VoIP Shim for RTP Payload Formats

draft-johansson-avt-rtp-shim

Ingemar Johansson, Ericsson AB

ERICSSON

TAKING YOU FORWARD

Outline

- MTSI in 3GPP Voice service requirements
- Problems with RTCP
- Why is inband signaling better in this application ?

Some buzzwords

- 3GPP (3rd Generation Partneship Project)
  - Standardization body for WCDMA and GSM
  - http://www.3gpp.org
- HSPA (High Speed Packet Access)
  - Aka ”Turbo 3G”
  - Constantly evolving
  - HSDPA (High Speed Downlink Packet Access)
  - EUL (Enhanced UpLink)
  - ~10Mbps (various figures)
- GPRS (General Packet Radio Service)
  - ”Internet for GSM”
  - Bitrates up to ~40kbps
- EDGE
  - Improvement to GSM-GPRS
  - More complex modulation techniques
  - Bitrates up to ~480kbps
- MTSI (Multimedia Telphony Service for IMS)
- IMS (IP multimedia System)
- AMR (Adaptive Multi-Rate codec)
  - Both narrow band (300-3500Hz) and wideband (50-7000Hz)
  - Payload format standardized in RFC3267

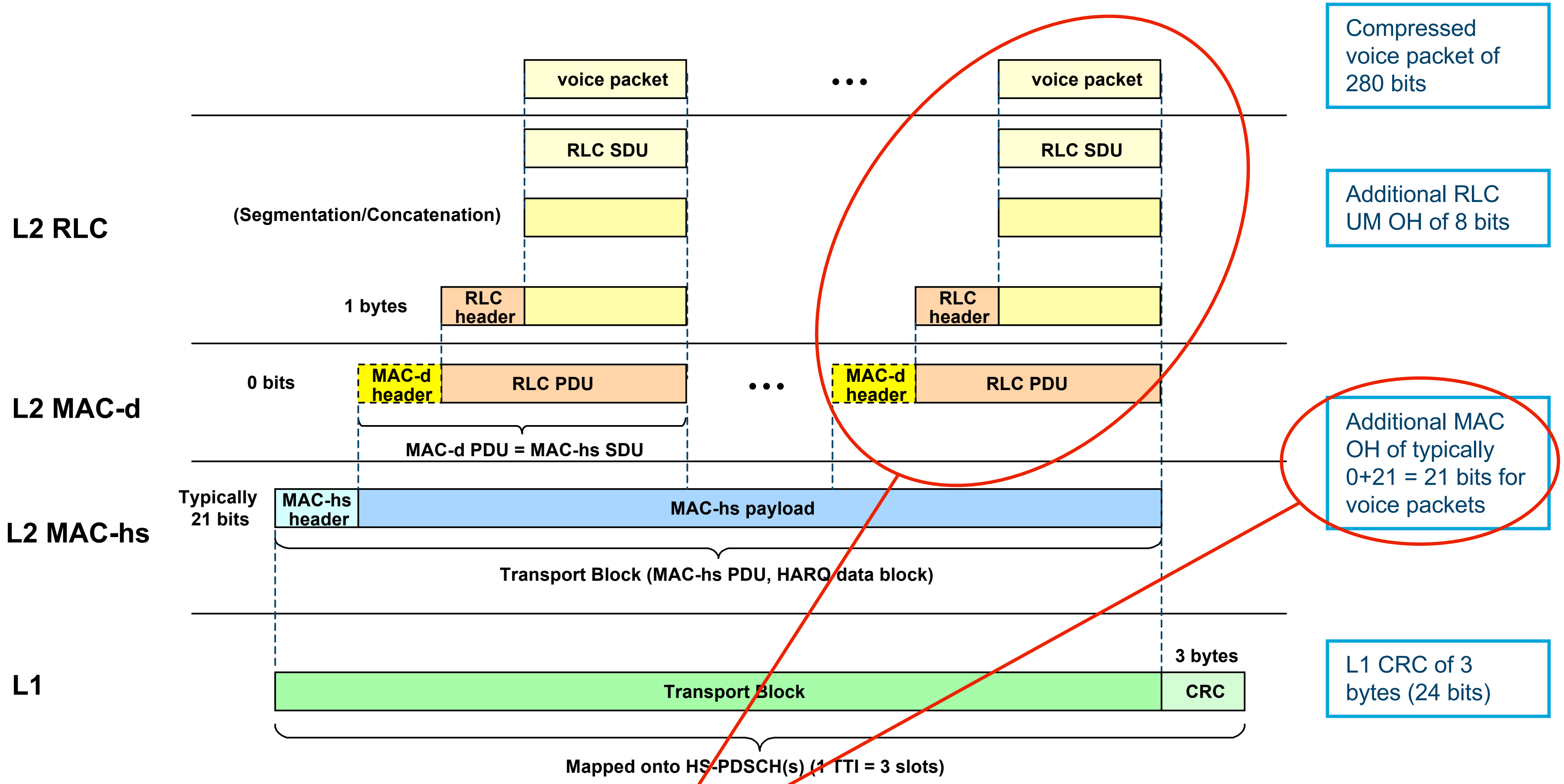
MTSI in 3GPP

- Multimedia telephony service specified in 3GPP
- Involves Video and Voice + other components such as Text.
- IP protocol based.
- Highly optimized radio bearers.
- Header compression for capacity boost.

3GPP VoIP bearers

- Highly optimized radio bearers tailored to fit with packet sizes for AMR.
  - Target is as many (satisfied) users as possible.
  - Minimized lower layer overhead.
  - IP/UDP/RTP header compressed to ~3byte RoHC overhead by means of header compression.
- Use of larger packet sizes comes with a cost :
  - In HSPA systems many retransmission at lower layers possible→ more jitter, greater risk of packet loss when user close to cell border (coverage issue).
  - EDGE non persistent mode limits to maximum one retransmission (delay req.) → Very high risk that large packets are lost in bad radio conditions.
  - Preferrable that packet sizes becomes smaller when users reach cell border.

HS-DSCH Data transmission



MAC-d PDU should be optimized for codecs used for MM Telephony to increase capacity. The number of PDU sizes is limited to 8 (by a 3 bit field in the MAC-hs header)

PDU optimization

- The speech codec determines PDU sizes

MAC-d PDU size	Comments
96, 112	To be used for SID frames
144	Optimized for AMR-NB 4.75+3 bytes ROHC
160	Optimized for AMR-NB 5.90+3 bytes ROHC
176	Optimized for AMR-WB 6.60+3 bytes ROHC
192	Optimized for AMR-NB 7.40+3 bytes ROHC
208	Optimized for AMR-NB 7.95+3 bytes ROHC
224	Optimized for AMR-WB 8.85+3 bytes ROHC
288	Optimized for AMR-NB 12.2+3 bytes ROHC
296	Optimized for AMR-WB 12.65+3 bytes ROHC
312	AMR-NB 12.2 + 6 bytes ROHC
344 (alt. 336)	To be used to convey larger packets (video, ROHC IR-DYN packets etc)

- Video or other real-time media creating large packets will be segmented and use the largest possible PDU size
- The use of large PDU sizes leads to less coverage than when using small PDU sizes

The need for adaptation

- Radio network performance is defined on a large scale basis.
  - Admission control determines if new users (VoIP calls) should be admitted.
  - Metric : percentage of satisfied users.
    - Even though retransmission scheduling algorithms do their best some unlucky users might experience poor quality.
- Fast application adaptation is needed
  - Fast reduction of codec rate enables better coverage for users at cell border.
  - Handover to other networks with different properties demands different application layer behavior.
    - EDGE or 802.11a gives better performance if the packet rate is reduced → frame aggregation preferrable
    - HSPA applies dynamic frame aggregation on lower layers by means of retransmission → frame aggregation gives no/litttle improvement.

Adaptation requests vs metrics

- Metrics is the established feedback entity esp. in IETF
  - Exceptions exists (eg. Full Intra Request in draft-ietf-avt-avpf-ccm)
- Requests
  - Simple entities such as
    - Reduce codec rate.
    - Employ redundancy.
    - Enable frame aggregation
    - Combinations of the above
  - Possible to send requests to other endpoint based of features that are only known in terminal.
    - Handover to different network that requires eg lower packet rate for optimum performance
    - High load
    - Close to cell border.
    - Many features are access specific → difficult to standardize as metrics to transmitted
  - Request does not mean that it is mandaded to follow.
    - A request for redundancy may very well be rejected in case the receiver of the request finds it inappropriate.
  - Can be transmitted inband (in the RTP flow) or out of band (RTCP)

Inband vs. Out of band

- Inband :
  - Signaling is transmitted in the RTP flow.
  - Example CMR bits in RFC3267 (AMR payload).
  - Inband signaling proposed in draft-johansson-avt-rtp-shim.
- Out of band :
  - Signaling is transmitted by means of a protocol (e.g RTCP), separated from the RTP flow.

Why inband ?

- Why not stick to RTCP ?
  - Inband signaling messes with the RTP architecture
  - RTCP has many benefits, solves many issues quite nicely
  - RTCP useful for performace monitoring

Problems with RTCP
Periodicity

- RTCP must be transmitted on a periodic basis even though the application does not need it.
- AVPF relaxes things a bit (immediate/early mode)
  – Still not possible to send ”only when necessary”
- Principle can be questioned
  – No large impact on capacity as RTCP bandwith is constrained

Problems with RTCP
Packet size

- A minimum required RTCP entities are mandated even though they are not (always) needed.
  - SR or RR with report block(s) + SDES(CNAME) mandated.
  - Minimum size is ~100 bytes (+IP/UDP), can be much larger however depending on CNAME string
  - IP/UDP can be compressed with RoHC but thats all.
- RTCP packet size typically ~3-4 times larger than VoIP RTP packets.
  - Will become segmented in optimized VoIP bearers → increased risk that RTCP packets are lost.
- ”Schrödinger’s cat” problem
  - Performance monitoring by means of RTCP in highly VoIP optimized transmission channels can affect the RTP flow considerably making performance monitoring this way questionable.

What are the options

- Inband signaling (as proposed in draft-johansson-avt-rtp-shim)
  - Useful for point to point communication
  - Small overhead
  - Controversial
- RTCP cheats (abuse)
  - Use RR instead of SR
    - NTP only needed for synch eg. between Voice and Video
  - Send RTCP with zero report blocks
    - Report blocks can be sent only once in awhile
  - Skip SDES-CNAME
    - ...or don't send it in every RTCP
    - Is there any need for this in a point to point application ?
  - Must be verified that middle boxes accept this..
- Additional RTCP compression ?
  - Unknown if it is possible to achive good compression

ERICSSON

TAKING YOU FORWARD