

Closed captions

What GStreamer Can Do

2023 GStreamer conference

Matthew Waters

25 September 2023

Closed Captions

- What are they?
- Why we use them?
- Note: this talk will only be an overview of the landscape

Closed Captions

Will only talk about CTA-608 and CTA-708 (and a little bit of SMPTE 334)

Not Teletext, not WebVTT, not ...

CTA-608

- Also known as 'line 21 captions' or CEA-608 or EIA-608
- Legacy format but still used for backward compatibility in many places
- 960bits/s transport
- More details in <https://en.wikipedia.org/wiki/EIA-608> or the relevant laws

CTA-608 display modes

- Roll up
 - Carriage return moves text up smoothly, text outside the window is discarded
- Paint On
 - Incoming text is displayed as received
- Pop On
 - Incoming text is buffered in memory, then popped onto the screen with a single command

CTA-608 features

- Captions
 - 4 caption channels (generally only max 2 used due to bitrate constraints)
 - base row changes, cursor indentation
 - 7 font colours
 - underlining
 - white italics
- Text mode
- XDS
 - TV/Movie rating, Warnings, Current program information, Audio/Caption services available, Network information, Date/Time, etc

CTA-608 data format

- A pair of bytes per field
- Each byte has a parity bit
- Each byte pair is a command, 1 or 2 ascii-ish characters or 1 extended character

So, what can GStreamer do with CTA-608 captions?

Demo

- `sccparse`
- `cea608overlay`

This is a test caption

So, what can GStreamer do with CTA-608 captions?

Demo

- `mccparse`
- `cea608overlay`

This is a test caption

So, what can GStreamer do with CTA-608 captions?

Demo

- `cea608tott`
- `tttocea608`

So, what can GStreamer do with CTA-608 captions?

Warning: a non-exhaustive list

- Combine two CTA-608 streams (`cea608mux`)
- Parse and encode SCC files (`sccparse` and `sccenc`)
- Parse and encode MCC files (`mccparse` and `mccenc`)
- Convert CTA-608 to Timed Text and back (`cea608tott` and `tttocea608`)
- Convert CTA-608 to WebVTT (`cea608tojson` ! `jsontovtt`)

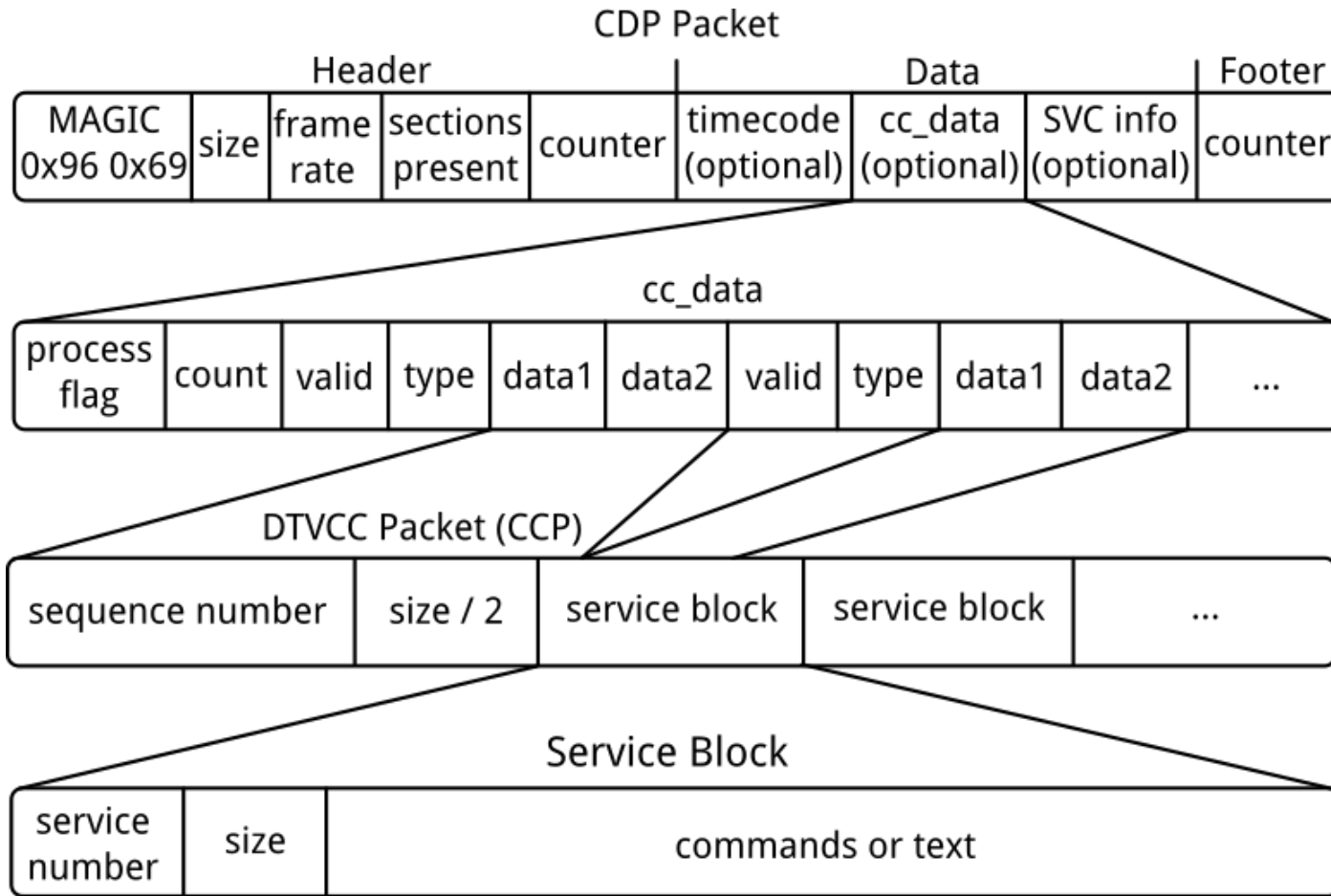
CTA-708

- Newer standard than CTA-608 designed for digital TV
- Contains CTA-608 compatibility section
- 9600bits/s transport
- More details in <https://en.wikipedia.org/wiki/CTA-708> and in the specification

CTA-708 features

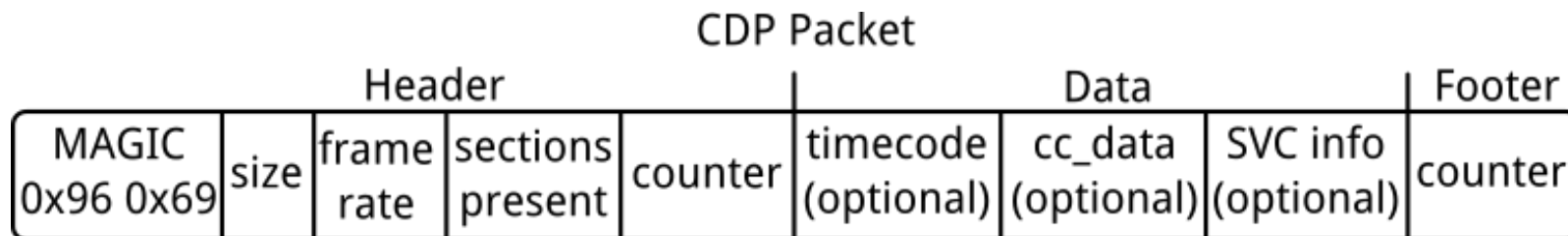
- Up to 8 windows
 - borders (64 colour, 5 styles), 64 background colour and 4 opacity, 3 transition effects, scroll and print direction, visible, columns and rows
- Pen
 - 64 colour background and foreground, 3 sizes, 8 fonts, super/subscript, italics, underline, 5 outline styles
 - Movement throughout the window

CTA-708 data format



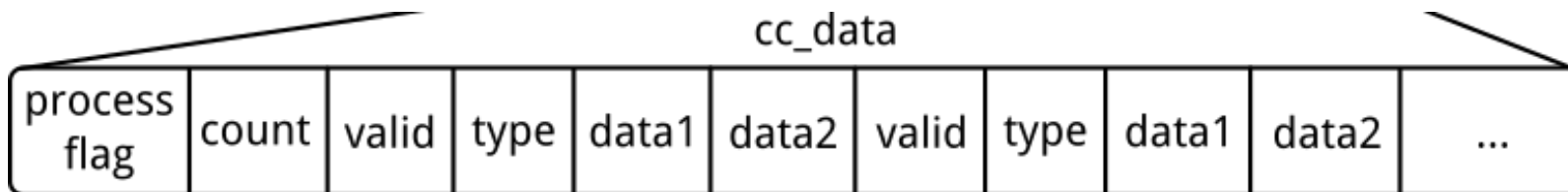
CTA-708 data format (CDP)

- Packet header:
 - magic 0x96 0x69
 - packet size, frame rate
 - data sections present bits (timecode, cc_data, service info)
 - packet counter



CTA-708 data format (cc_data)

- a collection of byte triples
- byte 0 contains cc_valid (1-bit) and cc_type (2-bits), rest marker/padding bits
 - cc_type determines whether this byte pair is CTA-608 data or part of a DTVCC packet
- byte 1 and 2 are the caption data



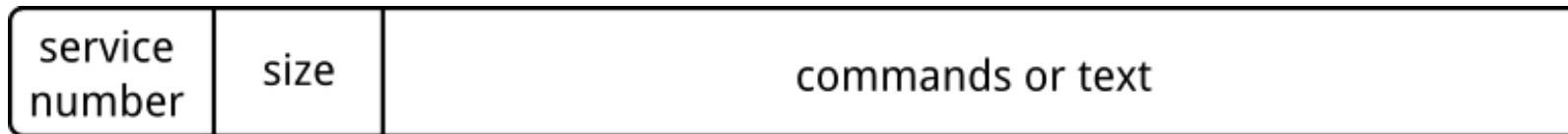
CTA-708 data format (DTVCC Packet (CCP))

- sequence number
- packet size
- packet data (contains one of 63 possible Service Block streams)



CTA-708 data format (Service Block)

- Service number
- Block size
- Caption data
 - Control codes
 - Characters



So, what can GStreamer do with captions?

Warning: a non-exhaustive list

- `GstVideoCaptionMeta` :
 - Parse captions from video streams (H264, H265, MPEG2, etc):
 - Split captions to a separate stream (`ccextractor`)
 - Combine caption stream with video frames (`cccombiner`)
- Convert between caption formats (`ccconverter`)
- Demo

So, what can GStreamer do with captions?

- Upconvert CTA-608 to CTA-708 (`cea608tocea708`)
 - Previously only implemented by expensive hardware
 - Uses <https://github.com/ystreet/cea708-types> for the CTA-708 code/data table
- Demo

So, what can GStreamer do with captions?

- Audio transcription (`transcriberbin`)
 - AWS implementation currently available
 - Other implementations possible
 - See Mathieu's talk for more details

Future possibilities

- `cea708mux` : mux together multiple CTA-708 streams
- Other transcription sources
- Convert from CTA-708 to WebVTT, Timed Text, Teletext, etc
- Other ideas?

Thanks

- ystreet00 on #gstreamer on OFTC
- <https://discourse.gstreamer.org/u/ystreet00>
- <https://gitlab.freedesktop.org/ystreet>
- ystreet00@floss.social on mastodon