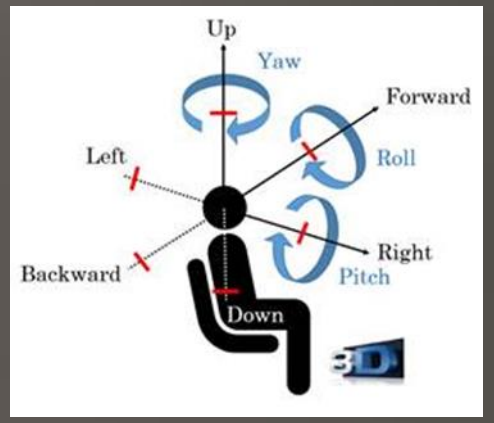
MPEG

2. OMAF - Delivery

- Encapsulation in ISO Base Media File Format
  - Adding timed text
- Transport
  - DASH and MMT
- Viewport-*independent* (or -*agnostic*) streaming
  - just send everything, no matter where the viewer looks
- Viewport-*dependent* streaming
  - Send viewport with better quality

2. OMAF v.2

- 3 DoF+: allowing a bit of lateral head motion



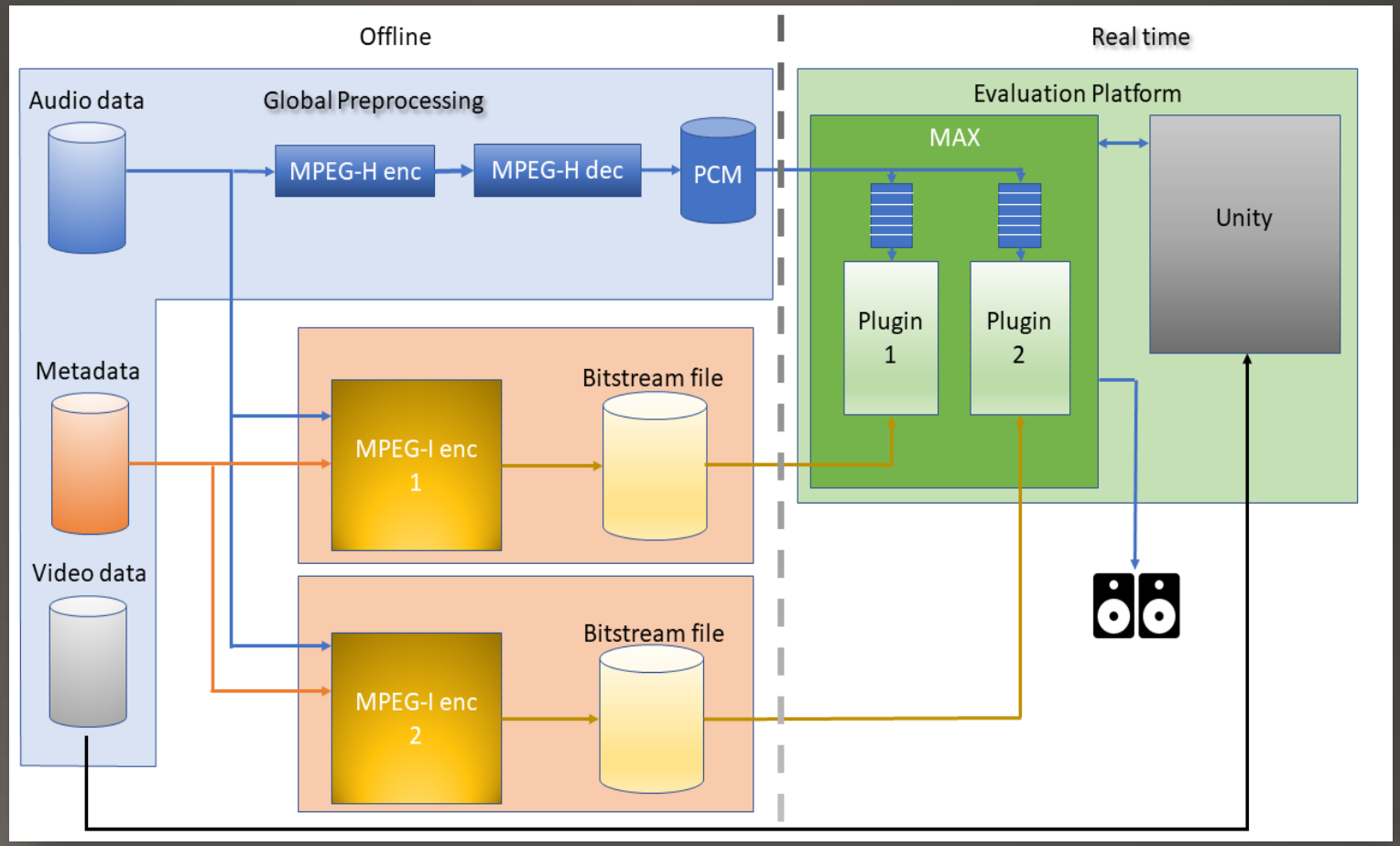
- Interactivity and Overlays
- Better viewport-adaptive streaming
  - "Late Binding" – allows for more flexibility and more sophisticated clients

3. Versatile Video Coding

- Next Generation Video Coding Standard
- Ready mid-2020
- 40% - 50% better than HEVC
- Useful for ever-larger media, including point clouds

4. Audio with 6 Degrees of Freedom

- Streaming Audio Objects + Model of Environment
- Goal: Define Immersive Rendering of Audio

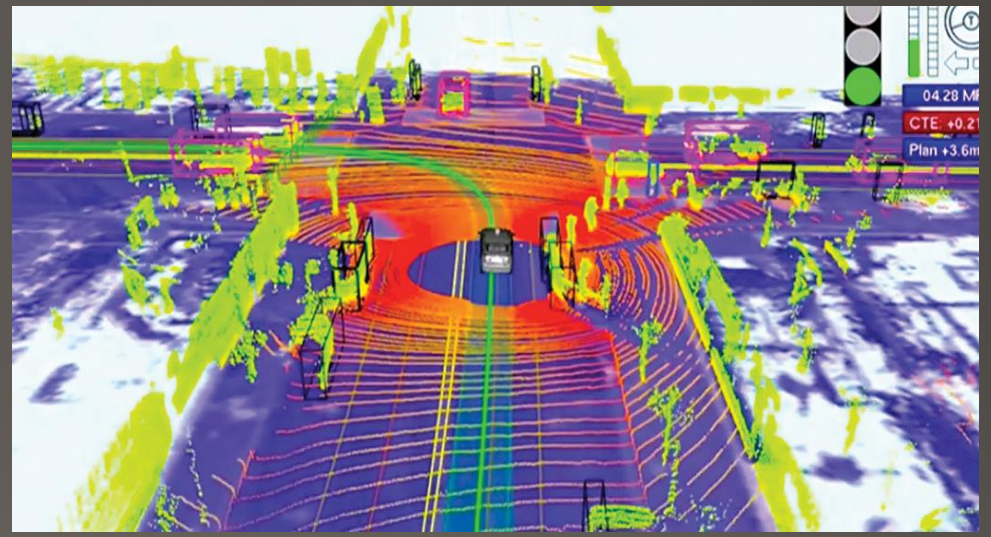


- Thanks to Philips for graphic

tiled media

Point Cloud Compression

- 5. Video-based
  - Dense / continuous point clouds
  - See Danillo's Talk!
- 9. Geometry-based
  - Sparse point clouds (e.g., Lidar scans)
  - 10:1 for lossless
    30:1 acceptable lossy

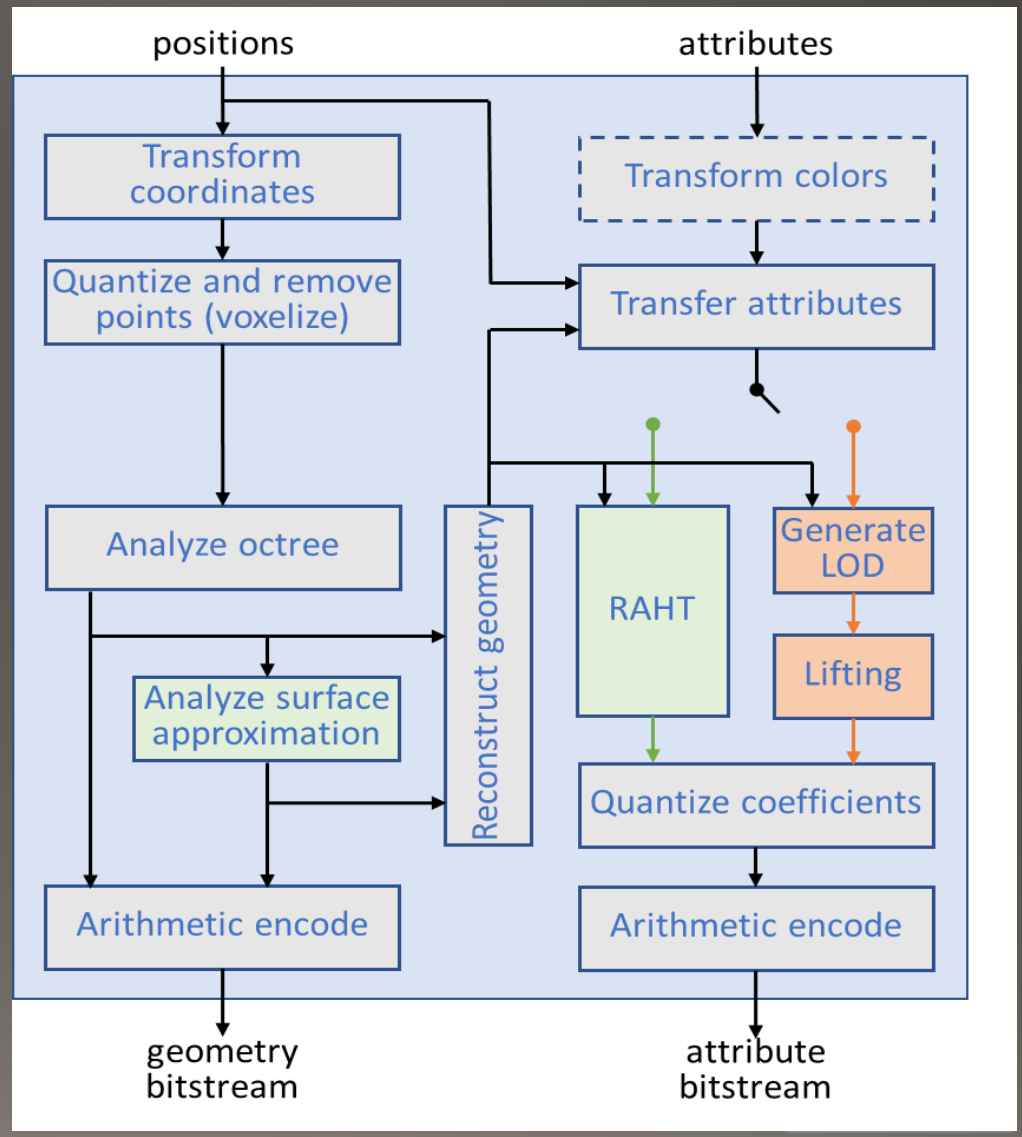


tiled media

MPEG

Point Cloud Compression

- 5. Video-based
  - Dense / continuous point clouds
  - See Danillo's Talk!
- 9. Geometry-based
  - Sparse point clouds (e.g., Lidar scans)
  - 10:1 for lossless
    30:1 acceptable lossy



tiled media

8. Network Based Media Processing

- Using the network and the edge to support media processing
  - Network-based, last-second media personalization
  - Video Stitching
  - AR rendering
  - Viewport extraction and encoding
  - Social VR Support
- Functions that could be carried out in a 5G Edge server

- Example: SK Telecom's 5G AR (23 March 19, Korean Baseball Game)
  - https://m.sports.naver.com/video.nhn?id=523565

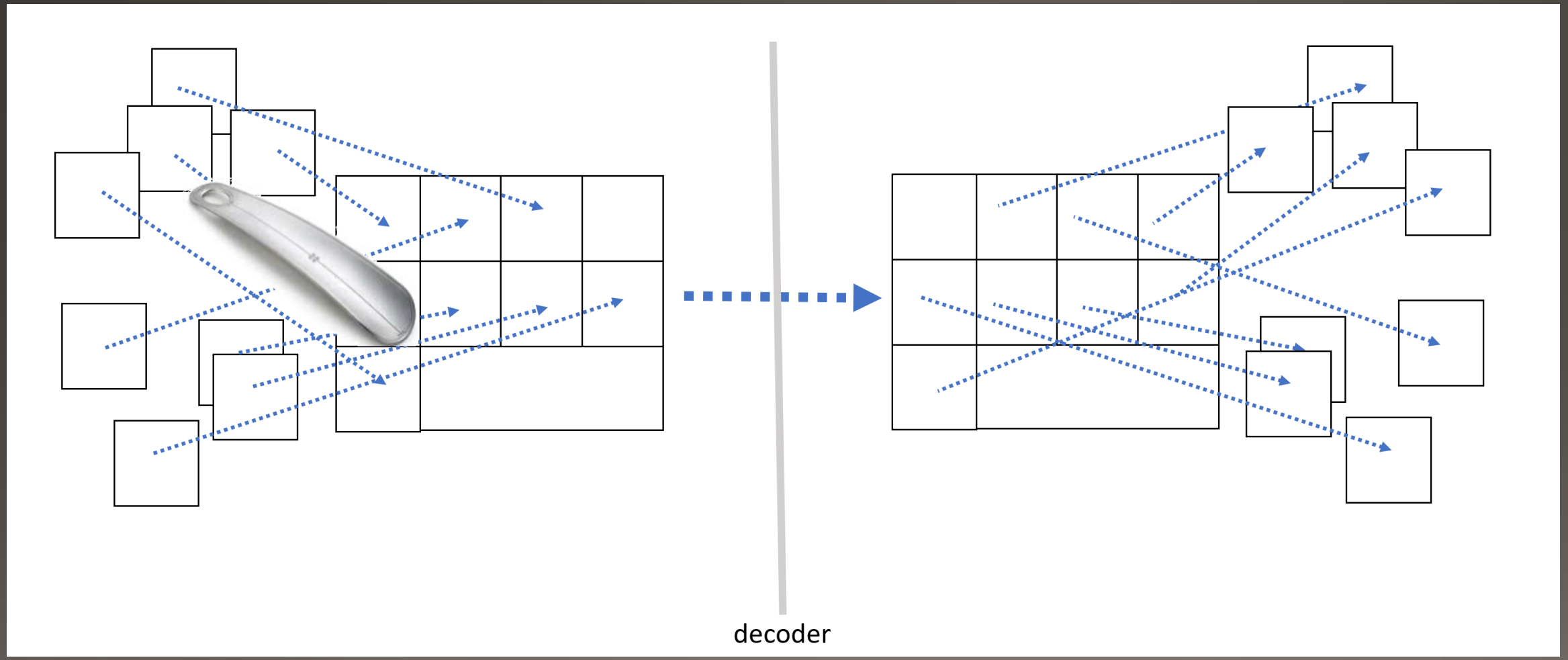
tiled media

MPEG

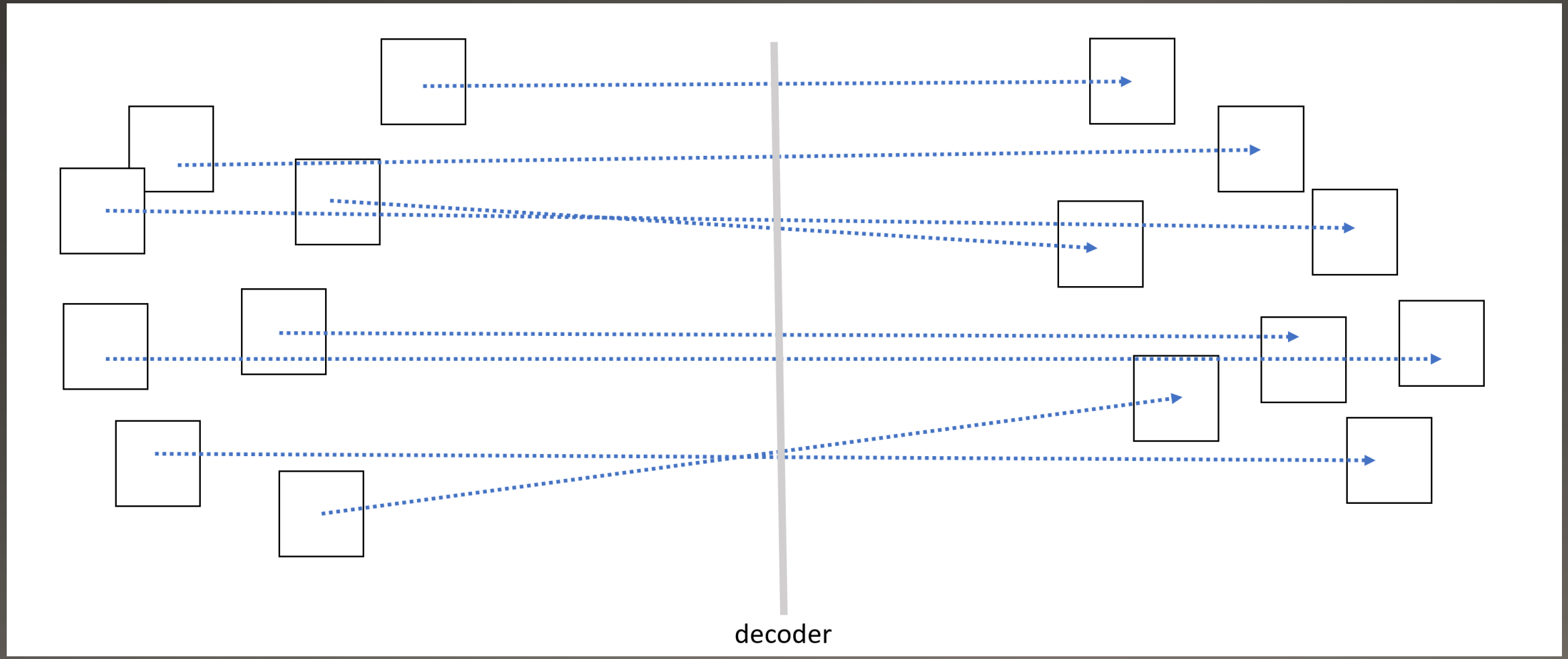
Immersive Media Access and Delivery

- "Exploration" – Design Considerations
- Larger media sizes lead to smaller access units!
- Partial retrieval of partially visible and audible scenes
  - Requires ultra-fast network response
- Edge Rendering, or Split Rendering
  - Heavy lifting in network, save on processing in mobile devices
  - Very sophisticated media with reasonable device complexity and power use
- New Decoder Architectures for Immersive Media
  - VR, Point Clouds, Hybrid Scenes

Towards Object-Based Decoding



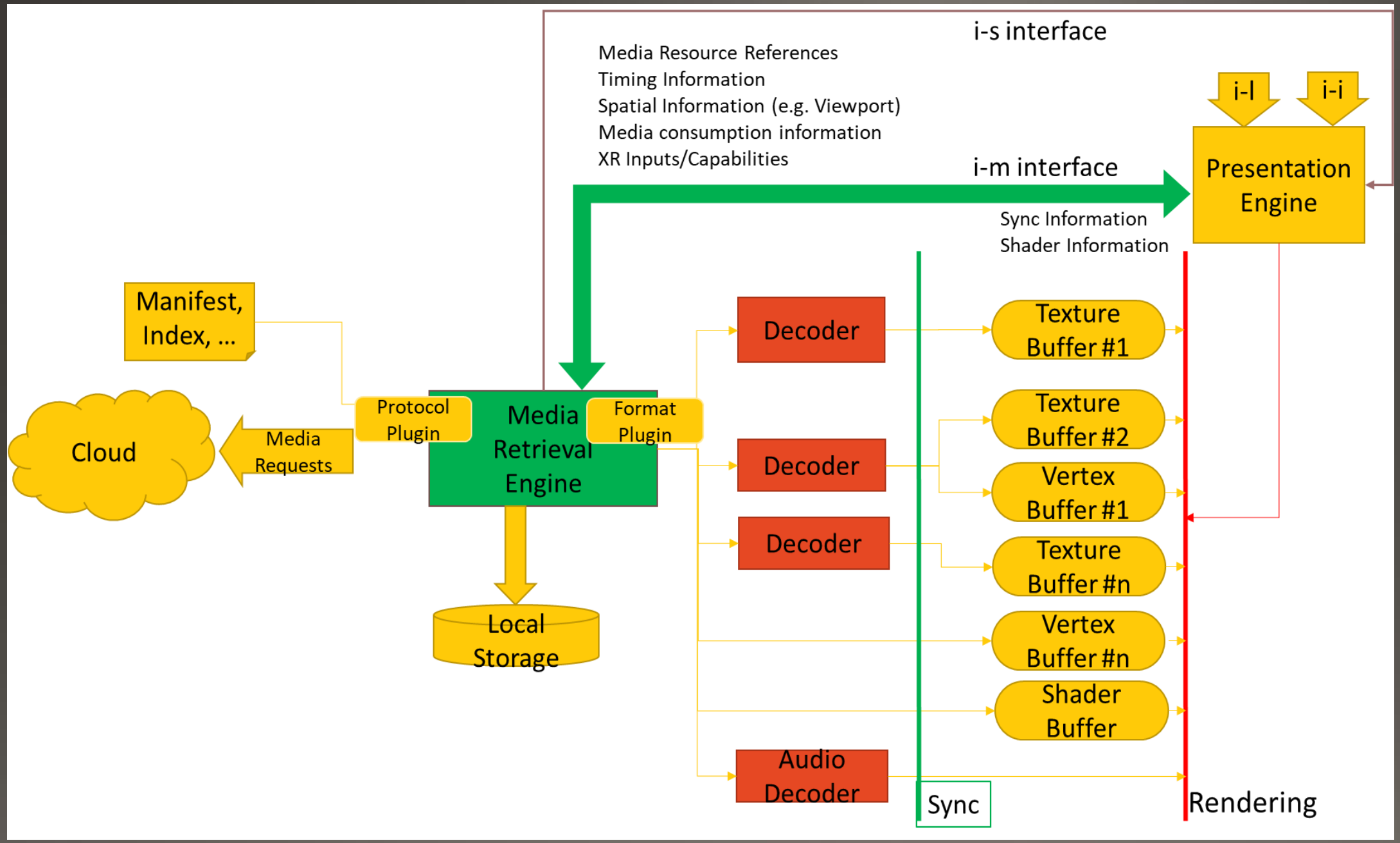
Towards Object-Based Decoding



Architectures and Scene Description

- Also “Exploration”

- Considering Existing Scene Description Formats
- Ideally merge these with timed media
  - Merging Web and AV Distribution paradigms proves hard


i-s interface
Media Resource References
Timing Information
Spatial Information (e.g. Viewport)
Media consumption information
XR Inputs/Capabilities
i-l
i-i
Presentation Engine
i-m interface
Sync Information
Shader Information
Manifest, Index, ...
Cloud
Media Requests
Protocol Plugin
Media Retrieval Engine
Format Plugin
Local Storage
Decoder
Decoder
Decoder
Audio Decoder
Texture Buffer #1
Texture Buffer #2
Vertex Buffer #1
Texture Buffer #n
Vertex Buffer #n
Shader Buffer
Sync
Rendering

Thanks!

tiled media

MPEG