draft-ramalho-g7110-segments-00

27 March 2012

Michael A. Ramalho, Ed.
Dan Wing
Noboru Harada
Hadriel Kaplan

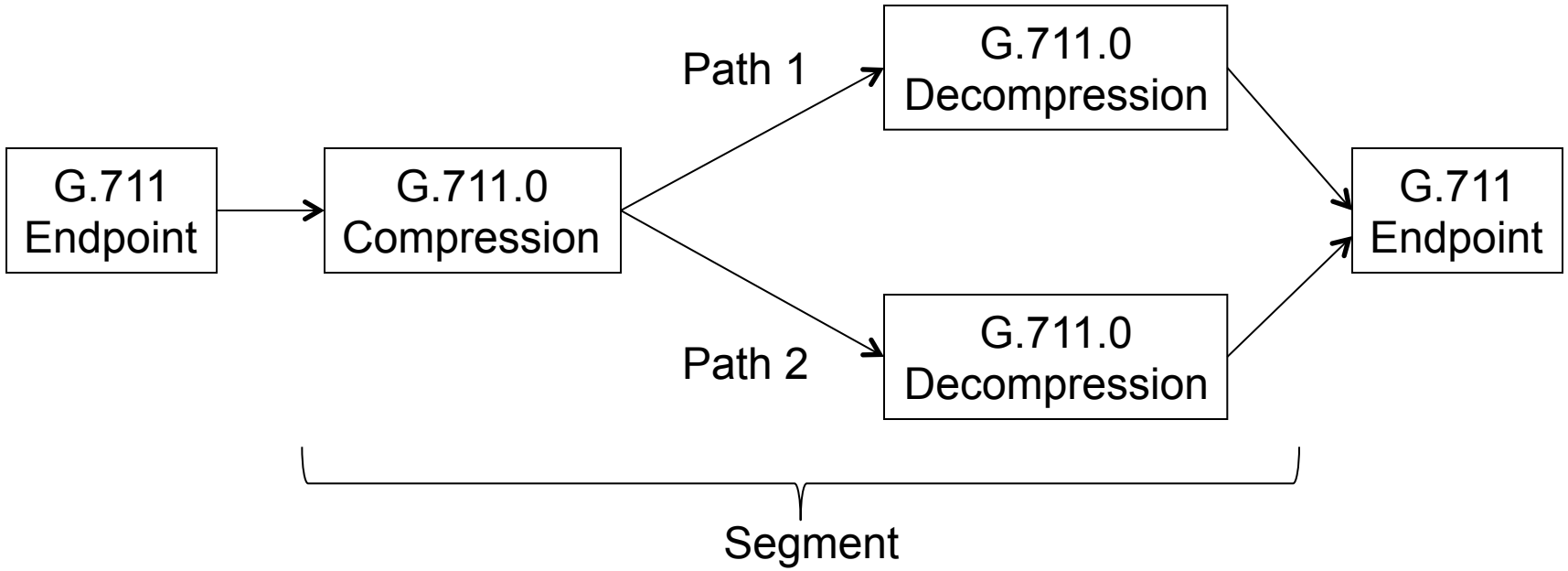
What this Draft is About

- Best Practices for when G.711.0 when used as *lossless* compression mechanism on *one or more segments* of an *end-to-end G.711* session.
- Guidelines for the following cases (for media):
  - Only G.711 RTP payload compression
  - Both RTP header and G.711 payload compression
  - Multiple “G.711 channels” in one RTP session
- Signaling considerations for when *G.711.0 is not negotiated end-to-end:*
  - G.711.0 re-negotiation on “media segments”

Essential G.711.0

- G.711.0 is a data compression algorithm especially designed for A-law or Mu-law G.711 VoIP payloads (i.e., not a generic compression).
- Lossless => Lossless for ANY payload (including random data in DS0s).
- Stateless => Compression not dependent on previous frames.
  - No error-propagation at decoder possible due to lost prior packets.
- "Self-describing" => G.711 regenerated WITHOUT access to signaling.
- Two Dominant Use Cases:
  - End-to-End: G.711.0 Payload Format (draft-ramalho-payload-g711) where G.711.0 is negotiated as "if it were a codec".
    - Nearly identical to G.711 RTP spec (exception is Payload Type)
  - In-The-Middle: This draft (draft-ramalho-g7110-segments) Can be employed multiple times within an end-to-end G.711 session.
    - Without endpoint or call agent knowledge.
    - With endpoint or call agent knowledge.
    - With no degradation of voice quality relative to G.711 (always lossless).

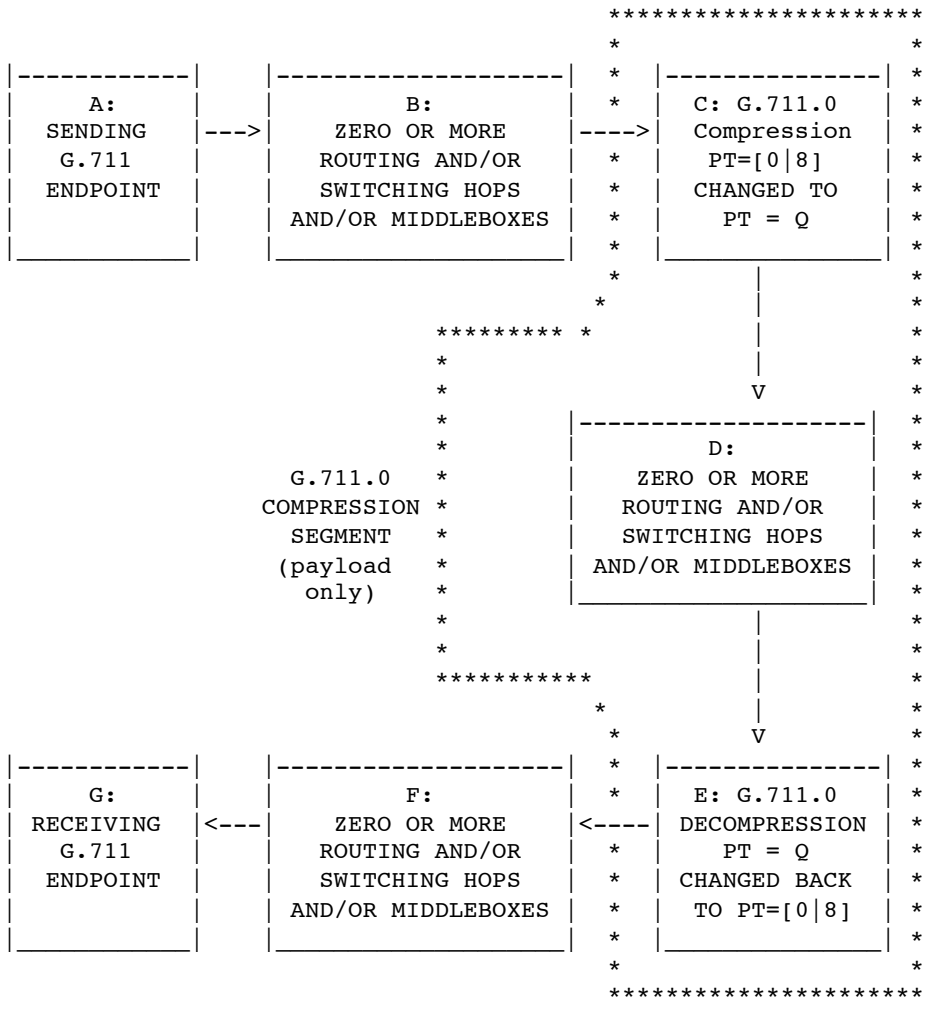
G.711.0 Statelessness



- (Compression and) Decompression is stateless.
- Packets need not traverse same IP routed path as long as decompression functionality can be assured prior to terminating G.711 endpoint.
- There are several topology scenarios where this can be assured (i.e., within a given Administrative Domain).
- Ditto for multiple compressors on possible egress paths (not shown).

G.711.0 Compression Segment

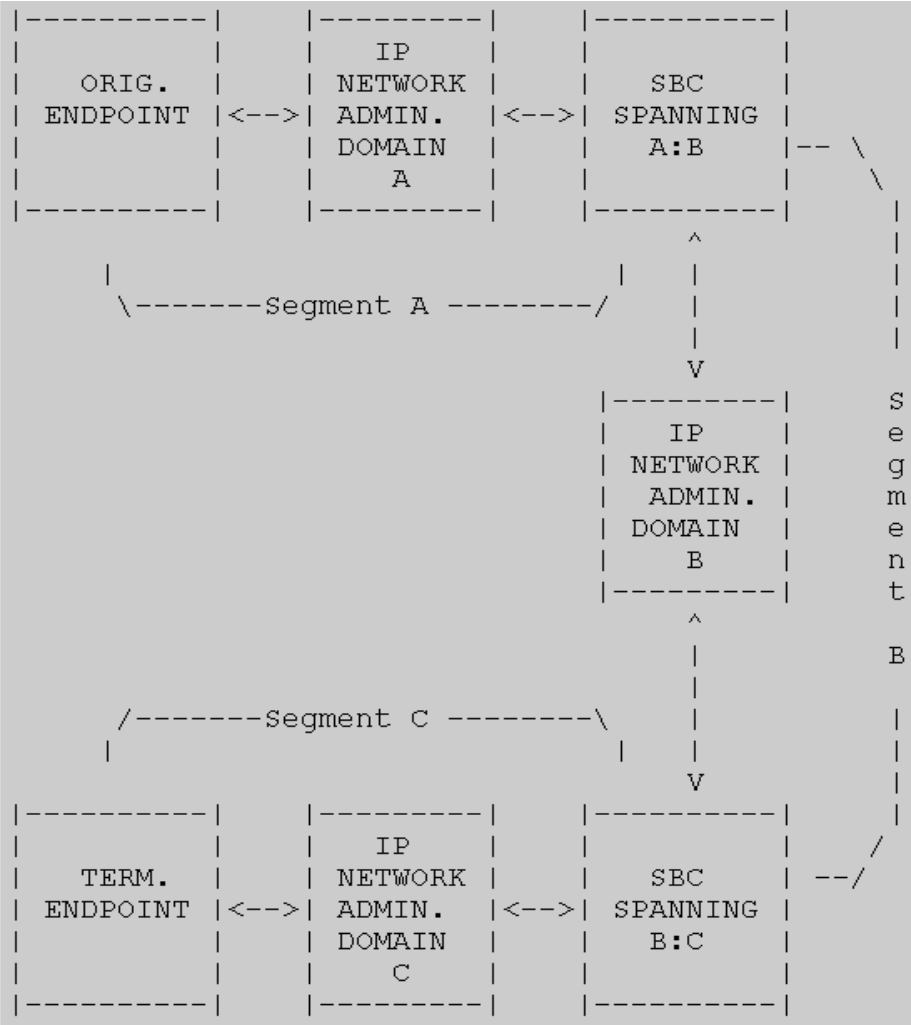
Media Only: No RTP Header Compression Case



- PT = 0; G.711 Mu-law (PCMU)
- PT = 8; G.711 A-law (PCMA)
- May have multiple compression segments on end-to-end connection.
- No Box D => Compression over single link.
- “Unaware Middleboxes” must be configured to let traffic other than what is expected to pass (in this case to let PT = Q pass).
- SBCs can re-negotiate G.711.0 on any UA-to-UA segment such that endpoints or other middleboxes outside of that segment do not know G.711.0 compression was performed (next slide).

G.711.0 Compression Segment

Signalling Issues



- Codec negotiated end-to-end is G.711 because one or more segments do not have G.711.0 as an option.

- At least one segment re-negotiates G.711.0.
  - Translation to/from G.711 on segments not supporting G.711.0.
  - Negotiation can use G.711.0 SDP parameters specified in G.711.0 RTP Payload Specification.

- No voice quality degradation (rel. to G.711).
  - Thus G.711 segments need not be informed of G.711.0 use on other segments.
  - Non-G.711.0 segments do NOT participate in any end-to-end SDP negotiation regarding codec change from G.711 to G.711.0.

draft-ramalho-g7110-segments-00

Other Issues Addressed in Draft

- RTP Payload and RTP header compression case
  - Header compression SHOULD be stateless.
- Implications for voice quality & added delay
  - No voice quality degradation relative to G.711.
  - Statelessness guarantees no media error propagation (burst loss).
  - Low complexity implies trivial added delay (< 1 ms).
- Multiplexing Multiple G.711 flows in G.711.0 payload
  - Not specified, except possibility to use channel parameter specified in existing G.711.0 RTP payload format draft.
- Translating G.711.0 RTP headers to/from G.711 RTP headers for "in-the-middle" G.711.0 segments.
  - G.711 RTP packets "appear" to have been transparently transported.

Known Open Issues

- How to choose PT = Q during re-negotiation?
  - MUST be outside of range of dynamic PT assignment
  - RECOMMENDED that Q be chosen from a static PT that is known never to be assigned within the scope of the G. 711.0 compression segment or from the range of unassigned PT?
  - Comments?
- Other Comments/Suggestions?

Thank You

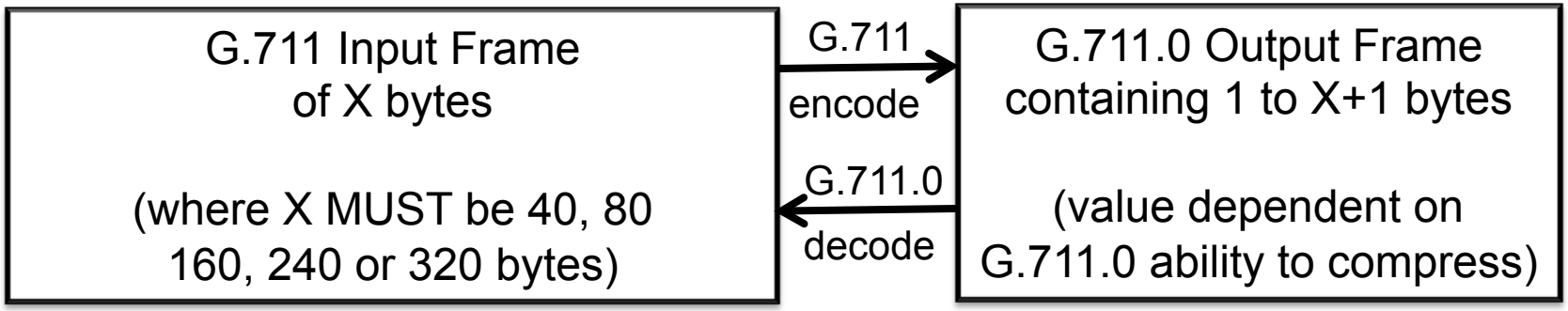
(Backup slides follow)

Design Requirements in ITU-T G.711.0 “Terms of Reference”

- Support both G.711 A-law and Mu-law.
- Lossless for ANY payload (including random data in DS0s).
- Accommodates G.711 payload sizes typically used in VoIP.
- Stateless: Compression not dependent on previous frames.
  - No error-propagation at decoder possible due to lost prior packets.
- Algorithmic delay equal to the time represented by G.711 input.
  - No “look-ahead” or per-channel state.
- Self describing G.711.0 output frame.
  - Decoder is NOT dependent on access to signaling.*
  - Encoder is NOT dependent on access to signaling.*
- Bounded expansion for “uncompressible G.711 input frames”.
- Low complexity (<1 WMOPS, 10k memory, 3.6k basic operations).

* Only needs to know A-law or mu-law. Encoder only needs the G.711 input frame.

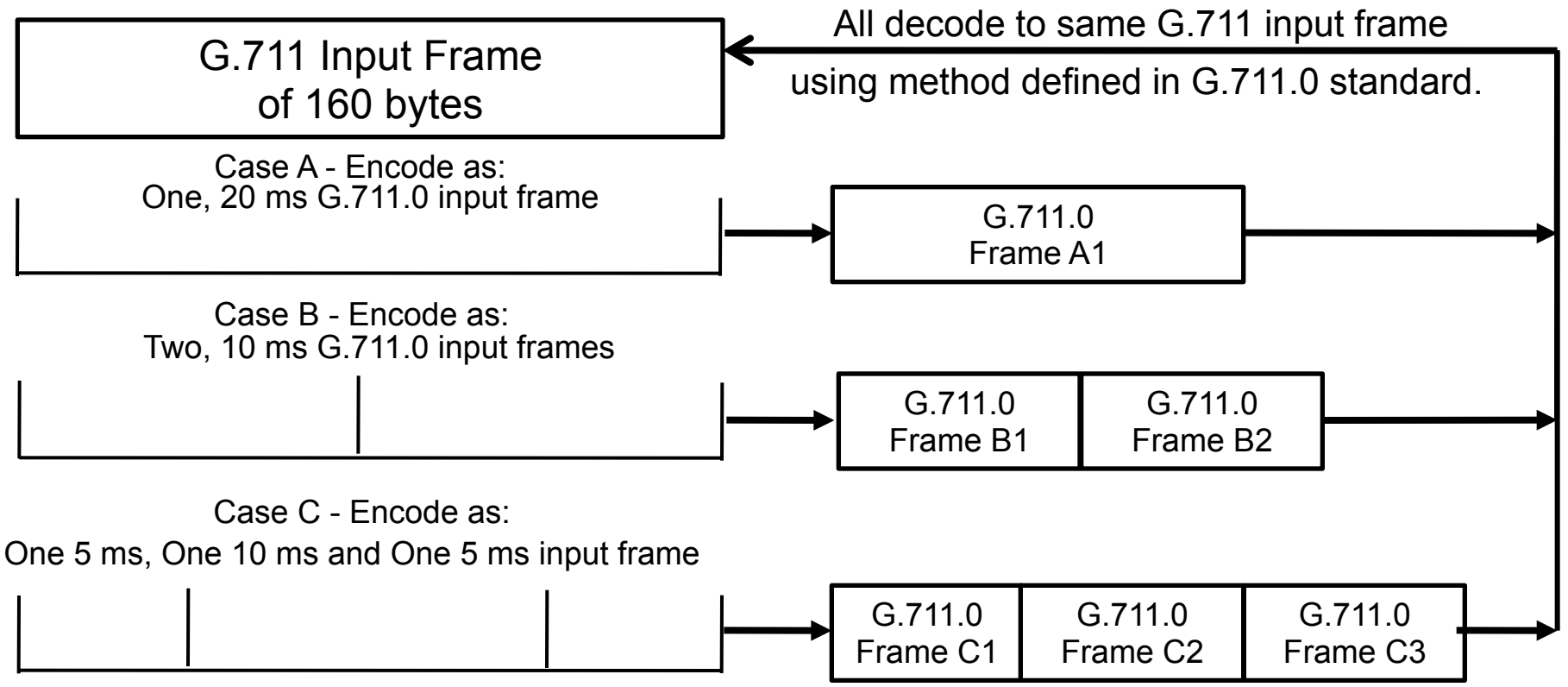
G.711.0 Basic Operation



At 8k sampling:
40 samples = 5 ms
80 samples = 10 ms
160 samples = 20 ms
240 samples = 30 ms
320 samples = 40 ms

- Mapping is 1:1 in both directions
- G.711.0 is a "Self Describing" encoding:
  - Decoder – without any signaling information - knows how many G.711 source samples to produce
- Optimized for zero-mean acoustic signals, however …
- Lossless for any G.711 input frame (including random data)

Complex G.711.0 Encoding Example:
20ms/160 bytes of G.711



- A smart encoder may choose ANY combination of sub input frame sizes to determine which compresses best (usually the largest does)
- As a result, ANY integer number of 5 ms of G.711 can be encoded and placed in a RTP payload

G711.0 Internal Design & Compression Results

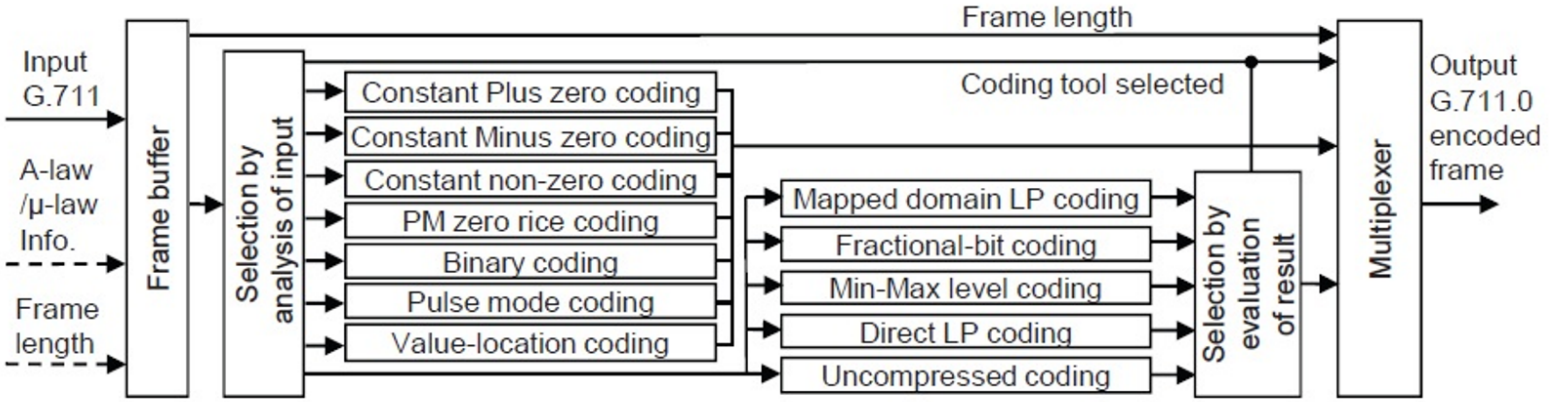


Fig. 1. High-level block diagram of the G.711.0 encoder.

Test category		Compression ratio [%]	
		A-law	μ-law
(a1): Clean speech	-16 dBoV	59.56 %	50.67 %
	-26 dBoV	69.39 %	60.62 %
	-36 dBoV	77.01 %	72.55 %
(a2): Noisy speech	SNR 15 dB	50.90 %	44.52 %
	SNR 20 dB	54.43 %	47.15 %
	SNR 25 dB	60.64 %	52.43 %
(a1) and (a2) conditions in total		57.55 %	50.24 %
(a3): Tandem conditions in total		60.08 %	54.52 %
(b): Recorded (NTT) μ-law corpus		-	50.83 %

Note: Conservative because averaged over all G.711.0 frame lengths (of 5ms, 10ms, 20ms and 30ms). Results for 20ms are better by about 2%. A-law compresses better due to coarser quantization at low levels.

N. Harada, Y. Yamamoto, T. Moriya, Y. Hiwasaki, M. A. Ramalho, L. Netsch, Y. Stachurski, Miao Lei, H. Taddei, and Q. Fengyan, "Emerging ITU-T Standard G.711.0 - Lossless Compression of G.711 Pulse Code Modulation" International Conference on Acoustics Speech and Signal Processing (ICASSP), March 2010, ISBN 978-1-4244-4244-4295-9