



GPAC
Licensing

SubTech 1

Symposium on Subtitling Technology

Subtitles in MP4 and DASH
Report from the MP4Box open-source team

Subtitles: historic usage

- ▶ Historic usages
 - ▶ Translation : different languages
 - ▶ Accessibility aspect:
 - ▶ Close captions
 - ▶ Subtitles for the deaf or hard-of-hearing (SDH)
 - ▶ Same-Language Subtitling (karaoke, presentation)

Subtitles: emergent usages

- ▶ Emergent usages
 - ▶ Watching a video without the audio on (eg. In public places, bars or on youtube, facebook)
 - ▶ - FB: subtitles increase the viewer engagement (40%)
 - ▶ Help for a better indexation of the video with search engines
 - ▶ Subtitles on 360 video (more on ImAc later)

Proliferation of subtitle formats (before TTML)



Software video players

- AQTitle
- JACOSub
- Gloss Subtitle
- MPSub
- Ogg Writ
- RealText
- VobSub
- SAMI
- SubViewer
- Phoenix Subtitle
- XSUB
- Structured Subtitle Format
- SubStation Alpha
- Universal Subtitle Format
- PowerDivX
- MPEG-4 Timed Text

• DVD Video

• MicroDVD

• Blu-ray Disc

DVD video

• D-Cinema

Cinéma

- DVB Subtitles
- Philips Overlay Graphics Text
- Imtext
- Teletext
- .ESY .XIF .X32 .PAC .RAC .CHK
- .AYA .890 .CIP .CAP .ULT .USF
- .CIN .L32 .ST4 .ST7 .TIT .STL

TV Broadcast

• SubRip

• Spruce subtitle format

• SMIL

• WebVTT

Internet Delivery

TTML: a new standard

- ▶ There are too many standards
 - ▶ Specify one standard that covers all existing standards
 - ▶ TTML : **one standard to rule them all !**
- ▶ New standards creation:



Proliferation of subtitle formats (after TTML)

GPAC
Licensing

Software video players

- AQTitle
- JACOSub
- Gloss Subtitle
- MPSub
- Ogg Writ
- RealText
- VobSub
- SAMI
- SubViewer
- Phoenix Subtitle
- XSUB
- Structured Subtitle Format
- SubStation Alpha
- Universal Subtitle Format
- PowerDivX
- MPEG-4 Timed Text

• DVD Video

• MicroDVD

• Blu-ray Disc

DVD video

• D-Cinema

Cinéma

- DVB Subtitles
- Philips Overlay Graphics Text
- Imtext
- Teletext
- .ESY .XIF .X32 .PAC .RAC .CHK
- .AYA .890 .CIP .CAP .ULT .USF
- .CIN .L32 .ST4 .ST7 .TIT .STL

TV Broadcast

• SubRip

• Spruce subtitle format

• **SMPTTE-TT** • **TTML (ex DFXP)**

• **EBU-TT** • **CFF-TT** • **SDP-US**

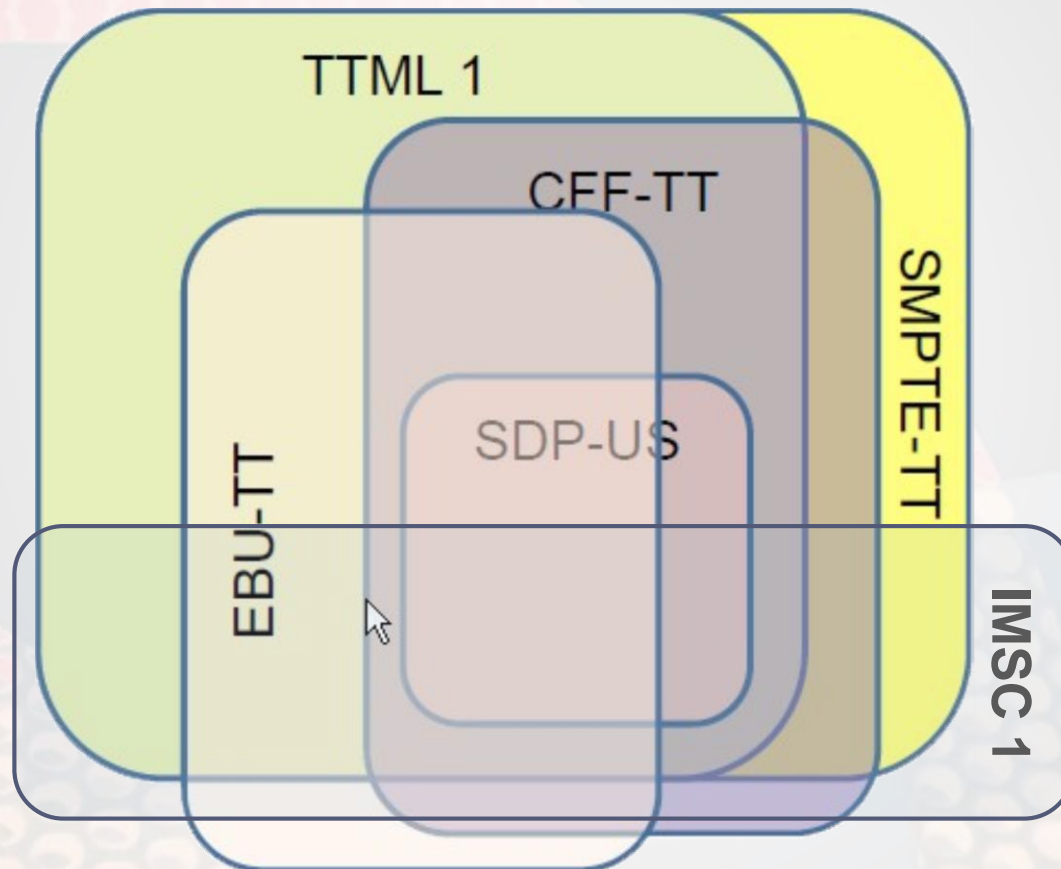
• SMIL

• WebVTT

Internet Delivery

TTML and derived profiles/captions formats

- ▶ Started in 2004.
- ▶ The Timed Text format is a Recommendation of [W3C](#) (2010)




What is GPAC ?

- ▶ GPAC is a multimedia framework started in 1999
 - ▶ Code became free and open-source in 2003
 - ▶ Used by major services (Youtube, Sony, IRT, ...)
 - ▶ Powered by a community of enthusiasts: x264, ...
 - Started a lot of companies
- ▶ **Softwares**
 - ▶ MP4Box packager
 - ▶ MP4Client player
 - ▶ Many other tools (mostly MP4/MPEG2-TS, DASH, live tools...)
- ▶ **A team**
 - ▶ Research and standards @Telecom ParisTech
 - ▶ A commercial arm: GPAC Licensing (SME) in 2012



GPAC and subtitles: a love story

- ▶ First subtitle support: MPEG-4 Timed Text (2001)
- ▶ Industry-grade conformant open-source tools
 - ▶ GPAC as a utility software at MPEG
 - ▶ Help the industry to produce better formats (GPAC Licensing):
 - ▶ [Guidelines to embed TTML in MP4 and DASH](#) (with IRT, BBC, EBU, TPT)
- ▶ ImAc H2020 project (Immersive Accessibility) 
 - ▶ Accessibility (subtitles, audio description, sign language) in VR
 - ▶ Push for “accessibility as a first-class citizen”
 - ▶ Consortium includes companies (Motion Spell, Anglatecnic), user associations (RNIB), broadcasters (RBB, CCMA) and academic partners (IRT, I2cat, UAB, USAL).

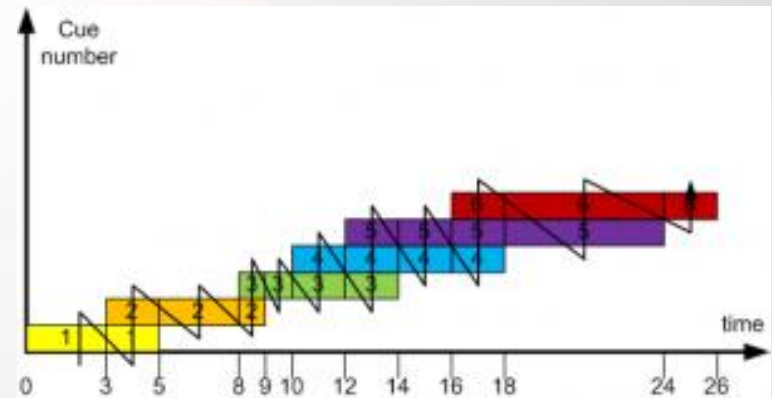
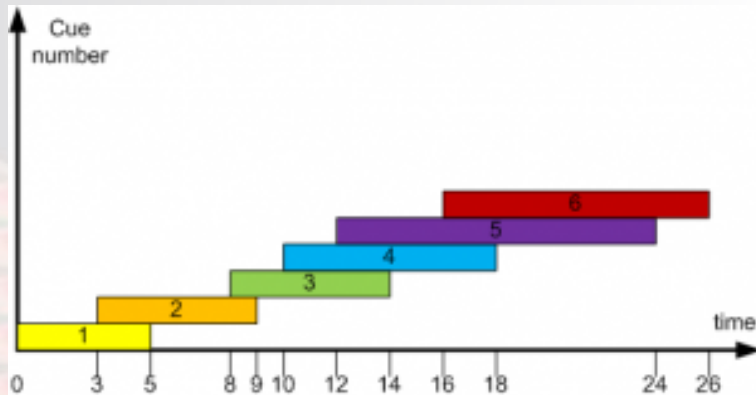
Subtitles and GPAC support

▶ ISOBMFF vision:

- ▶ Any human readable text (subtitles, closed captions...) is a subtitle.
- ▶ 2 major classes of formats :
 - ▶ ‘text ’: text processing capabilities only
 - e.g WebVTT
 - ▶ ‘subt’: text and image processing capabilities
 - e.g. TTML
- ▶ WebVTT and TTML in ISOBMFF
 - ▶ MPEG-4 part 30:
 - Title: “Timed Text and other visual overlays”
 - Standardized in 2014-03
 - Co-edited by D. Singer, M. Dolan and C. Concolato from GPAC

WebVTT support in GPAC

- ▶ First integration in Sept 2012
- ▶ Completed in July 2013 (import, muxing, dashing)
- ▶ Conform to ISOBMFF spirit:
 - ▶ Metadata are separated from the data
 - ▶ Output file is smaller than input
- ▶ Challenge for adaptive streaming segmentation:



<https://concolato.wp.imt.fr/2012/09/12/webvtt-streaming/>

TTML support in GPAC (1/2)

- ▶ TTML support in GPAC:
 - ▶ Started with NHML generic support in Jan 2014
 - ▶ MPEG-4 Part 30 is ok (see later)
 - ▶ GPAC Licencing commissioned by DASH-IF for EBU-TT-D initial support: Aug 2014
 - ▶ Some fixes with IRT about HbbTV in 2015

- ▶ New profiles keep appearing:
 - ▶ Faster than deployed.
 - ▶ Faster than we implement them.
 - ▶ IMSC will be the new standard to rule them all !!
 - ▶ IMSC 1 + text & image profiles
 - ▶ IMSC 2 is coming soon ?

TTML support in GPAC (2/2)

- ▶ Need to parse deeply to extract the useful information (timings, regions)
 - ▶ which then need to be semantically processed...
- ▶ TTML is extended from XML:
 - ▶ Heavy to handle properly in native C code (FFmpeg, GPAC, VLC, ...): deep hierarchy tree, namespaces, extensibility ...
- ▶ Redundancy
 - ▶ Same functionality described in many ways (timings, positions, etc.)
 - ▶ Same drawback as MPEG-DASH
- ▶ These are pain points for both muxers and players implementors.

TTML playback support in GPAC

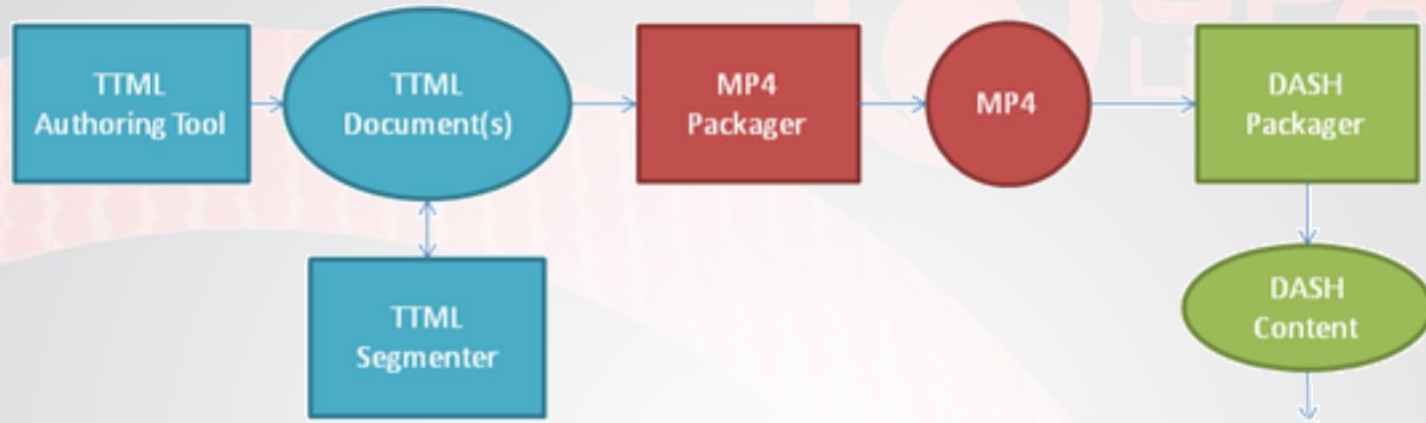
- ▶ No playback.
- ▶ Parsing is incomplete in GPAC.
- ▶ Conformance model.
- ▶ Need an elaborate model for pivotal representation:
 - ▶ Ex: ISD (Intermediate Synchronic Document)
- ▶ Rendering seems to expect the presence of a CSS stack.
 - ▶ No user asked for it.
 - ▶ We gave up, waiting for an industry input.

Segmenting TTML in MP4



- ▶ MPEG-4 part 30 for TTML:
 - ▶ Put the TTML document in one ISOBMFF sample.
 - ▶ Operations on MP4 file (cutting, extracting) require to rewrite the TTML content.
 - ▶ First time we need to re-encode content when manipulating MP4s.
 - Give me a H267 or AC-5 stream in MP4 and I can cut, DASH and encrypt it. Not for TTML.
 - ▶ The muxer need to map ISOBMFF timeline with TTML timeline.
 - Solution: data shouldn't be repeated over the layers.

Segmenting TTML in MP4



- ▶ Find a solution with industry leaders (GPAC Licensing):
 - ▶ [Guidelines to embed TTML in MP4 and DASH](#) (with IRT, BBC, EBU, TPT)
 - ▶ [Samples](#) (from Andreas Tai @ IRT)
- ▶ Authoring tools need to:
 - ▶ Prepare for segmentation (DASH).
 - ▶ Same as when video encoders communicate closely with muxer.
 - ▶ Add a pre-packager tools: no existing solution

Timeline issues for TTML streaming

- ▶ 3 possibilities, all deployed!
 - ▶ ISOBMFF « The Standard »
 - ▶ Time is relative to the media: streaming session start time is zero.
 - ▶ Not resilient
 - ▶ Smooth Streaming « The de-facto »
 - ▶ Each chunk/segment start time starts at zero.
 - ▶ Resilient
 - ▶ Non standard « The alternative »
 - ▶ Absolute UTC time.
 - ▶ Resilient
- ▶ Problems:
 - ▶ The packager needs to parse and rewrite the TTML
 - ▶ We have to maintain consistency between the TTML and ISOBMFF timelines
 - ▶ The player may need to guess the timeline

Thanks and acknowledgements

GPAC
Licensing



Horizon 2020
N°761974