

# CLUE Framework Status and Issues

IETF89 - London

March 5, 2014

Mark Duckworth

draft-ietf-clue-framework-14

# Changes in Framework-14

- Add Security section
- Add Participant Information, Participant Type, and Scene Information (was “Role” placeholder)
- Composition within an MCC doesn't use spatial information
- Clean up MCC example
- Remove editor's notes, delete unneeded text

# MCC attributes

- max-captures – add qualifier to say either “=” or “<=”
- add attributes **switched** and **composed**
  - proposed in data model draft
- Seems to be little support for either change
- proposal: do neither, don't make any change

# Global Capture Entry List

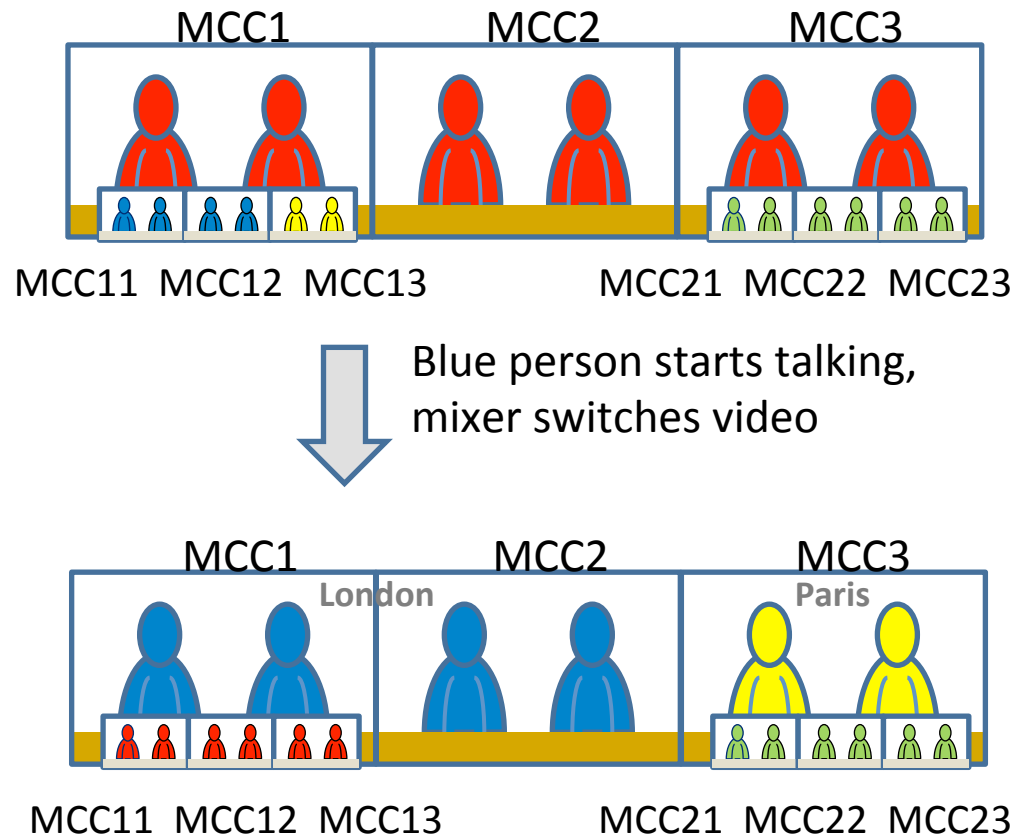
- Proposal – add optional list of “Global Capture Entries”
- A GCE references a set of media captures of a particular media type. The referenced captures can be from any capture scene in the advertisement.
- The set of captures included in a GCE are a suggestion from the provider as to which captures would be a good choice for the consumer to request, to represent a good view of the entire advertisement
- The Provider **MUST** be capable of encoding and sending all Captures in a single Global Capture Entry simultaneously
- The Media Consumer can choose to receive all Media Captures from one Global Capture Entry for each media type (e.g. audio and video), or it can pick and choose Media Captures regardless of how the Provider arranges them in Global Capture Entries.

# Switched Capture Example

- Topo-Mixer – media switching variety
  - see draft-ietf-clue-rtp-mapping  
and draft-ietf-avtcore-rtp-topologies-update
- Mixer (middle box) provides conceptual sources (Media Captures), selecting one source at a time from the original sources
- Mixer is not doing any video composition

# Example Video Layout

- This endpoint is receiving 9 video captures.
- MCC1, MCC2, MCC3 has the current talker
- MCC11 – MCC23 has other recent talkers



# Advertisement Approach

- Advertisement has several scenes, each with spatial information for switched MCC video captures
  - main scene for dominant talker, plus other scenes for next most recent talkers
  - consumer doesn't need to handle changing spatial information on the fly as switches occur, because the MCCs have their own spatial information. This is an advantage over the example in the framework.
- Each scene has multiple CSEs, with different number of MCCs, to accommodate consumers that want to receive a different number of captures
- “Global Alternative Capture List” is provider's suggestion for which captures to choose, across all scenes, again to accommodate consumers that want to receive a different number of captures

# Advertisement

Scene1	Red Room
VC1	Left; view=table
VC2	Center; view=table
VC3	Right; view=table
AC1	view=room
CSE1(VC1,VC2,VC3)	
CSE2(AC1)	

Scene2	Green Room
VC4	Left; view=table
VC5	Center; view=table
VC6	Right; view=table
AC2	view=room
CSE3(VC4,VC5,VC6)	
CSE4(AC2)	

Scene3	Blue Room
VC7	Left; view=table
VC8	Right; view=table
AC3	view=room
CSE5(VC7,VC8)	
CSE6(AC3)	

Scene4	Yellow Room
VC9	view=table
AC4	view=room
CSE7(VC9)	
CSE8(AC4)	

- None of these media captures has an encoding group



# Advertisement

Scene5	Main Scene
MCC1(VC1,VC4,VC7)	Left; maxCaptures=1; SyncID=1; Policy=soundlevel:0; EncodingGroup=1
MCC2(VC2,VC5,VC8)	Center; maxCaptures=1; SyncID=1; Policy=soundlevel:0; EncodingGroup=1
MCC3(VC3,VC6,VC9)	Right; maxCaptures=1; SyncID=1; Policy=soundlevel:0; EncodingGroup=1
MCC4(VC1,VC4,VC7)	Left; maxCaptures=1; SyncID=2; Policy=soundlevel:0; EncodingGroup=1
MCC5(VC2,VC3,VC5,VC6,VC8,VC9)	Right; maxCaptures=1; SyncID=2; Policy=soundlevel:0; EncodingGroup=1
MCC6(<all VCs>)	maxCaptures=1; Policy=soundlevel:0; EncodingGroup=1
MCC7(<all ACs>)	maxCaptures=1; Policy=soundlevel:0; EncodingGroup=2
MCC8(<all ACs>)	maxCaptures=1; Policy=soundlevel:1; EncodingGroup=2
MCC9(<all ACs>)	maxCaptures=1; Policy=soundlevel:2; EncodingGroup=2
CSE101(MCC1, MCC2, MCC3)	
CSE102(MCC4, MCC5)	
CSE103(MCC6)	
CSE104(MCC7,MCC8,MCC9)	
CSE105(MCC7,MCC8)	
CSE106(MCC7)	

# Advertisement

Scene6	PiP Scene 1
MCC11(VC1,VC4,VC7)	Left; maxCaptures=1; SyncID=3; Policy=soundlevel:1; EncodingGroup=2
MCC12(VC2,VC5,VC8)	Center; maxCaptures=1; SyncID=3; Policy=soundlevel:1; EncodingGroup=2
MCC13(VC3,VC6,VC9)	Right; maxCaptures=1; SyncID=3; Policy=soundlevel:1; EncodingGroup=2
MCC14(VC1,VC4,VC7)	Left; maxCaptures=1; SyncID=4; Policy=soundlevel:1; EncodingGroup=2
MCC15(VC2,VC3,VC5,VC6,VC8,VC9)	Right; maxCaptures=1; SyncID=4; Policy=soundlevel:1; EncodingGroup=2
MCC16(<all VCs>)	maxCaptures=1; Policy=soundlevel:1; EncodingGroup=2
CSE111(MCC11, MCC12, MCC13)	
CSE112(MCC14, MCC15)	
CSE113(MCC16)	

- Scene7, PiP Scene 2, is similar.
- It uses MCC21 to MCC26, with Policy=soundlevel:2; EncodingGroup=3.

# Advertisement

## Global Capture Entry List

(MCC1, MCC2, MCC3, MCC11, MCC12, MCC13, MCC21, MCC22, MCC23)

(MCC1, MCC2, MCC3, MCC11, MCC12, MCC13, MCC24, MCC25)

(MCC1, MCC2, MCC3, MCC11, MCC12, MCC13, MCC26)

(MCC1, MCC2, MCC3, MCC11, MCC12, MCC13)

(MCC1, MCC2, MCC3, MCC14, MCC15)

(MCC1, MCC2, MCC3, MCC16)

(MCC1, MCC2, MCC3)

(MCC4, MCC5)

(MCC6)

- The provider is suggesting all these to accommodate consumers that want to receive different number of captures.
- Each entry in the list provides a good view of the entire advertisement, for the given number of captures.

# Need “active capture” info

- Signaling the “active video capture”, not using just audio energy in the audio stream. Discussed at October 2011 interim meeting, where the group supported this need, but we never followed up on it.
- Example: endpoint provider sends 3 video captures, and 1 mono audio
  - Consumer (MCU topo-mixer) wants to know which video has the loudest talker, so it can send that one to a simple endpoint that just receives one video stream
  - How does this MCU consumer know which one it is?
- Provider can signal this somehow, if it knows locally which video is associated with the talker
  - in RTP?
  - in CLUE channel?
  - with separate switched MCC capture for the “active capture”?
    - but don’t want to send duplicate encodings wasting bandwidth

# “active capture” Proposal

- Use separate switched MCC to represent the active capture
- Example, endpoint provider advertisement:
  - VC1 left; VC2 center; VC3 right
  - MCC1(VC1,VC2,VC3) maxCaptures=1; policy=soundlevel:0
  - CSE1(VC1,VC2,VC3)
  - CSE2(MCC1)
- MCU receives all four captures, so it has enough information to send the right one(s) to different consumers
- Desire optimization from RTP Mapping draft, Media-9 requirement:
  - If a given source is being sent on the same transport flow for more than one reason (e.g. if it corresponds to more than one switched capture at once, or to a static capture), it should be possible for a sender to send only one copy of the source.

# Open Tickets

- #10 sufficient info for receiver – ready to close this?
- #19, #26 site switching – we have MCC syncID, can we close this?
- #32 Roles – Added to framework, can close
- #33 Security – Added to framework, can close
- #35 consistency with data model - ongoing