ETSI TR 126 916 V14.2.0 (2017-05)

TECHNICAL REPORT

Universal Mobile Telecommunications System (UMTS);
LTE;
Media handling and Quality aspects of SRVCC
(3GPP TR 26.916 version 14.2.0 Release 14)

Reference

DTR/TSGS-0426916ve20

Keywords

LTE,UMTS

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

The present document can be downloaded from:
http://www.etsi.org/standards-search

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the only prevailing document is the print of the Portable Document Format (PDF) version kept on a specific network drive within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at
https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

If you find errors in the present document, please send your comment to one of the following services:
https://portal.etsi.org/People/CommiteeSupportStaff.aspx

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Technical Report (TR) has been produced by ETSI 3rd Generation Partnership Project (3GPP).

The present document may refer to technical specifications or reports using their 3GPP identities, UMTS identities or GSM identities. These should be interpreted as being references to the corresponding ETSI deliverables.

The cross reference between GSM, UMTS, 3GPP and ETSI identities can be found under http://webapp.etsi.org/key/queryform.asp.

Modal verbs terminology

In the present document "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Contents

Foreword

This Technical Report has been produced by the 3rd Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

- x the first digit:
  - 1 presented to TSG for information;
  - 2 presented to TSG for approval;
  - 3 or greater indicates TSG approved document under change control.
- y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.
- z the third digit is incremented when editorial only changes have been incorporated in the document.

Introduction

Single Radio - Voice Call Continuity (SRVCC) is an existing standard ([3], [4]), specifying the handover of a Voice or Video call from LTE access to CS radio access, either to GERAN (2G) or to UTRAN (3G) or other CS networks. The present document considers only eSRVCC for voice calls between 3GPP accesses.

In the IMS Core Network, the voice call is typically anchored in the ATCF/ATGW (Access Transfer Control Function / Access Transfer Gate Way). The eSRVCC procedure, as specified, may cause additional transcoding between the target radio leg and the ATGW, even though in theory it would be possible to avoid it. As a result, the eSRVCC procedures may add one or more unnecessary transcoding point(s) for the call and thereby degrade the quality of the ongoing call unnecessarily.

Transcoding-Less Codec Interworking (TLCI) is always desirable to achieve good voice quality. Furthermore TLCI preserves network resources, i.e. by avoiding transcoding. TLCI is especially important for HD Voice.

The Mobility Management Entity (MME) of the LTE-RAN, which sends the PS-to-CS Handover Request to the Target Network, does not know the **IMS Selected Codec**, which is in use before the eSRVCC in the ongoing call towards the remote end. Thus the MME cannot support the Target Network for selecting the optimal **Target RAN Codec**. The Target Network thus selects this Target RAN Codec on own criteria; often the Target RAN Codec is then not compatible to the IMS Selected Codec. Transcoding is then the immediate reaction.

While it is possible for the ATCF, based on the current procedure, to renegotiate the IMS Selected Codec with the remote end to fit any selected Target RAN Codec at call transfer, this may extend the perceived time it will take to conclude the call transfer and this might extend the speech interruption time that might result due to the time the additional negotiation with the remote end will take. The ATCF was introduced for exactly that reason: avoid renegotiation with the remote end - accelerate eSRVCC.

But even worse: in a substantial number of call scenarios the remote end may not be able to support the arbitrarily chosen Target RAN Codec and the transcoding cannot even be avoided by that renegotiation.

The first attempt will optimize the Target RAN Codec to fit the IMS Selected Codec. If that is impossible or not optimal, then the renegotiation with the remote end might be attempted. The last resort has to be transcoding; sometimes it is unavoidable.

1 Scope

Enhanced Single Radio - Voice Call Continuity (eSRVCC) is an existing standard ([3], [4]) specifying the handover of a Voice or Video call from LTE access to CS-radio access, either to GERAN (2G) or to UTRAN (3G) or other CS networks. The present document considers only enhanced SRVCC for voice calls between 3GPP accesses.

This study assumes that the Codecs defined in TS 26.114 are used on the LTE access and the Codecs defined in 3GPP TS 26.103 [7] on the CS accesses. Since Rel-13, the specifications for CS networks include the Codec Type UMTS_EVS with several Configurations, called UMTS_EVS (Set 0) to UMTS_EVS (Set 3).

In the IMS Core Network, the voice call is typically anchored in the ATCF/ATGW (Access Transfer Control Function/ Access Transfer Gate Way).

The eSRVCC procedure, as specified, may cause additional transcoding between the target radio leg and the ATGW, even though in theory it would be possible to avoid it. As a result, the eSRVCC procedures may add one or more unnecessary transcoding point(s) for the call and thereby degrade the quality of the ongoing call unnecessarily.

The main objectives of this study are to analyse example call scenarios and find potential solutions to minimize the number of transcoding cases. Another objective is to optimize the interworking and the transition between EVS and AMR-WB during eSRVCC. The study should also show the reasons and potential solutions for too long speech path interruptions during eSRVCC.

The present Technical Report has the following detailed objectives:

- Identify relevant eSRVCC scenarios, especially with Codec Mode Control;
  from AMR-WB and/or EVS in VoLTE to AMR-WB and/or EVS in CS;
  but include also other important Codecs, such as AMR and G.722.

- Analyse Speech Quality Aspects and Media Handling Aspects, based on these scenarios.

- Analyse Codec Mode Control before, during and after eSRVCC;
  recently SA4 has clarified some essential details on Rate Control for AMR and AMR-WB;
  Rate Control and Audio Bandwidth Control for EVS are still under discussion to some extent.

- Analyse the existing SDP Offer - Answer protocol between Target MSC and Anchor-ATCF during eSRVCC, as specified in 3GPP TS 23.216 [3], Stage 2;
  This analysis will include the whole eSRVCC procedure for at least one essential scenario (e.g. eSRVCC to GERAN) and will identify the potential reasons for transcoding and too long speech path interruptions.

- Clarify the existing Codec Compatibility aspects for eSRVCC;
  especially the interworking between CS and IMS for AMR, AMR-WB and EVS needs to be documented.

- Propose enhancements for media and quality aspects of eSRVCC with the aims:
  a) to avoid transcoding cases as much as possible;
  b) to minimize the speech path interruption time during eSRVCC;

- Support the SA2 SETA work by SA4 expertise in speech quality and media handling.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[2] 3GPP TR 41.001: "GSM Specification set".

[3] 3GPP TS 23.216 (V12.1.0): "Single Radio Voice Call Continuity (SRVCC); Stage 2 (Release 12)".

[4] 3GPP TS 24.237 (V13.0.0): "IP Multimedia (IM) Core Network (CN) subsystem IP Multimedia Subsystem (IMS) service continuity; Stage 3 (Release 13)".

[5] 3GPP TS 26.114 (V12.10.0): "IP Multimedia Subsystem (IMS); Multimedia telephony; Media handling and interaction (Release 12)".

[6] 3GPP TR 23.717 (V1.0.0): "Enabling Transcoder Free Operation During SRVCC (PS to CS)".

[7] 3GPP TS 26.103: "Speech codec list for GSM and UMTS".

[8] 3GPP TS 26.445: "Codec for Enhanced Voice Services (EVS); Detailed algorithmic description".

[9] IETF RFC 4867: "RTP Payload Format and File Storage Format for the Adaptive Multi-Rate (AMR) and Adaptive Multi-Rate Wideband (AMR-WB) Audio Codecs".

[10] 3GPP TS 26.454: " Codec for Enhanced Voice Services (EVS); Interface to Iu, Uu, Nb and Mb".

3 Definitions, symbols and abbreviations

3.1 Definitions

For the purposes of the present document, the terms and definitions given in 3GPP TR 21.905 [1] and the following apply. A term defined in the present document takes precedence over the definition of the same term, if any, in 3GPP TR 21.905 [1].

Active EVS Configuration: is always as big as or smaller than the EVS Framework Configuration, never bigger. It may be modified by CMR and - maybe - other influences, like RNC Max-Rate-Control. A voice call has two Active EVS Configurations, one in each direction.

Assumption: client hosting the EVS Encoder/Decoder may send CMR anytime to influence the media-stream it receives. The sent CMR value is always within the limits of the local EVS Configuration; it may be outside the perceived Active EVS Configuration in receiving direction. It may happen that a node (e.g. MGW) receives CMR-values outside the local EVS Configuration of next following link. The node then limits the received CMR values to the next local EVS Configuration. This guarantees that the CMR-receiving media-sender gets in error free cases only CMR values within its own local EVS Configuration.

Codec: used for the combination of Codec Type plus Codec Configuration, as used in Codec Negotiation, like in the SIP/SDP Offer - Answer procedure or in BICC IAM - APM signalling

Codec Configuration: defines the full set of attributes to a certain Codec Type, e.g. the set of Codec Modes

Codec Mode: defines a specific mode of a Codec Type, e.g. the 12,2 kbps mode of the AMR

Codec Type: defines a specific type of a speech coding algorithm, applied on a specific radio or other transport technology, e.g. GSM FR, FR_AMR, AMR, AMR-WB, EVS, G.722, G.711, see also 3GPP TS 26.103 [7]

CS PS Codec: Codec for the Interface between CS- and IMS-network

EXAMPLE: G.711, AMR(0,2,4,7), AMR-WB(0,1,2), UMTS_EVS(...).

EVS Framework Configuration: is selected by the Offer-Answer Codec Negotiation at call setup or in mid-call modifications. It is the intersection of all Local EVS Configurations along the speech path. It is not explicitly known to every node in the path.

IMS Selected Codec: Codec selected for the call before SRVCC from the ATGW towards the remote end

EXAMPLE: AMR(0,2,4,7), AMR-WB(), EVS(), G.722, G.711.

Local EVS Configuration: is sent to the EVS client by SIP/SDP or CS Signalling after Codec (re-) Negotiation, or to a MGW within the path. There might be different local EVS Configurations along the speech path for different sub-links.

LTE Used Codec: Codec used on the LTE RAN leg before SRVCC between local UE and ATGW

EXAMPLE: AMR(0,2,4,7), AMR-WB(), EVS().

Naming convention: EVS Codec is named by its main SDP parameters in the SDP Answer, put in brackets ()

EXAMPLE: "**EVS (br=5.9-128; bw=nb-fb)**". This example means: The EVS primary mode of operation is selected with all audio bandwidths allowed. The "mode-set" parameter for EVS-IO need not (always) to be present (Open Offer, Open Answer).

Target RAN Codec: Codec chosen by the Target Network for the Target RAN leg after SRVCC

EXAMPLE FR_AMR(0,2,4,7), HR_AMR(0,2,4), UMTS_AMR2(0,2,4,7).

Transcoding-Less Codec Interworking: Interworking between Codecs in a gateway without decoding and re-encoding the Speech and SID contents, but with potentially modifying rate control commands

3.2 Symbols

For the purposes of the present document, the following symbols apply:

()	without a mode-set, e.g. in the Open Offer
(0,2,4,7)	with mode-set=0,2,4,7
(...)	with or without a mode-set
AMR()	AMR Codec without a mode-set
AMR (0,2,4,7)	AMR Codec with mode-set=0,2,4,7
FR_AMR(...)	AMR Codec on the Full Rate GERAN traffic channel
HR_AMR(...)	AMR Codec on the Half Rate GERAN traffic channel
UMTS_AMR2(...)	AMR Codec on the UTRAN traffic channel
AMR-WB()	AMR-WB Codec without a mode-set
AMR-WB (0,1,2)	AMR-WB Codec with mode-set=0,1,2
FR_AMR-WB(...)	AMR-WB Codec on the Full Rate GERAN traffic channel
UMTS_AMR-WB(...)	AMR-WB Codec on the UTRAN traffic channel
EVS()	EVS Codec with all its operational modes, i.e. in the Open Offer
EVS (br=5.9-128; bw=nb-fb)	EVS Codec with all its operational modes, i.e. in the Open Offer
EVS-NB (...)	EVS Codec in Narrow-Band operation
EVS-WB (...)	EVS Codec in Wide-Band operation
EVS-SWB(...)	EVS Codec in Super-Wide-Band operation
EVS-FB(...)	EVS Codec in Full-Band operation
EVS-IO (...)	EVS in AMR-WB Inter-Operable operation
UMTS_EVS (Set x)	EVS over CS (UMTS) with Configuration Set x; x=0,1,2,3
UMTS_EVS (…)	EVS over CS (UMTS) with any Configuration Set
<=>	is used when two Codecs are TLCI-compatible, i.e. no transcoding is needed
EXAMPLE 1:	AMR(0,2,4,7) <=> HR_AMR(0,2,4)

EXAMPLE 2:	EVS-IO(0,1,2) <=> AMR-WB(0,1,2)

<=/=>	is used when transcoding is needed
EXAMPLE 1:	AMR(0,2,4,7) <=/=> UMTS_AMR2(0,2,5,7)
EXAMPLE 2:	EVS-NB() <=/=> FR_AMR(0,2,4,7)

3.3 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905 [1] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in 3GPP TR 21.905 [1].

AMR	Adaptive Multi-Rate (Codec)
AMR-WB	Adaptive Multi-Rate WideBand (Codec)
AMR-WB-CMR	AMR-WB Codec Mode Request (if needed to differentiate from EVS-CMR)
APM	Application Transport Mechanism (functionality-wise like SIP Response)
ATCF	Access Transfer Control Function (on Control Plane)
ATGW	Access Transfer Gate Way (on User Plane)
BICC	Bearer Independent Call Control
CMR	Codec Mode Request (used AMR and AMR-WB and EVS)
eNB	evolved Node Base-station
EVS	Enhanced Voice Services (Codec)
EVS-CMR	EVS Codec Mode Request (if needed to differentiate from AMR-WB-CMR)
EVS-ICM	EVS Initial Codec Mode
IAM	Initial Application Message (functionality-wise like SIP Invite)
MSC	Mobile Switching Center
RAN	Radio Access Network
SID	Silence Descriptor
SID-Con	SID-Conversion between EFR-SID and AMR-SID
sMSC	SRVCC MSC
TLCI	Transcoding-Less Codec Interworking
tMGW	Target Media GateWay
tRAN	Target RAN
UMTS_EVS	EVS Codec Type, applied in CS (UMTS) networks

4 eSRVCC Reference Architecture

Figure 4-1 shows the Reference Architecture for eSRVCC, as used in the present document. In this Reference Architecture the "SRVCC MSC" (sMSC) has direct control over the "Target RAN" (tRAN).

NOTE: In many life networks there is, however, another "Target MSC" inserted between the SRVCC MSC and the Target RAN. This has the advantage that only the SRVCC MSC has to be updated for the communication with MME and ATCF, while the Target MSC can be left SRVCC-agnostic. The interface between SRVCC MSC and Target MSC is as for any legacy Inter-MSC handover. It can be regarded in the context of the present document as a "solved problem" and so it is sufficient to concentrate on the shown Reference Architecture.

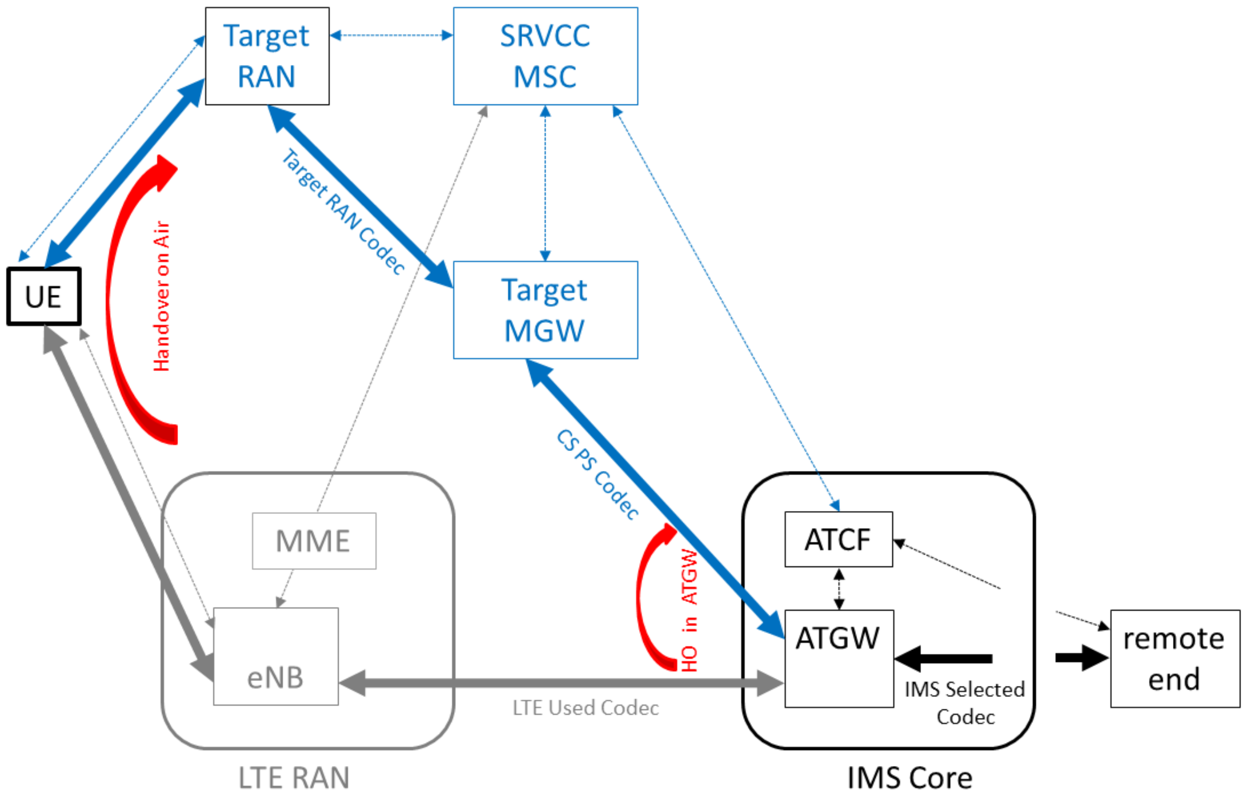

Figure 4-1: Reference Architecture for eSRVCC

Figure 4-1 introduces also terms to be used within the present document.

It is assumed that there is a VoLTE call already set up and ongoing between the UE at the "left side" of the ATGW and a partner at the remote end. The ATCF/ATGW are inserted in the call as Anchor, if eSRVCC is supported by all necessary nodes, especially the UE.

On the "left side" of the ATGW the so called "LTE Used Codec" is chosen. Candidates for the LTE Used Codec are primarily AMR(...), AMR-WB(...) and EVS(...).

On the "right side" of the ATGW the so called "IMS Selected Codec" is used to transport voice to/from the remote end. Candidates for the IMS Selected Codec are AMR(...), AMR-WB(...) and EVS(...), but also G.711, G.722 (e.g. if the remote party is fixed access terminal). Transcoding may be performed in the ATGW already before eSRVCC.

If all Codecs in the voice path are identical or TLCI-compatible (see chapter 11), then end-to-end TLCI is reached with the best possible voice quality under the given constraints. If LTE Used Codec and IMS Selected Codec are not TLCI-compatible, then the ATGW inserts transcoding.

Real life call scenarios at VoLTE setup might be quite complex. The control and media path between ATGW and remote end might be "long", e.g. due to call forwarding or roaming.

In order to keep eSRVCC execution delay and speech path interruption short, the ATCF and ATGW are inserted into the voice path, "as close as possible" to the local LTE RAN. This measure isolates the local eSRVCC from the rest of the control and media path, until eSRVCC is completed. ATCF and ATGW are the "Anchors" at this side of the call. They stay in the media and signalling path before, during and after eSRVCC.

Figure 4-1 defines also the terms "Target RAN Codec" and "CS PS Codec". Those codecs are used after eSRVCC on the interfaces indicated in the figure. If the chosen Target RAN Codec and the IMS Selected Codec are not TLCI-compatible, then either the Target MGW or the ATGW has to transcode. In the worst case there is a third non-compatible codec between them and two transcoding stages are required. In the best case Target RAN Codec and IMS Selected Codec are TLCI-compatible and no transcoding is needed.

Note that two "Handover Switching Points" exist, as in every handover.

- One is the "Handover on Air": The local UE disconnects from the LTE RAN and reconnects to the Target RAN (here GERAN or UTRAN).
- The other is the "Handover in the ATGW". It is theoretically and practically impossible (!) to synchronize both Handover Switching Points in time exactly.

Please note that the local UE is not connected to both radio accesses simultaneously, as the figure seems to suggest. "Single Radio" connectivity is the basis for eSRVCC.

5 eSRVCC Reference Procedure

5.1 General

Figure 5.1-1 is a direct reprint of 3GPP TS 23.216 [3], figure 6.2.2.1-1, showing the essential eSRVCC for the simplest case of an active voice call, without a parallel data session, from LTE to GERAN.

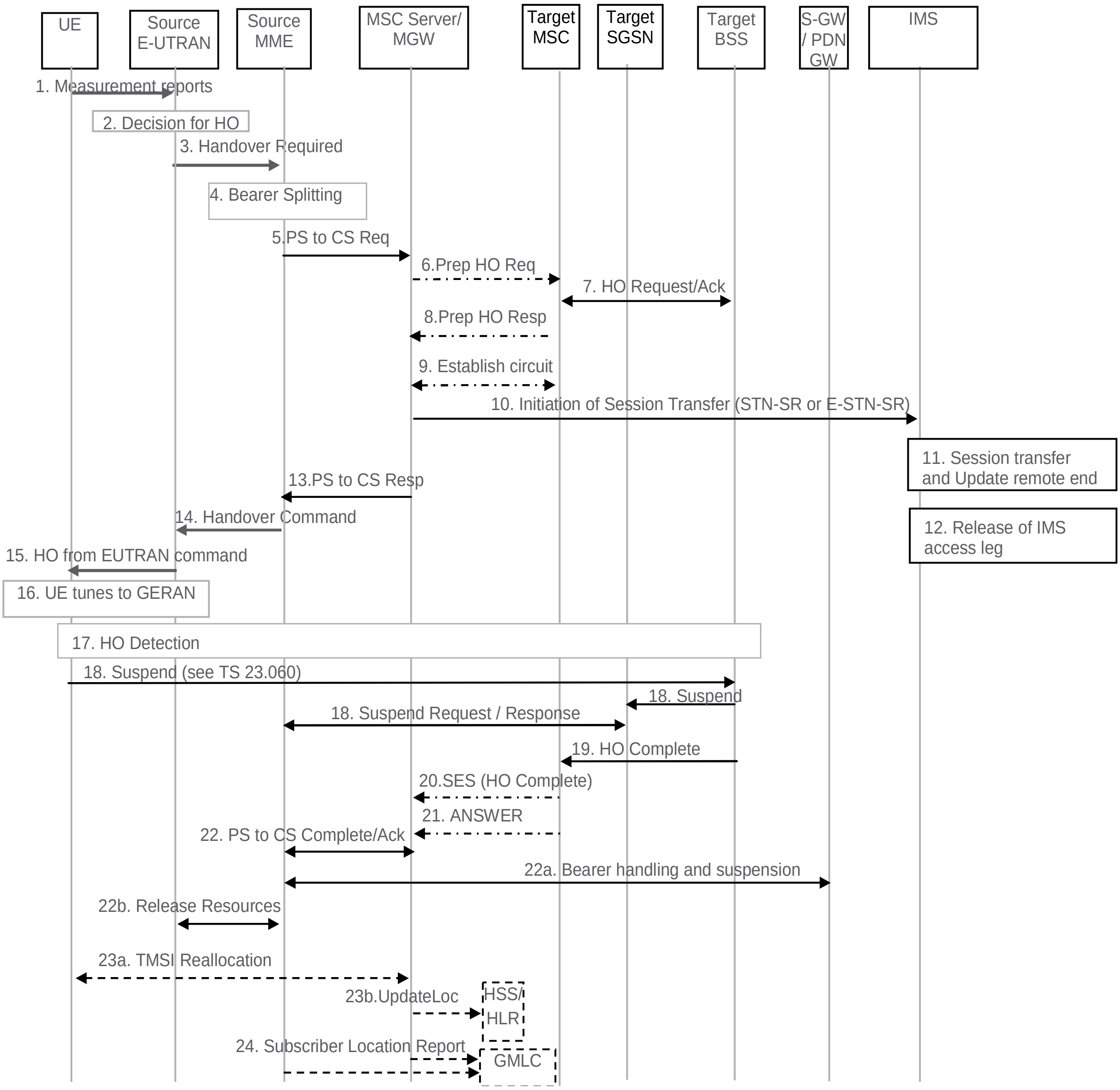

Figure 5.1-1: SRVCC from E-UTRAN to GERANwithout DTM support

Figure 5.1-2 is a substantially simplified version of 3GPP TS 23.216 [3], figure 6.2.2.1-1, focusing on the purpose of the present document, referring to the simplified Reference Architecture and the introduced terms.

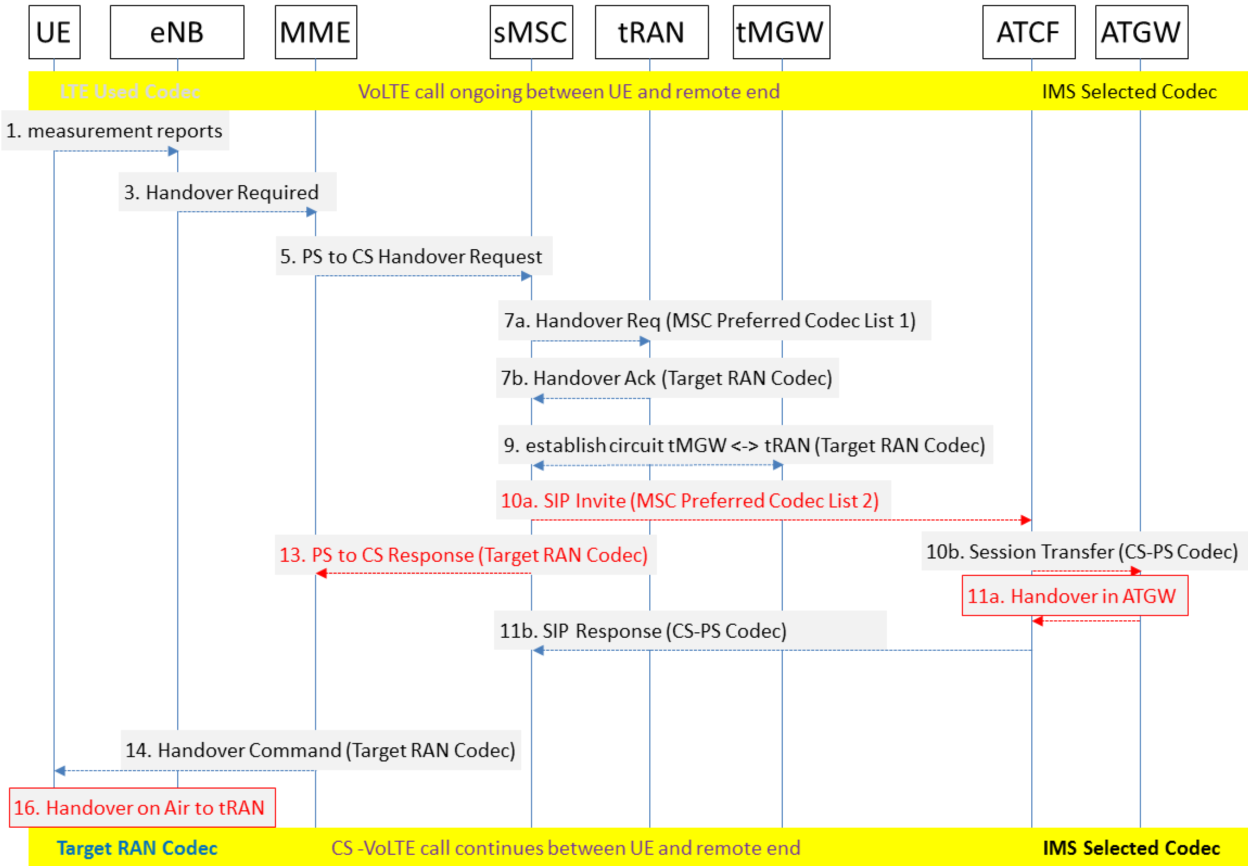

Figure 5.1-2: Reference Procedure of eSRVCC from LTE to GERAN

In this simplified version message 13 "PS to CS Response" is sent by the Target MSC **before** it got confirmation from the ATCF by message 11b "SIP Response". This is Stage 2 behaviour. The idea behind this timing sequence is to synchronize the handover in the ATGW as close as possible with the handover on air.

This Reference Procedure, shown in Figure 5.1-2, will be used as basis in the present document.

5.2 Codec Selection during eSRVCC

The local UE is moving through the radio networks and is continuously observing and measuring its radio environment. It is reporting these measurements to the LTE base station (eNB). Some when the eNB may decide that a GERAN (or UTRAN) radio cell is better suited for the voice call and may send a "Handover Required" message to the MME, including the wanted Target Radio. The MME sends this information to the relevant SRVCC MSC as PS-to-CS Handover Request message.

This **PS-to-CS Handover Request** message contains also the "**UE Supported Codec List**" (UE-SCL), as supported by the Local UE for the Target Radio Network(s), i.e. for GERAN and/or UTRAN.

The UE-SCL may contain all specified GERAN Codecs:

- FR_AMR-WB, FR_AMR, HR_AMR, EFR, HR, FR.

The UE-SCL may contain all specified UTRAN Codecs:

- UMTS_AMR-WB, UMTS_AMR2, UMTS_AMR and UMTS_EVS.

This PS-to-CS Handover Request message does not include the IMS Selected Codec and not the LTE Used Codec, because the MME has no knowledge about the Application Layer.

The Target MSC decides, based on the received UE-SCL and the known Target RAN Capabilities, which Codec is the **locally optimal** Codec for the Target RAN. This Target RAN Codec is based on local Target RAN criteria, without sufficient knowledge about the ongoing call.

The MSC takes the best possible Codec Type and Configuration, as locally preferred (set by the operator) for the Target RAN, given the received UE-SCL. Assignment Request is sent to GERAN (or RAB Assignment to UTRAN) and the voice path between Target RAN and Target MGW is setup, including all necessary details on Target MGW Context, MGW Termination properties, IP addresses1, UDP Ports1 and whatever is required.

Then, when all these preparations are done, the MSC sends a **SIP Invite** message to the ATCF to initiate the session transfer. This SIP Invite contains the so called "**MSC Preferred Codec List2**" (MSC-PCL2), with the Target RAN Codec on first place (i.e. most preferred). It also contains the connectivity data of the Target MGW (IP Address2 and UDP Ports2, etc.).

This MSC-PCL may contain at least the Target RAN Codec (or the SIP representative of it). Typically it contains many more Codecs, like AMR-WB(0,1,2), G.711, G.722, maybe more, depending on the Target MGW and its Transcoding capabilities. In some implementations even different Configurations of the AMR are included, like AMR(0,2,4,7), AMR(0,2,4), AMR(0,2), AMR(7), AMR(), even AMR(0,1,2,3,4,5,6,7) has been observed; similar for AMR-WB and UMTS_EVS.

The ATCF/ATGW-pair takes the MSC-PCL2 and decides on own capabilities (ATGW Supported Codecs, ATGW Supported Transcodings, whatever), considering the IMS Selected Codec, which Codec to use as CS PS Codec.

Then the ATCF sends the SIP Response back to the SRVCC MSC, including the CS PS Codec, including the connectivity data of the ATGW (IP Address3 and UDP Ports3, etc.).

In the ideal case IMS Selected Codec, CS PS Codec and Target RAN Codec are TCLI-compatible and the call continues after eSRVCC without Transcoding (at least at this end of the call).

5.3 Voice Path Switching during eSRVCC

As long as the SRVCC MSC prepares the Target RAN leg and the voice path between Target RAN and Target MGW, the call continues on the LTE access leg without disturbance by these eSRVCC preparation procedures. If these procedures take some longer time, e.g. due to network load, then the voice path switching is shifted in time, but this has no influence on the voice path interruption. This phase of eSRVCC preparation is rather uncritical. Of course: waiting too long might result in a lost LTE connection, before the new connection is up; in that case the call is lost.

Then at some time (denoted as "T_0" in what follows) the MSC sends the **SIP Invite to the ATCF** and the **PS-to-CS Handover Response to the MME**. According to Stage 2 description both messages are sent more or less at the same time.

The PS-to-CS Handover Response forwards the necessary parameters, like Target Cell and Target RAN Codec to the UE in the **Handover Command**. The ATGW switches the call leg from the LTE access towards the Target MGW, when the ATCF sends the **SIP Response** back to the MSC.

The Handover Command, after travelling through the LTE access, triggers the UE to change to the prepared Target RAN channel. How fast the UE changes, is implementation dependent.

Shortly after T_0 the voice path downlink to the LTE access is interrupted by the ATGW. The LTE "pipe", notably the sender buffers in ATGW and eNB may have still some few speech packets stored to be sent. So the speech path interruption will be observed some time later at the radio input of the UE and some processing time later at the loudspeaker output of the UE, here at "T_1". A substantial part of the processing time might be hidden inside the Adaptive Jitter Buffer (AJB) within the UE. The time difference between "T_0" and "T_1" varies, depending on LTE parameter setting, the cell load and actual LTE radio performance, between about 40 ms and (much) more than 100 ms.

Shortly after "T_0" also the voice path uplink to the remote end is interrupted in the ATGW. There could be still some speech packets in the pipe from the UE to the ATGW, notably inside the UE, but these are ignored by the ATGW.

The pipe from ATGW to the remote end might have a long delay, depending on the voice path and the remote access technology. At time "T_2" the Decoder at the remote end runs empty and the voice output gets muted. The delay between ATGW and remote end has no influence on the **duration** of the interruption at the remote end.

Shortly after "T_0" also the voice path pipe downlink to the Target MGW is filled with speech packets, coming from the remote end. So the downlink pipe of the Target Radio leg is started to be filled. It takes in the order of 100ms, until the first speech frame can be sent onto air.

Like the speech packets travelling with finite speed through the LTE RAN, also the Handover Command takes a while across the LTE radio access, depending on load and radio conditions. There is a "racing problem" between Control Plane and User Plane. In fact the race starts, when the SRVCC MSC sends the SIP Invite and PS-to-CS Handover Response. Ideally the UE would get the Handover Command at the same time as the last speech frame from the ATGW and would switch immediately after that to the new Target RAN leg. In real life networks that cannot be guaranteed.

As soon as the UE accesses the Target RAN the radio connection is established and downlink speech packets may arrive at the UE - depending how fast the downlink pipe is filled. Also in uplink the UE starts to send speech packets and fill the uplink pipe.

According to the eSRVCC standard, however, the uplink path in the Target MGW is blocked, until the MSC has received a "Handover Complete" message from the UE. Then speech packets are through-connected. They arrive at the ATGW and are forwarded to the remote end. When they finally arrive at the remote end the uplink speech break ends.

5.4 Possibilities to adjust codecs after eSRVCC without standards extensions

5.4.1 IMS Selected Codec re-negotiation towards the remote end

Figure 5.4.1-1 is applicable when the remote end supports the selected Target RAN codec (B) in the Re-INVITE.

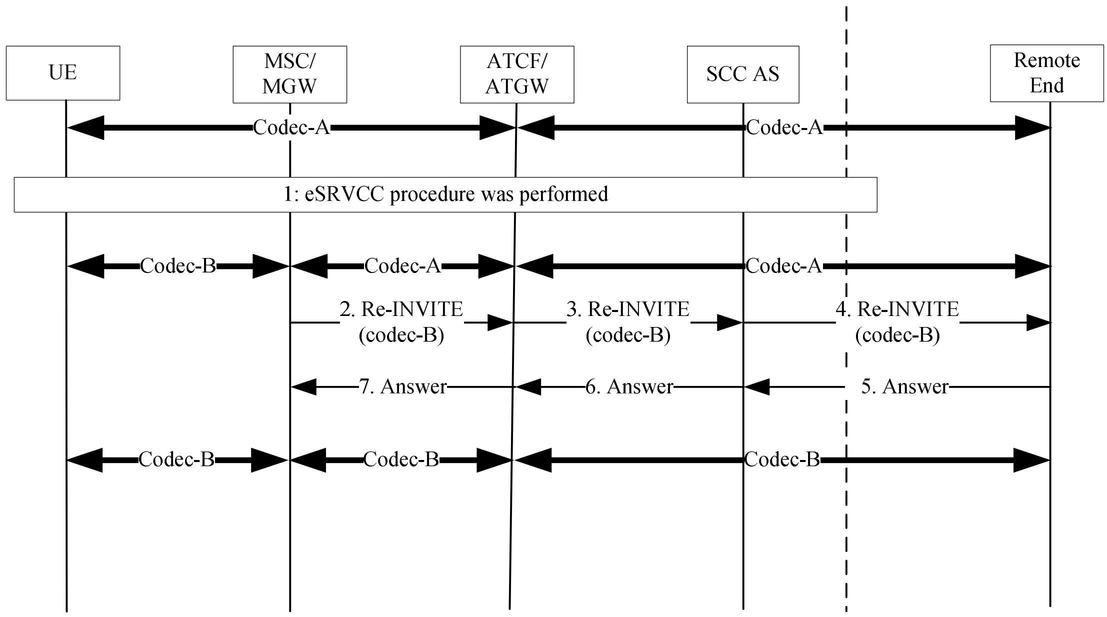

Figure 5.4.1-1: Re-negotiation method towards the remote end

1. eSRVCC is performed as standardized. A Target RAN Codec (B) is selected "blindly" that is not TLCI-compatible to the IMS Selected Codec. The SRVCC MSC has included this Target RAN Codec and all other supported codecs into SIP the session transfer request to the ATCF. The MSC Supported codec list includes also the IMS Selected Codec that is currently used in the ongoing IMS session (or a TLCI-compatible one). The ATCF has selected this IMS Selected Codec in the session transfer response, therefore there is no transcoding in ATGW, but there is transcoding in the CS-MGW. The session between UE and CS-MGW uses the Target RAN Codec (B). The session between CS-MGW, ATGW and remote end uses the IMS Selected Codec (A).

2. The MSC server sends a Re-INVITE towards the remote end with the list of supported codecs in the SRVCC MSC to ATCF, with the Target RAN Codec (B) as the most preferred codec in the list.

3. ATCF passes the Re-INVITE towards the SCC AS with the codec list.

4. SCC AS performs a remote leg update towards the remote end.

5-7. The remote end accepts the offer and selects the most preferred codec it can support, in this case (hopefully) the Target RAN Codec B (or a TLCI-compatible one) was selected. From now on the Target RAN Codec (B) is used e2e in TLCI manner.

NOTE 1: The second Codec change can interrupt the voice path a second time.

NOTE 2: It cannot be excluded that the Target RAN Codec (B) is not supported by the remote end (or a path in between), although potentially a third Codec could be a common Codec end-to-end.

5.4.2 Codec re-negotiation towards the SRVCC UE

Figure 5.4.2-1 is applicable when the remote end does not support any of the offered codec in the Re-INVITE, or the remote end selects a codec that was not in use in RAN and SRVCC UE, i.e. the new IMS Selected Codec is not TLCI-compatible to the Target RAN Codec (B) and the re-negotition was not successful.

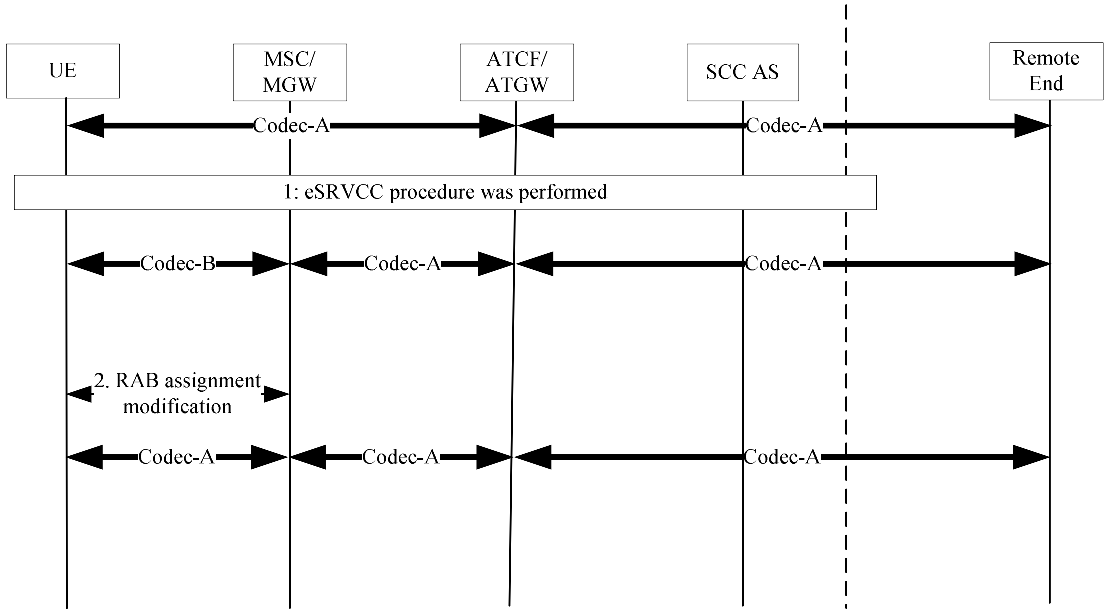

Figure 5.4.2-1: Re-negotiation of the Target RAN Codec towards the SRVCC UE

1. As in figure 5.4.1-1. The SRVCC MSC may also attempt via re-Invite to modify the IMS Selected Codec at the remote end towards the Target RAN Codec (B) and only execute step 2, if this attempt fails or ends in a new IMS Selected Codec (A*) that is again not TLCI-compatible (see steps 2 to 7 in figure 5.4.1-1).
2. The SRVCC MSC determines that TLCI was not achieved yet, and it decides to update the Target RAN Codec by Mid-call modification and RAB assignment modification procedure. UE and RAN accept the new Target RAN Codec A*. Codec A* is now used e2e in TLCI manner.

NOTE: This Mid-call modification in step 2 interrupts the voice path again, potentially a third time. The signalling effort is substantial.

6 Selected example scenarios for eSRVCC

6.1 General

In the following clauses a series of example scenarios is presented. The clause headlines have the following convention:

6.x eSRVCC <IMS Selected Codec> to <Target RAN Codec>.

EXAMPLE: 6.2 eSRVCC AMR(...) to FR_AMR(...)

In many cases it is immediately obvious that transcoding is required after eSRVCC, in some cases transcoding depends on the Codec Configurations, like AMR(0,2,4,7) to UMTS_AMR2(0,2,5,7), which requires transcoding, although the Codec Types are identical or at least from the same Codec Family.

The first scenarios are de facto the prototypes for all the others. These will be discussed more intensively; the others follow then the same principles, with differences.

In all scenarios a voice call is setup and in operation, with an LTE RAN on the local side, as shown in Figure 4-1. Local side means: the side, where the eSRVCC is executed. For simplicity of the discussion it is assumed than no other session to this local UE is setup. The local UE indicated support for eSRVCC and the IMS Core has inserted an ATCF/ATGW pair as local Anchor of the call. The call setup negotiation ended in the IMS Selected Codec as assumed in each scenario. The local UE is assumed to support all currently standardized 3GPP Codecs in 2G and 3G and 4G.

6.2 eSRVCC AMR(...) to AMR(...)

The **IMS Selected Codec** is in this example **AMR(...)**, with different possible mode-sets. There are more than 50 AMR Configurations thinkable, only few of them have real life relevance and only one of these is recommended, even mandatory for 3GPP GERAN networks: mode-set=0,2,4,7. Operators have the choice to influence the AMR Configuration in the IMS Core. Inter-Operator calls should be considered in this choice, as well as subsequent eSRVCC to CS networks and sub-subsequent Intra-CS Handovers.

The **LTE Used Codec** is here also **AMR(...)**, typically with the same Configuration as for the IMS Selected Codec. In fact there is no obvious reason, why the configurations should be different; in principle it is possible. The LTE Used Codec will discontinue existing due to eSRVCC; remaining is the IMS Selected Codec. If there would be a difference between LTE Used Codec and IMS Selected Codec and transcoding would exist in the ATGW, then this would be irrelevant after eSRVCC. It is assumed here that the LTE Used Codec and the IMS Selected Codec use the same AMR Configuration.

The remote end determines to a large extent the IMS Selected Codec, assuming that the local UE and the IMS network are capable of all mandated and recommended 3GPP Codecs: AMR(...), AMR-WB(...) and EVS(...) including EVS-IO(...).

Also the voice path between the shown IMS Core and the remote end has substantial influence, especially, if the call crosses network boundaries. These questions are, however, not discussed in the present document.

In the ideal case IMS Selected Codec, CS PS Codec and Target RAN Codec are TLCI-compatible and the call continues after eSRVCC without Transcoding; recommended: AMR(0,2,4,7), or subsets, everywhere, see table 6.2-1.

Other used AMR Codec Configurations are AMR (0,2,5,7) and AMR (7). Also these may be used in homogenous 3G- and IMS- networks in such an "ideal eSRVCC" scenario. However, both of these AMR Configurations are not supported in GERAN and not used in many UTRAN deployments, and thus frequently necessitate transcoding when interworking with other networks.

Table 6.2-1: eSRVCC result for the recommended AMR(0,2,4,7) to AMR(0,2,4,7) or a sub-set

Target RAN Codec	TLCI ?	CS PS Codec	TLCI?	IMS Selected Codec
UMTS_AMR2 (0,2,4,7)	**yes**	AMR (0,2,4,7)	**yes**	AMR (0,2,4,7)
FR_AMR (0,2,4,7)	**yes**	AMR (0,2,4,7)	**yes**	AMR (0,2,4,7)
HR_AMR (0,2,4)	**yes**	AMR (0,2,4,7)	**yes**	AMR (0,2,4,7)
UMTS_AMR2 (0,2)	**yes**	AMR (0,2,4,7)	**yes**	AMR (0,2,4,7)

Although the call continues in all these cases without Transcoding, the maximum bit rate may be very different, depending on the load situation in the Target RAN. The effects of these differences are discussed in <clause xxx>.

AMR() or AMR(0,1,2,3,4,5,6,7) as IMS Selected Codec would not be TLCI-compatible to any CS network.

A discussion on this issue is ongoing in 3GPP and GSMA.

Either the IMS network or the terminating PS-UE can select an AMR configuration. The IMS network can select a specific AMR configuration, like AMR (0,2,4,7) by modifying the original Open SDP offer.

If the terminating PS-UE selects the default AMR(0,2,4,7), see 3GPP TS 26.114 [5]), then this can avoid transcoding in subsequent eSRVCC in networks that support the default configuration, but may necessitate transcoding after eSRVCC in networks where other AMR configurations are used in the CS radio network.

If the originating network select any configuration suitable for local eSRVCC in the offer phase, it increases the risk of interworking problems (a need for transcoding) with other networks even before eSRVCC.

The only - and simple - solution to this "Gordian-knot" is to agree to one unique AMR Configuration across all operator networks. The best "Golden Compromise" is AMR (0,2,4,7).

6.3 eSRVCC AMR(...) to AMR-WB(...)

As in scenario 6.2 the **IMS Selected Codec** is **AMR(...)**, let's assume it is AMR(0,2,4,7), the recommended Codec.

As described in clause 5.2 the SRVCC MSC determines the Target RAN Codec based on the received UE-SCL and the known Target RAN Capabilities without knowledge about the IMS Selected Codec. The MSC takes the best possible Codec and Configuration, as locally preferred (set by the operator) for the Target RAN, given the received UE-SCL.

If the Target RAN is updated to FR_AMR-WB(0,1,2) and/or UMTS_AMR-WB(0,1,2), but not to even better Codecs, then one of these will be selected by the SRVCC MSC as Target RAN Codec and the Target RAN leg will be prepared. In SIP Invite towards the ATCF this Codec will be listed as AMR-WB(0,1,2).

The SRVCC MSC will send the SIP Invite to the ATCF, with the MSC-PCL containing the AMR-WB(0,1,2) on first place, followed by other Codecs, see clause 5.2.

The ATCF has no other possibility than to insert Transcoding between Target RAN Codec and IMS Selected Codec; the only freedom left is where to place the transcoding. From call setup it is obvious that the remote end does not support a WB Codec, because otherwise AMR-WB would have been the IMS Selected Codec. **Therefore it is not reasonable trying to re-negotiate the IMS Selected Codec with the remote end.**

The ATCF could select the AMR-WB(0,1,2) as CS PS Codec, taking the burden of Transcoding fully into the ATGW. The ATCF could select the AMR(0,2,4,7) as CS PS Codec, shifting the burden of Transcoding fully into the Target MGW.

The third choice, for completeness, if offered by the MSC, would be to select an "intermediate" Codec as CS PS Codec, such as G.711 or G.722 or "lin.PCM128", with 8 kHz sampling and 16 bit "linear" resolution == 128 kbps.

Table 6.3-1: eSRVCC result for the recommended AMR(0,2,4,7) to AMR-WB(0,1,2)

Target RAN Codec	TLCI ?	CS PS Codec	TLCI?	IMS Selected Codec
AMR-WB (0,1,2)	yes	AMR-WB (0,1,2)	no	AMR (0,2,4,7)
AMR-WB (0,1,2)	**no**	**AMR (0,2,4,7)**	**yes**	**AMR (0,2,4,7)**
AMR-WB (0,1,2)	no	lin.PCM128	no	AMR (0,2,4,7)

The choice is implementation dependent. Often the ATCF selects the IMS Selected Codec also as CS PS Codec. This is "egoistic", as the burden is shifted to the Target MGW. But it has a substantial advantage: it indicates to the SRVCC MSC that the choice of the Target RAN Codec was not optimal. The SRVC MSC has then the opportunity to execute a Mid-Call Modification of the Target RAN Codec to reach TLCI again, after eSRVCC is successfully executed.

So in this scenario eSRVCC is executed and transcoding resources are added, typically in the Target MGW. Then, after a short while, Mid-Call Modification of the Target RAN leg may remove the inserted Transcoder again. This additional Mid-Call Modification is implementation specific.

6.4 eSRVCC AMR(...) to UMTS_EVS(...)

As in scenario 6.1 and 6.2 the **IMS Selected Codec** is **AMR(...)**, e.g. AMR(0,2,4,7), the recommended Codec. The SRVCC MSC determines the Target RAN Codec based on the received UE-SCL and the known Target RAN Capabilities without knowledge about the IMS Selected Codec.

If the Target RAN is updated to UMTS_EVS, then this may be selected as Target RAN Codec. But, which of the Configurations (Set 0, Set 1, Set2 or Set 3, see TS 26.103[7]) would the SRVCC MSC select?

Set 0: with Spreading Factor SF=256, needs least radio capacity, provides lowest voice quality of all standardized Configurations: The UMTS_EVS (Set 0) Codec is equivalent to EVS (br=5.9-8; bw=nb-wb; mode-set=0). Narrowband and Wideband voice quality is provided up to 8 kbps, including EVS-IO (0), as well a Variable Bit Rate coding at an average rate of 5,9 kbps.

Set 1: with Spreading Factor SF=128, needs more radio capacity and is a decent compromise. The UMTS_EVS (Set 1) Codec is equivalent to EVS (br=5.9-13.2; bw=nb-swb; mode-set=0,1,2). Narrowband, Wideband and Super-Wideband voice quality is provided up to 13,2 kbps. EVS Variable Bit Rate, EVS Channel-Aware Mode of operation and EVS-IO up to 12,65 are supported.

Set 2: with Spreading Factor SF=64 provides the best possible quality in UTRAN and is optimal, if the IMS Selected Codec is EVS (br=5.9-24.4; bw=nb-fb; mode-set=0,1,2) or better. This UMTS_EVS (Set 2) is the most costly alternative for the Target RAN, including all the features of UMTS_EVS (Set 1) on a higher quality level.

Set 3: with Spreading Factor SF=128, tailor-made to guarantee Super-Wideband quality. The UMTS_EVS (Set 3) Codec is equivalent to EVS (br=9.6-13.2; bw=swb; mode-set=0,1,2). It has radio capacity demands comparable to or slightly higher than UMTS_EVS (Set 1); EVS Channel Aware Modes of operation is supported.. For interworking with legacy networks, EVS-IO up to 12,65 is supported.

The decision could and will be based on the load in the Target RAN. Sometimes there is no other choice than UMTS_EVS (Set 0), except the operator prefers UMTS_EVS (Set 3) and provides always sufficient radio capacity. Note that the SRVCC MSC may select UMTS_EVS (Set 2) as Target RAN Codec and the Target RAN restricts the RAB assignment to UMTS_EVS (Set 0) and informs the network by Rate Control commands.

The problems and solutions are similar, a bit more negative, compared to the scenario 6.3. The temporarily inserted Transcoder (EVS <=/=> AMR) is even more complex and resource hungry. The temporary radio load is potentially high without gain.

An optional Mid-Call Modification of the wrongly selected Target RAN Codec is the only escape, after such an eSRVCC as specified currently.

6.5 eSRVCC AMR-WB(...) to AMR(...)

In this scenario the call setup resulted in the IMS Selected Codec being AMR-WB(...). Maybe even AMR-WB() is selected, with all 9 modes allowed. This is an important scenario today in VoLTE<=>VoLTE calls. But also AMR-WB(0,1,2) provides impressive HD Voice quality.

Unfortunately, in this scenario, the Target RAN is not updated and does not support AMR-WB yet. The SRVCC MSC selects AMR(0,2,4,7) instead. Transcoding is required between Target RAN Codec and IMS Selected Codec.

Other than in the scenarios before (6.2 - 6.4) there is a chance to renegotiate the IMS Selected Codec with the remote end and achieve end-to-end TLCI again, although in AMR(0,2,4,7) quality.

This Codec Renegotiation is optional. In any case it should be performed after eSRVCC is successfully finished.

6.6 eSRVCC EVS (br=5.9-128; bw=nb-fb) to UMTS_EVS (Set1)

In this example scenario the call setup resulted in the **IMS Selected Codec** being EVS (br=5.9-128; bw=nb-fb), the biggest EVS Configuration with all four audio bandwidths included and all the bit rates, ranging from the lowest rate, 5,9 kbps (average), up to the highest, 128 kbps. In addition, the EVS-VBR and the EVS-CA modes are included, as well as the EVS-IO with all modes.

The call is ongoing with this biggest possible **EVS Framework Configuration**. Mode Control may be ongoing and the EVS modes in both directions may be different and lower than maximally possible, depending on external factors, such as audio-I/O capabilities and network load situations. The **active EVS Configurations** may be temporarily smaller and different in both directions, but transcoding is in no case needed.

Now eSRVCC is requested. The Target RAN supports UMTS_EVS.

Case 1: The Target RAN is not loaded. The Target MSC determines UMTS_EVS (Set 2) as **Target RAN Codec**, based on local RAN Capabilities and the UE Supported Codec List, but without any knowledge about the IMS Selected Codec or the LTE Used Codec. It is mainly by coincidence that the Target RAN Codec fits so well in this example. It can be easily shown, that all EVS Configuration, which include all modes and rates below an upper corner, are all TLCI-compatible to each other. Therefore the call continues after eSRVCC without transcoding, although the EVS Framework Configuration shrinks to EVS (br=5.9-24.4; bw=nb-fb; mode-set=0,1,2), still providing FB quality.

Case 2: The load in the Target RAN is higher. The MSC selects UMTS_EVS (Set 1) as Target RAN Codec. The call continues without transcoding in the best possible SWB quality under these conditions.

Case 3: The load in the Target RAN is much higher. The MSC selects UMTS_EVS (Set 0) as Target RAN Codec. The call continues without transcoding in the best possible WB quality, the best possible under these harsh load conditions.

Case 4: Although the MSC selects UMTS_EVS (Set 2) as Target RAN Codec, the RNC has the freedom (according to the strategy in life networks) to allow only a sub-set of the Target RAN Codec. This may end in the de facto Configuration of UMTS_EVS (Set 1) or even UMTS_EVS (Set 0) and the call would still continue in TLCI. The RNC would send Mode Control commands to keep the Codec Modes within these limits. Case 4 has the advantage that the RNC may subsequently modify the de facto Configuration up to UMTS_EVS (Set 2) without informing the MSC, by that upgrading the call quality seamless to the highest possible.

If only EVS Bottom up Configurations are used, in IMS and CS, which include all modes and rates below their individual upper corner of Rate and Bandwidth, then TLCI is always guaranteed before and after handover.

Important is that the MSC selects EVS only as Target RAN Codec, if the IMS Selected Codec is compatible. In order to do that it is indispensable that the MSC knows the IMS Selected Codec.

Mode Control keeps the active EVS Configurations within this new EVS Framework Configuration, although the IMS Selected Codec is still EVS (br=5.9-128; bw=nb-fb). There is no need to modify that from a speech quality point of view. Of course, it may happen during the call that the RNC restricts the upper bit rate temporarily due to varying cell load conditions, or the UE goes down in rate due to TX power problems, see TS 26.454 [10]. In these cases, the speech quality goes down or up as necessary. This is not different to the situation in a pure VoLTE call. In all cases the speech quality remains as high as possible.

Important is that the remote UE receives the necessary EVS-CMR, requesting the maximum bit rate and maximum audio bandwidth, as soon as possible and follows this EVS-CMR as soon as possible. If done well, it is possible to command the remote EVS client to use EVS modes within the range of the Target RAN Codec long enough **before** the local UE performs the eSRVCC handover on air.

This so-called "Pre-SRVCC Mode Control" could be triggered by the ATGW, if the ATGW gets early information about the Target RAN Codec. It may also be triggered by the Target MGW, after the ATGW has switched the radio legs.

6.7 eSRVCC EVS (br=16.4-128; bw=fb) to UMTS_EVS (Set1)

In this example scenario, the call setup by SIP/SDP negotiation resulted in the IMS Selected Codec being the biggest EVS FB-only Configuration, EVS (br=16,4-128; bw=fb). SDP excluded all bandwidths below FB and all bit rates below 16,4 kbps. It is generally not allowed that EVS-CMR could change this FB-only Configuration during the call.

The call quality may reach the same quality as in the EVS (br=5.9-128; bw=nb-fb) Configuration scenario in clause 6.6, using the highest EVS mode with 128 kbps and full band audio, but not higher. Transcoding is not needed. Mode Control may be ongoing, but the rate cannot be set lower than 16,4 kbps and the audio bandwidth is fixed to Fullband. High quality seems to be guaranteed. This is fact not the full truth. The following paragraph discusses this.

Due to the EVS algorithm design the EVS Encoder classifies the input audio signal and decides frame by frame, which audio bandwidth is actually given and where to put the "coding bit resources". It may well use a NB Codec mode and achieve optimal quality for a NB input signal. The adaptation follows the audio-input quite well - also for non-Full-band signals. The EVS FB-only Configuration does not prevent the media-sender using lower bandwidth modes. The Transport Plane (here RTP) and the MGWs in the path will support this. The quality is optimal, if the media-receiver has FB audio output capabilities.

The inband EVS-CMR cannot change the audio bandwidth, even if the audio output on the remote side would require it, e.g. because the remote user connects a legacy handsfree kit with lower bandwidth. Because coding bit resources are wasted by the local media-sender in audio signal regions, which the remote media-receiver cannot play back, the voice quality may not be optimal, but lower than in the scenario with EVS (br=5.9-128; bw=nb-fb).

If there would be a capacity problem along the speech path, rates below 16,4 are not available, also the EVS-CA mode is forbidden. The voice quality may well fall below the quality of the other Configuration due to a higher residual frame loss rate.

The high quality expectation is already without eSRVCC not always fulfilled by this (and other) punctured EVS Configuration EVS (br=16.4-128; bw=fb).

Now the network has to execute eSRVCC with this EVS (br=16.4-128; bw=fb) as IMS Selected Codec.

Remember: the Target MSC does not know the IMS Selected Codec.

The Target RAN supports UMTS_EVS and the load on the Target RAN is not too high, so for example the Configuration EVS (br=5.9-13.2; bw=nb-swb; mode-set=0,1,2), i.e. UMTS_EVS (Set 1) is determined as Target RAN Codec, same as in clause 6.6.

The IMS Selected Codec is not TLCI-compatible to this Target RAN Codec, because there is no common audio band and the lower bit rates are not common. The ATGW (or Target MGW) will insert Transcoding! Transcoding resources are quite expensive for EVS, involving two EVS Decoders and two EVS Encoders in the ATGW or Target MGW.
The SWB quality after eSRVCC is degraded below the maximum quality of the Target RAN Codec, it is lower than in the scenario with the Bottom up Configuration EVS (br=5.9-128; bw=nb-fb) as IMS Selected Codec.

Discussion of potential alternatives to avoid transcoding:

In this scenario the knowledge about the IMS Selected Codec would not help much, if the Target RAN had no other choice than SF=128, as there is no TLCI-compatible Codec available in the Target RAN for this EVS FB-only Configuration of the IMS Selected Codec. However, if the MSC would get knowledge about alternatives to the IMS Selected Codec, then an overall optimization could be considered by selecting first an optimal Target RAN Codec, followed then after eSRVCC by a renegotiation of the IMS Selected Codec. **The effort would be rather high, the resulting quality no better than with the Bottom up Configuration already at call setup**.

If the Target RAN would support SF=64, then the MSC could try deploying this, without knowing the IMS Selected Codec. Allocating this double radio capacity "blindly" is maybe not commercially reasonable, if the IMS Selected Codec would be EVS (br=5.9-128; bw=nb-fb) and SWB Quality would be a good enough compromise for 3G under the given load conditions.

In one alternative approach, the MSC could be tempted to select EVS (br=5.9-24.4; bw=nb-fb), i.e. UMTS_EVS (Set 2), as Target RAN Codec with Guaranteed Bit Rate (GBR)=16.4 kbps and "hope" the Target RAN would be able to accept and support it. In case of too high load, the Target RAN would reject this RAB Assignment.
The ATGW (or Target MGW) could send Pre-SRVCC Mode Control to steer the remote UE into EVS (br=16.4-24.4; bw=fb). The call **could** continue seamless in FB quality! However, as soon as the Target RAN would need to restrict the bit rate in downlink below 16,4 kbps **the call would break**, respectively end in one way muting. In order to avoid that, the MSC would have to set the **Guaranteed Bit Rate** in the Target RAN to 16,4 kbps.

The UE, however, could be tempted to improve uplink radio quality in case of TX power limitations. Without a clear rule, the UE could use lower rate and lower audio bandwidth in uplink. Clause 7.2 of TS 26.454 [10] has set such a rule in REL-13 for the "UE autonomous rate": it is indispensable that the UE obeys the commanded audio-bandwidth. Example: if the 3G-UE receives EVS-CMR (br=16.4; bw=fb) from the network, then it obeys the commanded bw=fb, even if the uplink TX power limit is reached and even if lower rates would be available in UMTS_EVS (Set 2). As a result, the frame loss rate in uplink (and downlink) may be high in marginal radio conditions. **This alternative is not satisfying** and not according to the EVS compatibility rules.

This Target RAN Codec UMTS_EVS (Set 2), with GBR=16.4 would also be sub-optimal for an IMS Selected Codec EVS (br=5.9-128; bw=nb-fb).

In another alternative the ATGW (or other MGW in the path) could send EVS-CMR commands to bring both ends into the **EVS-IO mode of operation**. This would bring the call into TLCI as well, with AMR-WB quality. It depends on the EVS Configurations, if the resulting WB quality is preferred. **In this example IMS Selected Codec, it would not be better**.

In this scenario, it would be clearly better to use an EVS Bottom up Configuration for the IMS Selected Codec.
All discussed alternatives are worse.

Without knowledge about the IMS Selected Codec, the Target network cannot decide, which Target RAN Configuration for EVS is optimal. Without knowledge about the Target RAN Capabilities, the ATCF/ATGW cannot decide on Pre-SRVCC Mode Control either.

6.8 eSRVCC EVS (br=9.6-24.4; bw=swb) to UMTS_EVS (Set1)

Here the **IMS Selected Codec** has the punctured Configuration **EVS (br=9.6-24.4; bw=swb)**, based on operator policy**.** This is TCLI-compatible to UMTS_EVS (Set 3). Assumedly, the operator sets the parameters in all his network parts consistently, in IMS and in CS. Interworking with other operators should be taken into account.

The network has to execute eSRVCC.

Case 1: The Target RAN supports UMTS_EVS and the load on the Target RAN is not too high. Based on operator policy the MSC prefers UMTS_EVS (Set 3) as Target RAN Codec. This fits perfectly to the IMS Selected Codec, by some coincidence, as the IMS Selected Codec was unknown. It could have been AMR or AMR-WB or other, then this Target RAN Codec would be not that good.

Pre-SRVCC Mode Control is necessary to bring the remote end into the Target Codec bit rate range before the handover is performed.

Case 2: If the Target RAN is highly loaded and another EVS Configuration will be chosen, like UMTS_EVS (Set 0), then transcoding is neededd. The quality ends up below the quality of the Target RAN Codec.

NOTE: Since the operator has, based on his policy, provided sufficient capacity in Target RAN, case 2 will not occur often or not at all in this network. Nevertheless: Under such good radio conditions, which avoid case 2, also the Bottom up Configurations EVS (br=5.9-24.4; bw=nb-swb) as IMS Selected Codec and EVS (br=5.9-13.2; bw=nb-swb), i.e. UMTS_EVS (Set 1) as Target RAN Codec would not use bit rates and bandwidth worse than SWB. If radio conditions would be worse and even bad, as in the unlikely case 2, then the Mode Control would automatically use a smaller Button up Configuration, like EVS (br=5.9-8, bw=nb-wb) without transcoding, providing best possible quality in this bad conditions.

In all conditions, the resulting quality with the Bottom up Configurations up to SWB would be as good as or better than with the punctured SWB-only Configurations.

6.9 eSRVCC EVS (...) to AMR-WB (...)

Here any EVS Configuration could be selected as **IMS Selected Codec**, because all include the mandatory EVS AMR-WB IO mode of operation. Important is that the mode-set was reasonably set to include the lower modes of EVS AMR-WB IO, ideally mode-set=0,1,2. Additional modes may be included, maybe all.

The network has to execute eSRVCC.

The Target RAN supports AMR-WB and the load on the Target RAN is not too high. Based on operator policy the SRVCC MSC selects AMR-WB (0,1,2) or AMR-WB (0,1,2,4) or AMR-WB (0,1,2,8) as Target RAN Codec. All these configurations do not require transcoding towards an IMS Selected Codec as recommended above; the AMR-WB IO modes can be adjusted to a lower range via EVS-CMR.

Pre-eSRVCC Mode Control is preferable to bring the remote end into the Target RAN Codec bit rate range and into the EVS AMR-WB IO mode of operation, **before** the handover on air is performed. If Pre-eSRVCC Mode Control is not possible (as today), then the voice path interruption is longer than necessary, but the call will continue in TLCI end-to-end.

6.10 eSRVCC EVS (...) to AMR (...)

This scenario is similar to scenarios above, where Transcoding is needed immediately after eSRVCC. The reasons, why the eSRVCC choses the AMR (...) as Target RAN Codec may be either overload in the Target RAN or missing support for AMR-WB and UMTS_EVS in the target RAN. Or - of course - the missing information about the IMS Selected Codec.

In any case, it can be assumed that the remote end supports AMR with high likelihood (otherwise EVS would not be the IMS Selected Codec). A Re-Invite towards the remote end seems to be promising, see clause 5.4. This re-Invite could be triggered by the ATCF or SRVCC MSC.

6.11 eSRVCC EVS (br=5.9-128; bw=nb-fb) to UMTS_EVS (Set 1) and subsequent Handover to AMR-WB(0,1,2)

In this example scenario, the **IMS Selected Codec** is EVS (br=5.9-128; bw=nb-fb), the biggest EVS Configuration with all four audio bandwidths included and the bit rate ranging from the lowest rate, 5,6 kbps, up to the highest, 128 kbps. The call is ongoing with FB quality.

The local mobile is moving and leaving LTE coverage. The network performs eSRVCC as in clause 6.6 to the UMTS_EVS (Set 1) as **Target RAN Codec** and the call continues after eSRVCC without transcoding in SWB quality. EVS-CMR controls the now reduced Framework Configuration.

However, the mobile is moving on and is even leaving 3G-coverage into 2G-coverage. Another handover follows, this time a CS-internal Inter-RAT handover, to a Target RAN2, with AMR-WB(0,1,2) as Target RAN2 Codec. Without going into details here, the call may continue in HD Voice quality (WB quality), without transcoding, with the EVS Primary mode of operation in the IMS Selected Codec replaced seamlessly by the EVS AMR-WB IO (0,1,2) mode of operation. The Target RAN sends AMR-WB-CMR=2 (or smaller) towards the remote end, together with AMR-WB-coded speech in RTP packets according to IETF RFC 4867 [9]. A MGW in the path (e.g. the Target MGW of the preceding eSRVCC) repacks these AMR-WB-RTP packets into EVS-RTP packets according to 3GPP TS 26.445 [8] and translates the AMR-WB-CMR ≤ 2 into the EVS-CMR for the EVS AMR-WB IO mode with maximum bit rate 2 (or smaller).

These two handovers reduced the voice quality from FB to SWB and finally to WB. In all these scenarios, the quality was and is as good as possible under the given circumstances, always transcoding free. The eSRVCC used by coincidence a TLCI-compatible Target RAN Codec, while the Inter-RAT handover from UTRAN to GERAN has exact knowledge about the Selected Codec and selects the Target RAN2 Codec precisely.

Although the remote LTE UE (or a remote client in a wireline terminal) may have still excellent (radio) link quality, allowing EVS (br=5.9-128; bw=nb-fb) still, it is indispensable that the remote UE (client) obeys the received EVS AMR-WB IO EVS-CMR as soon as possible and strictly. Only then, the eSRVCC and subsequent CS-internal handover are executable with minimal speech break time and without Transcoding. If the remote LTE UE would not follow the received EVS-CMR strictly, then the call would go muting on the side, where the handover reduced the EVS Configuration in size. It is inacceptable that the remote UE would change from the EVS AMR-WB IO mode to an EVS primary mode without explicit command by EVS-CMR or a SIP renegotiation.

After a while, the UE moves back into 3G coverage. The CS-network performs another Inter-RAT handover, selecting the UMTS_EVS (Set 1) as **Target RAN3 Codec.** Mode Control takes care that the remote end remains in the EVS AMR-WB IO mode, until the UE safely landed in the 3G network. Then the 3G UE sends EVS-CMR to the remote end to switch to EVS (br=5.9-13.2; bw=nb-swb) and the call continues in both directions in SWB quality.

It should be noted that in an "upgrading" handover, the Mode Control follows (ideally) the handover and in a "downgrading" handover the Mode Control precedes (ideally) the handover.

6.12 eSRVCC and Handover in speech pauses

With quite some likelihood, these local handover may occur in phases, where the local UE detected a speech pause and does not send anything in uplink, except a SID frame every now and then.

In such a case, the handover-handling MGW should send CMR≤x towards the remote end in several CMR-Only frames and SID frames to accelerate the fall-back to lower modes as much as possible. Without these inserted CMR-Only frames (CMR-Only are No_Data frames including only the CMR), the handover-handling MGW would have to wait for the next SID frame and that might take quite a while. This would increase the speech break time in the local downlink direction. A lost SID or a lost CMR-Only frame would also mean the CMR is lost, which would cause a delay of the adaptation and therefore a longer speech break. Therefore this CMR is repeated several times (forward error correction by repetition) in several CMR-Only and/or SID frames. The repetition could be continued, until the remote end reacted accordingly. It is important that these CMR-Only frames are carried in the RTP packets all the way to the remote LTE UE.

The EVS standards allow extracting the CMR from the received RTP packets and sending CMR in EVS RTCP-APP, if AVPF is allowed. This requires more effort, more transport bandwidth and takes in general a noticeable longer time to reach the remote UE. **CMR within RTP is substantially faster, more error robust and simpler to handle**.

Important to note: the current text in 3GPP TS 26.445 [8], clause A.2.2.1.2 ToC byte, states:
Begin of cite (important part in bold):

*"**Packets containing only NO_DATA frames should not be transmitted in any payload format configuration**. Frame-blocks containing only NO_DATA frames at the end of the packet should not be transmitted in any payload format configuration. In addition, frame blocks containing only NO_DATA frames in the beginning of the packet should not be included in the payload."*

End of cite.

This paragraph could potentially be misunderstood in a way that RTP packets including only the CMR byte should not be transmitted. In order to avoid misunderstanding, it should be added:

"Packets without speech data, containing only the CMR byte, are to be transmitted."

The details, when and how often CMR-only packets are to be sent, are for further study, see discussion above.

7 Identified Problems with current eSRVCC

7.1 General

This clause summarizes the identified problems with the current 3GPP standard procedures for eSRVCC, based on the discussion of the example eSRVCC scenarios in clause 6.

7.2 IMS Selected Codec not known in Target RAN

7.2.0 General

Figure 5.1-2, Reference Procedure for eSRVCC, shows that the MME informs the eSRVCC MSC (sMSC) with the PS to CS Handover Request in Message 5. Message 5 contains the UE Supported Codec List and the Target RAN cell(s), but it contains no information about the ongoing call, except that it is a voice call and which call it is (call identifier), but no information about the IMS Selected Codec.

This is in contrast to legacy CS handover procedures, where the Source Network informs the Target Network about the Source Used Codec or the (CS-) Selected Codec. The current eSRVCC procedure therefore cannot match the performance of legacy handover.

Without knowledge about the IMS Selected Codec, the Target RAN Codec cannot be selected optimally.

7.2.1 Remote Access Network supports only lower quality codecs than Target RAN

In most of today's networks, the interworking with legacy CS network(s) of many flavours and capabilities on the remote end is essential. Often the local VoLTE-UE is connected via IMS to a remote partner with reduced Codec capabilities, such as NB Codecs, e.g. AMR or G.711 (PCM). Thus, the IMS Selected Codec has reduced capabilities. The voice quality of the call before eSRVCC is optimal under the given circumstances, but worse than the local UE supports.

Then the eSRVCC to a Target RAN with **better** capabilities, like AMR-WB or even UMTS_EVS, unavoidably ends in the selection of a Target RAN Codec that is too good (!), with the unexpected consequence of **even lower quality** at higher resource cost and higher speech path delay, due to the necessary transcoding.

The early knowledge of the lower quality IMS Selected Codec would improve the situation noticeably in all respects. The Target RAN Codec would match optimally to the IMS Selected Codec, avoiding transcoding, achieving better quality than by using the best available Target RAN Codec.

The eSRVCC would then immediately land in the best possible voice quality, given the constraints of the remote end. The voice quality would typically not change due to eSRVCC.

If the IMS Selected Codec is equal or worse than the Target RAN capabilities and the eSRVCC SC is informed about the IMS Selected Codec in due time, then the Target RAN Codec can be selected optimally in one step.

7.2.2 Remote Access Network supports higher quality codecs than Target RAN

However, just informing the Target Network about the IMS Selected Codec is not sufficient for many scenarios, where the remote end has better capabilities, than the Target RAN. An example is the VoLTE <=> VoLTE call with AMR-WB() or EVS (br=5.9-24.4; bw=nb-swb) as IMS Selected Codec and the subsequent eSRVCC to a Target RAN with only NB Codecs, like AMR or EFR.

In such a case, the Target Network may not even be in a position to understand the IMS Selected Codec. For example, the legacy SRVCC MSC in a GSM Network does not know EVS. Therefore it would be helpful, even necessary, to inform the SRVCC MSC also about alternative Codec candidates with the "IMS Preferred Codec List", where the IMS Selected Codec is at first place, followed by (all) other Codec candidates. The Target Network needs this IMS Preferred Codec List before it selects the Target RAN Codec.

The SRVCC MSC can match the Target RAN Codec (one of the MSC Supported Codec List) optimally to the best matching Codec of the IMS Preferred Codec List. Since the SRVCC MSC does not know the IMS Selected Codec, it cannot avoid transcoding immediately. The SRVCC MSC provides the optimal Target RAN Codec under these conditions to the ATCF for fast eSRVCC. The ATCF/ATGW can then remove the Transcoding after the UE landed safely in the Target RAN by a subsequent SIP/SDP Re-Invite, modifying the IMS Selected Codec and the Remote Used Codec to match the new Target RAN Codec.

This scenario is more complex, but it is unavoidable in real life networks.
The voice quality unavoidably goes down to the quality of the new Target RAN Codec in transcoding free operation.

If the IMS Selected Codec is "better" than the Target RAN capabilities, then it is important that the ATCF sends the IMS Preferred Codec List to the eSRVCC MSC.

Then the SRVCC MSC can select the Target RAN Codec optimally, although transcoding is temporarily necessary. The subsequent Re-Negotiation of the IMS Selected Codec may achieve TLCI, because the SRVCC MSC selected the Target RAN Codec for that purpose. Only in cases, where the remote end is not supporting any 3GPP Codec, transcoding is unavoidable.

7.2.3 Assemble the remote IMS Preferred Codec List

7.2.3.1 General

One side problem in this scenario, where the Remote Access Network has better capabilities than the Target RAN, is to assemble the (remote) IMS Preferred Codec List. The Codec Negotiation procedure in the CS-world calls this list the (remote) "Alternative Codec List". The present document differentiates two cases, depending on the call setup direction.

7.2.3.2 Call Setup Scenario 1: from remote to local

Per definition, the local side performs the eSRVCC. The local ATCF got in the initial SIP Invite a List of Codec Candidates from the remote end, the "remote IMS Supported Codec List", stemming from the "remote UE Supported Codec List", filtered by all nodes in the path. The local ATCF (at the terminating side) may filter this list further and send the result as initial SIP Invite Offer to the local, terminating VoLTE UE. This selects finally the local LTE Used Codec. Based on this SIP Response from the local UE the ATCF determines the IMS Selected Codec. The ATCF sends only this IMS Selected Codec to the remote, originating end.

Important is in this scenario 1: the local ATCF may remember all the other Codec candidates from the remote IMS Supported Codec List. Together with the IMS Selected Codec, **the remote IMS Preferred Codec List can be assembled.**

EXAMPLE: Where the local UE does not support EVS, but AMR-WB and AMR.
remote UE Supported Codec List = {EVS-FB-11+EVS-IO(), AMR-WB(), AMR()}
remote IMS Supported Codec List = {EVS-SWB-6+EVS-IO(), AMR-WB(), G.722, AMR(0,2,4,7), G.711}
local initial SIP Invite Offer = {EVS-SWB-6+EVS-IO(), AMR-WB(), AMR(0,2,4,7)}
selected local LTE Used Codec = {AMR-WB()}
IMS Selected Codec = {AMR-WB()}
remote IMS Preferred Codec List = {AMR-WB(), EVS-SWB-6+EVS-IO(), AMR(0,2,4,7), G.711}.

NOTE: It is not likely that the local UE supports EVS in CS, while it does not support EVS in LTE. It is therefore most likely not essential that EVS is offered to the SRVCC MSC. Nevertheless, the complete remote IMS Preferred Codec List in this example contains EVS, although on second place, after the IMS Selected Codec.

7.2.3.3 Call Setup Scenario 2: from local to remote

The local, originating UE sends the initial SIP Invite with its local UE Supported Codec List and the local ATCF filters this according to local policy. The ATCF sends this further as "Local IMS Supported Codec List" to the remote end. The SIP Response from that remote end contains, however, only the IMS Selected Codec. The remote ATCF does not even report the Remote Used Codec. In contrast to Codec Negotiation in the CS Networks, the IMS Offer-Answer procedure returns only one Codec, not the Alternative Codec List in addition.

From this SIP Response, the remote IMS Preferred Codec List would contain only one entry, but not the whole list, in contrast to the call setup from remote to local. The local ATCF may undertake some "intelligent guessing", but in principle some important information is missing.

EXAMPLE 1: Where the remote UE does not support EVS, but AMR-WB and AMR.
local UE Supported Codec List = {EVS-FB-11+EVS-IO-8, AMR-WB(), AMR()}
local IMS Supported Codec List = {EVS-SWB-6+EVS-IO-8, AMR-WB(), G.722, AMR(0,2,4,7), G.711}
remote initial SIP Invite Offer = {EVS-SWB-6+EVS-IO-8, AMR-WB(), AMR(0,2,4,7)}
selected remote Used Codec = {AMR-WB()}
IMS Selected Codec = {AMR-WB()}
remote IMS Preferred Codec List = {AMR-WB(), AMR(0,2,4,7), G.711} - by guessing.
This result in this example is a good guessing, but this guessing may not be complete and correct in all cases.

EXAMPLE 2: Where the remote UE supports more than the local UE: EVS, AMR-WB and AMR.
local UE Supported Codec List = {AMR-WB(), AMR()}
local IMS Supported Codec List = {AMR-WB(), G.722, AMR(0,2,4,7), G.711}
remote initial SIP Invite Offer = {AMR-WB(), AMR(0,2,4,7)}
selected remote Used Codec = {AMR-WB()}
IMS Selected Codec = {AMR-WB()}
local Used LTE Codec = {AMR-WB()}
remote IMS Preferred Codec List = {AMR-WB(), AMR(0,2,4,7), G.711}.
This result in this example is not complete, but maybe good enough. The full remote IMS Preferred Codec List could be {AMR-WB(), EVS-FB-11+EVS-IO-8, G.722, AMR(0,2,4,7), G.711}.
The local ATCF cannot assemble the remote IMS Preferred Codec List correctly in all cases. This is a result of the Codec Negotiation rules in IMS, which mandates to return only the IMS Selected Codec in SIP Response, without alternative candidates.

7.3 Late Information about the Target RAN Codec

According to figure 5.1-2, Reference Procedure for eSRVCC, the SRVCC MSC informs the ATCF in message 10a, "SIP Invite (MSC Preferred Codec List)" about the Selected Target RAN Codec (first Codec in the list) and some alternative Codec candidates for the CS-PS-Codec. At that moment, the target radio leg is already setup and not changeable; in addition some noticeable time has passed since eNB as taken the decision for the eSRVCC. The ATCF has no now other alternative than to accept one of the offered Codecs from the MSC Preferred Codec List.

If necessary - in a noticeable number of cases - the MGWs insert transcoding, in either the Target MGW, or the ATGW, or both. The ATCF informs the ATGW about that decision in message 10b, Session Transfer (CS-PS-Codec), see figure 5.1-2.

Message 10b immediately also starts the session transfer, stopping the communication with the local LTE leg and starts the communication with the local CS leg. That is at least Stage 2 procedure and real life networks show this.

Even if transcoding is not necessary, the IMS Selected Codec has often a wider range of capabilities, than the Target RAN Codec in terms of bit rate or audio bandwidth; or the IMS Selected Codec is operating in another, non-compatible mode of operation, than supported by the Target RAN Codec. One example is the eSRVCC from EVS-FB-11 to AMR-WB-2. Although the call can continue after eSRVCC without transcoding, the transition is cumbersome. The Target RAN Codec cannot understand the speech data coming from the Remote Used Codec in that moment, immediately after session transfer, as long as the Remote Used Codec received no CMR command to use EVS-IO-2.

In one alternative approach the ATGW may insert a pair of Transcoders (e.g. AMR-WB-2 <=/=> EVS-FB-11) for a certain transient time to keep the speech break during eSRVCC small. After the successful execution of the eSRVCC and the successful Mode Control of the Remote Used Codec, the ATGW removes this pair of Transcoders again. Inserting into and removing transcoders from a speech path is expensive, complex to handle and in any case, it causes speech path distortions and jumps in the speech path delay. Both effects, distortions and delay jumps, are clearly measurable by objective tools and are of course often audible.

Inserting transcoding for a short while and removing it later is expensive and is degrading the voice quality.

In another alternative approach the ATGW may immediately started the Mode Control of the Remote Used Codec, as soon as the ATCF informs the ATGW (message 10b). Due to the unavoidable round trip delay, from the ATGW to the remote UE and back, the speech break during eSRVCC can still be substantial, far beyond the target of 300 ms. Also this is clearly audible and measurable.

Starting Rate- and Band-Control too late is degrading the voice quality during eSRVCC.

In order to achieve an optimal solution the ATCF would have to inform the ATGW a while **before** the session transfer, to trigger the Pre-SRVCC Mode Control. It would not matter, if the remote end would send in the reduced Codec Mode already before the handover interrupts the link to the local LTE leg, because the local VoLTE UE can receive these speech data frames as well. The example is here again: Remote Used Codec is EVS-FB-11, the ATGW sends CMR-IO-2 in due time and the Remote Used Codec falls back to EVS-IO-2, before the local eSRVCC-handover to AMR-WB-2 happens.

Pre-SRVCC Mode Control is necessary for the optimal eSRVCC.

7.4 Access Transfer and Handover Command

Figure 5.1-2 shows the Stage 2 procedure, where message 10a, SIP Invite (MSC Preferred Codec List 2), is sent to the ATCF at the same time as message 13, PS to CS Response (Target RAN Codec). The idea behind that was to synchronize the Access Transfer in ATGW with the handover on air. This idea is, however, not realistic for several reasons.

In "sunshine" situations, where the network links and network nodes are not loaded with traffic and the radio interface is excellent, without delay and transmission errors, the timing may be trim-able, such that the handover on air and the access transfer in the ATGW (HO in ATGW) occur at roughly the same time.

Real life networks, however, have to work also well under realistic, partly high load situations.

Case 1: Maybe the messages between MSC and ATCF/ATGW are delayed, queued or otherwise the execution may be shifted in time. Sometimes (e.g. in eSRVCC during setup) the necessary resources are not available. The ATCF delays then the handover in the ATGW. The ATGW still maintains the link to the local LTE leg after the local UE has left the LTE access, because the Handover Command was faster. The speech break is longer than wanted.

Case 2: Maybe the handover command is delayed, e.g. because the LTE leg is already disturbed (eSRVCC is necessary, because the LTE leg is weak) and the Handover Command is repeated one or several times. Then the ATGW has already broken the LTE leg and uses the CS access leg, although the UE is still LTE connected. This causes a longer speech break, too.

This legacy procedure design for eSRVCC is not fail save and falls short compared to legacy handover handling in CS networks:

A legacy MGW starts "bi-casting" the speech data, coming from the remote end, downlink to both, the old and the new access leg. This guarantees that the speech interruption in downlink is minimal, independent of the timing of the handover on air.

Similarly, the legacy MGW starts listening to all speech data coming in uplink from both, the old and the new access leg. The MGW forwards valid speech data to the remote end, regardless on which access the MGW received them. This "intelligent combining" in uplink guarantees, that the speech interruption in uplink is minimal.

Prerequisite for minimal speech path interruption during eSRVCC is a successful bi-casting in downlink and intelligent combining in uplink. This may be ensured only, if the Handover Command is triggered after the MGW resources are successfully allocated.

This handover handling within and by the ATGW stops after the local UE performed the handover successfully. Another advantage of this legacy handover handling is that the old radio leg is still active in the MGW in case the handover fails.

7.5 Target MGW is blocked in Uplink

According to the eSRVCC standard, the uplink path in the Target MGW is blocked (is set to one-way, downlink-only), until the MSC has received a "Handover Complete" message from the UE via the new Target RAN leg. Then the MSC commands the Target MGW to pass speech frames in uplink. They arrive at the ATGW, which forwards them to the remote end, maybe after repacking or even transcoding. The uplink speech break ends, when these speech frames finally arrive at the remote end.

This control (blocking) of the Target MGW is unusual and not necessary. It blocks the uplink speech path in the Target RAN too long and causes an unnecessary uplink interruption. The target base stations have strong error detection mechanisms, allowing differentiating good speech frames in uplink from garbage quite well. These base stations send only valid speech frames uplink and the Target MGW should let them pass immediately. The "Handover Complete" message from the UE is just the confirmation that the handover was successful. After that, the network may shut down the old radio leg safely.

7.6 The remote UE does not follow CMR commands

Lab- and field-tests showed that some remote UE did not follow the Codec Mode Requests at all. In this case muting on the local UE was unavoidable after eSRVCC, if the network did not insert transcoding. There are currently no means to detect such a faulty remote UE.

In other cases, some remote UE did follow the CMR, but only after e.g. three repeated CMRs. This caused an additional delay to the round trip time of at least 60ms. This is measurable; often not easy to detect during active speech at the side of the local UE, because in that case the ATGW sends the new CMR in consecutive RTP packets of 20 ms distance. This unusual behaviour of such a remote UE gets problematic, in case the local UE sends only SID frames, when it detected a local speech pause: then three consecutive new CMR take at least 320 ms more than needed. The speech break is then very long in local downlink.

Meanwhile 3GPP TS 26.114 [5] clarifies in REL-12 for AMR and AMR-WB: every MTSI client has to follow each received CMR as soon as possible and so the problem will - hopefully - not appear in new terminals. The same clarification is necessary for EVS.

This CMR problem is not only an eSRVCC problem; it is a serious misbehaviour in many situations.

8 Speech Quality and Media Handling Aspects

8.1 General

This clause discusses the Speech Quality and Media Handling Aspects of the current 3GPP standard procedures for eSRVCC, based on the discussion of the example eSRVCC scenarios in clause 6. It shows the deficits and reasons.

8.2 Blind Selection of the Target RAN Codec

As explained in clause 6 the SRVCC MSC has to select the Target RAN Codec without sufficient knowledge about the ongoing call and therefore in many scenarios SRVCC MSC and/or ATCF insert transcoding, although TLCI would be possible.

Unnecessary transcoding does not only waste MGW resources, either in the CS-MGW or in the ATGW (in worst case in both), but increases also the speech path delay, with negative influence on the overall user perception of the communication.

The additional intrinsic voice quality distortion is the most important negative influence, caused by this transcoding.

8.3 Unnecessary speech break by missing Rate Control

Even in scenarios, where the Target RAN Codec is TLCI-compatible to the IMS Selected Codec, the speech break during eSRVCC may be longer than necessary due to high Codec Modes on the IMS side, which the CS-Side cannot handle.

Examples are eSRVCC from AMR (0,2,4,7) to HR_ AMR (0,2,4); or eSRVCC from AMR-WB () to UMTS_AMR-WB (0,1,2); or eSRVCC from EVS (br=5.9-64; bw=nb-fb) to UMTS_EVS (Set 2) or eSRVCC from EVS (br=9.6-24.4;bw=swb) to UMTS_EVS (Set 3).

Although all these Codec pairs are TLCI-compatible, the CS-side receives for a short while (round trip time) too high Codec Modes, until the Maximum Mode Control with CMR has brought the Codec Modes used in the remote end into the common Configuration. During this time, the CS-side mutes the loudspeaker during active speech segments, while the IMS-side does not perceive a problem. Even worse, the CS-side handles received SID frames as usual and generates Comfort Noise in speech pauses, while muting occurs in active speech parts.UEs, which do not follow the Codec Mode Requests, or not fast enough, intensify this problem.

Speech muting is obviously the worst thinkable effect, especially if only one side perceives it, while the other side experiences undisturbed reception.

8.4 Unsynchronized, early Handover switching by ATGW

Figure 8.4-1 shows the relationship of speech signals travelling in various segments of the speech path before, during and after eSRVCC for the case, where the ATGW switches the speech path earlier than the UE changes the radio access.

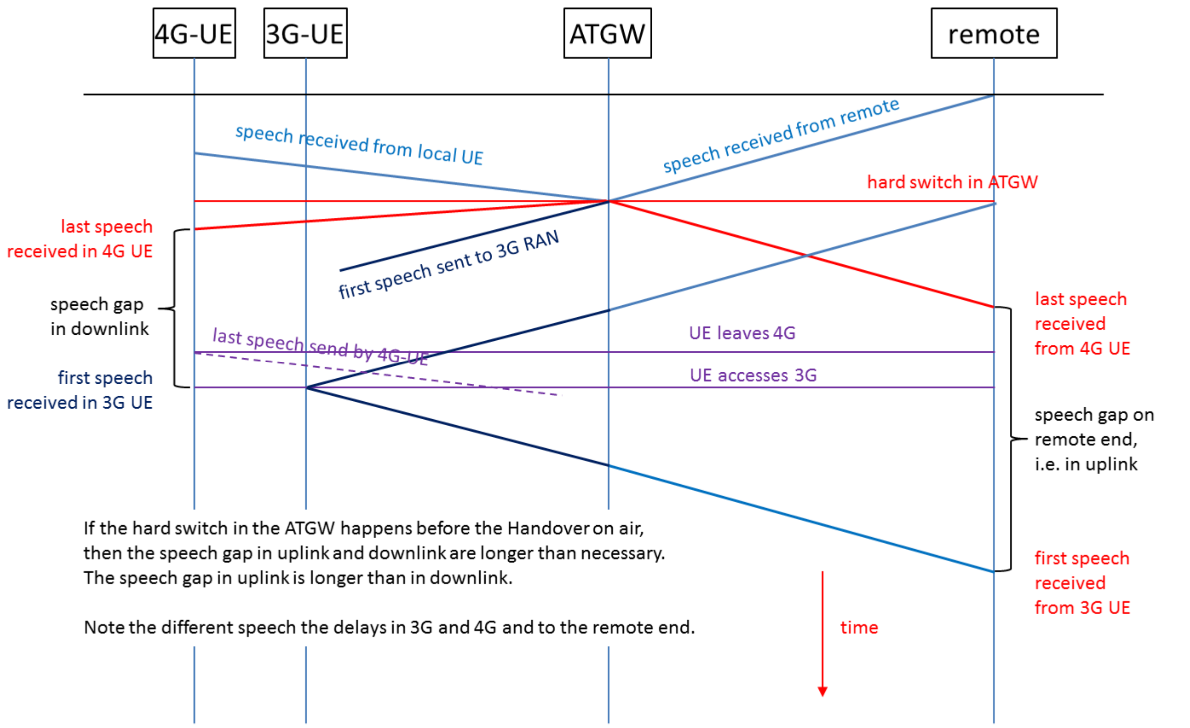

Figure 8.4-1: Timing relations for eSRVCC with hard switching in the ATGW before the UE changes the RAT

In figure 8.4-1, time is running downwards. Every speech path has an unavoidable limited transport speed and therefore a speech path delay. The higher the speech path delay, the steeper the lines in this timing diagram.

Speech sent by the local 4G-UE travels for a while, until it reaches the ATGW. In this assumed example, the 4G cell is unloaded and the delay is comparably small in uplink and downlink, smaller than the corresponding delay in 3G. The path to and from the remote is "long" in terms of speech path delay, so it takes a while to receive from or send the remote end. This longer speech path delay with respect to the remote end is not immediately perceivable. Only in case of an active communication (Question - Answer), or in case of a Codec Mode Request from one side and the reaction to it back to this side, this speech path delay is observable (round trip delay).

At a certain point in time the ATGW gets the command from the ATCF (not shown) to switch from 4G to 3G, in uplink and in downlink. Some speech packets are still travelling downlink and reach the 4G-UE, before it mutes its output. The next frames after switching are travelling to the 3G RAN (and maybe onto air), but the UE is still in 4G and does not get these first frames.

The hard switching in the ATGW cuts the uplink path from 4G-UE to ATGW sharp, packets in this uplink pipe are lost, as well as several following packet, which the 4G-UE sends until the UE leaves the 4G access. After a while, the remote side notices this sharp break and goes muting.

Some time span after the ATGW performed the switching the Handover Command reaches the 4G-UE and the UE leave 4G access and connects to the 3G access: it becomes a 3G-UE.

Because the 3G RAN receives downlink speech since some time from the ATGW the 3G-UE may quickly start decoding and unmuting its output. The speech break in downlink ends.

The first frames from the 3G-UE in uplink need to travel the uplink pipe, until they reach the ATGW, which then forwards them to the remote side. Then the uplink speech break ends.

The example in figure 8.4-1 does not show the Target MGW. It considers the Target MGW as through-connected both-ways, not blocking the uplink path. In real networks, following the current eSRVCC stage 2 specification this Target MGW blocks, however the speech path and by that increases the uplink speech break even more.

In general, the speech break in uplink is noticeably longer than the break in downlink. Both are far longer than the time span in which the UE "disappears" on air.

The more the system or UE delay the handover on air with respect to the hard switching by the ATGW, the longer both speech breaks are. This fact leads to the attempt to "synchronize" both events by sending the Handover command earlier to the 4G-UE, before the ATCF/ATGW could perform the switching.

8.5 Synchronized hard Handover

Figure 8.5-1 illustrates the (theoretical) example, where the hard switching within the ATGW exactly synchronized to the handover on air. The speech breaks in uplink and downlink are smaller than in the (more realistic) example before, but they are still not as small as they could be.

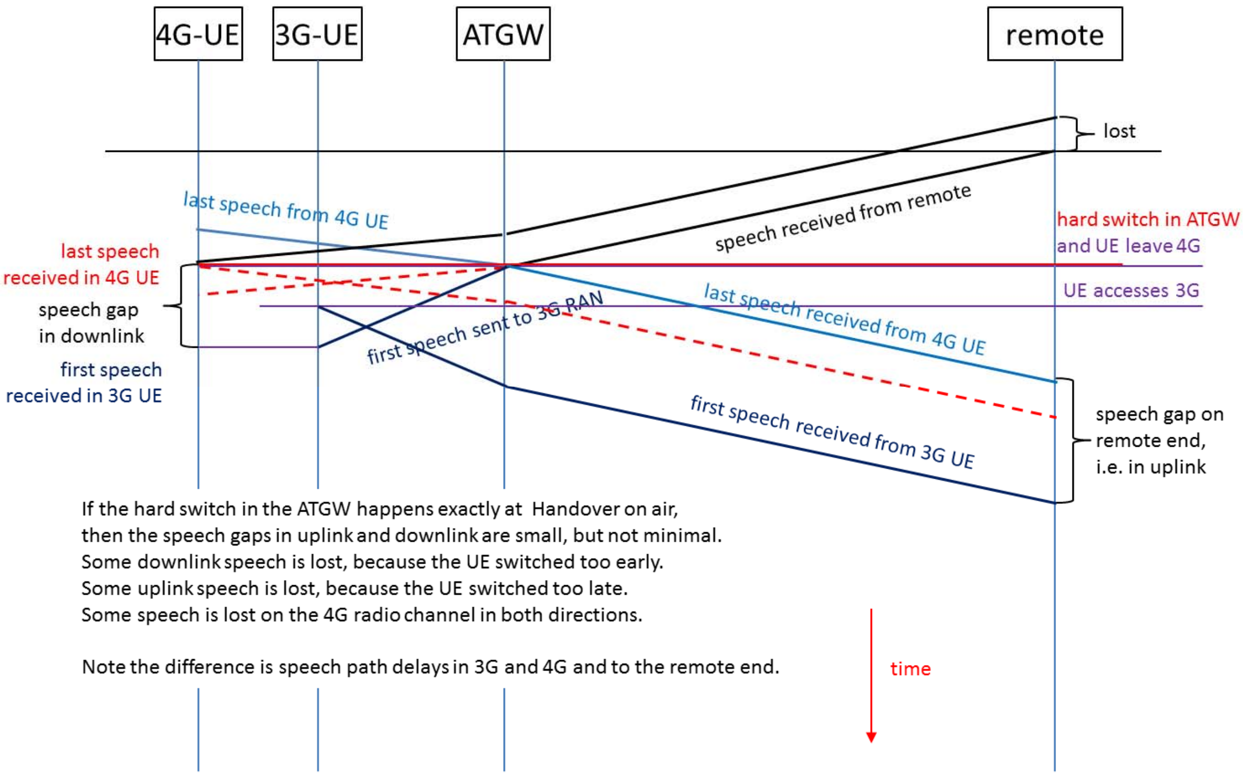

Figure 8.5-1: Timing relations for eSRVCC with synchronized hard switching

Some last speech packets from 4G-UE are lost in the pipe, because the ATGW ignores them. Some speech packets in downlink are lost, because the 4U-UE does no longer listen. In addition, the speech break in downlink is increased by the longer speech path delay in the 3G access (at least in this example, where the 4G cell is not loaded).

In uplink the ATGW ignores some speech frames from the 4G-UE, which are still in the uplink pipe; the longer speech path delay in 3G increases the uplink break, too.

The dominant disadvantage of this approach: the eSRVCC MSC sends the Handover Command before the ATCF reports the successful allocation of resources in the ATGW. This leads in some situations to handover failure and call break.

The third approach, described in the following clause avoids that too early (unconfirmed) sending of the Handover Command and minimizes the speech break in both directions, regardless when the ATGW or UE execute the switching: synchronization between both events is not a prerequisite.

8.6 Ideal eSRVCC Handover

As described in the previous clause, it is essential for a save eSRVCC, that the SRVCC MSC waits, until the ATCF responds **positively** to the SIP Invite, indicating that all resources are available. At this point in time, the ATGW has already started to handle the eSRVCC handover, similar to the case in clause 8.4.

If, however, mainly at call setup and in the Pre-Alerting State, the ATCF response negatively, indicating that the resources are (still) not available, then the SRVCC MSC does not send the Handover Command, but either waits for a (short) while or rejects the PS-to-CS Handover Request. The call continues in 4G access, until the resources are available and the 4G access request the eSRVCC handover again.

The Handover Command reaches in that approach unavoidably the 4G-UE some span after the ATGW started the handover handling. However, different to the approach in clause 8.4 the ATGW continues to send speech in downlink to the 4G radio access. The speech break in downlink starts exactly then, when the UE leaves the 4G access. Meanwhile the ATGW send speech already also to the 3G access and speech is "on air" in both, 4G and 3G simultaneously for a short while. **The present document calls this "bi-casting"**.

Similar in uplink: the ATGW continues to listen to the 4G uplink path and forwards all speech packets as they arrive to the remote side. Simultaneously the ATGW starts to listen also to the 3G access. When the UE leaves the 4G access, the ATGW gets for a short while (uplink pipe) the last speech packets, before this 4G uplink stream stops. Later the ATGW receives then speech frames from the 3G access. The uplink break is as short as can be and only determined by the time, the UE "disappears" on air and the time-difference between 3G-uplink-delay and 4G-uplink-delay. The present document calls this handling "**intelligent-combining**".

In the good case (majority), the eSRVCC Handover is successful. Then the ATGW does not get speech from 4G after 3G and it never gets speech from both uplink channels. In the bad case, when the UE cannot access the 3G radio, the UE falls back to the 4G access, and stays there. In both cases, the ATGW may autonomously, or on command from the ATCF, stop bi-casting and intelligent-combining after a while and return to its normal operation.

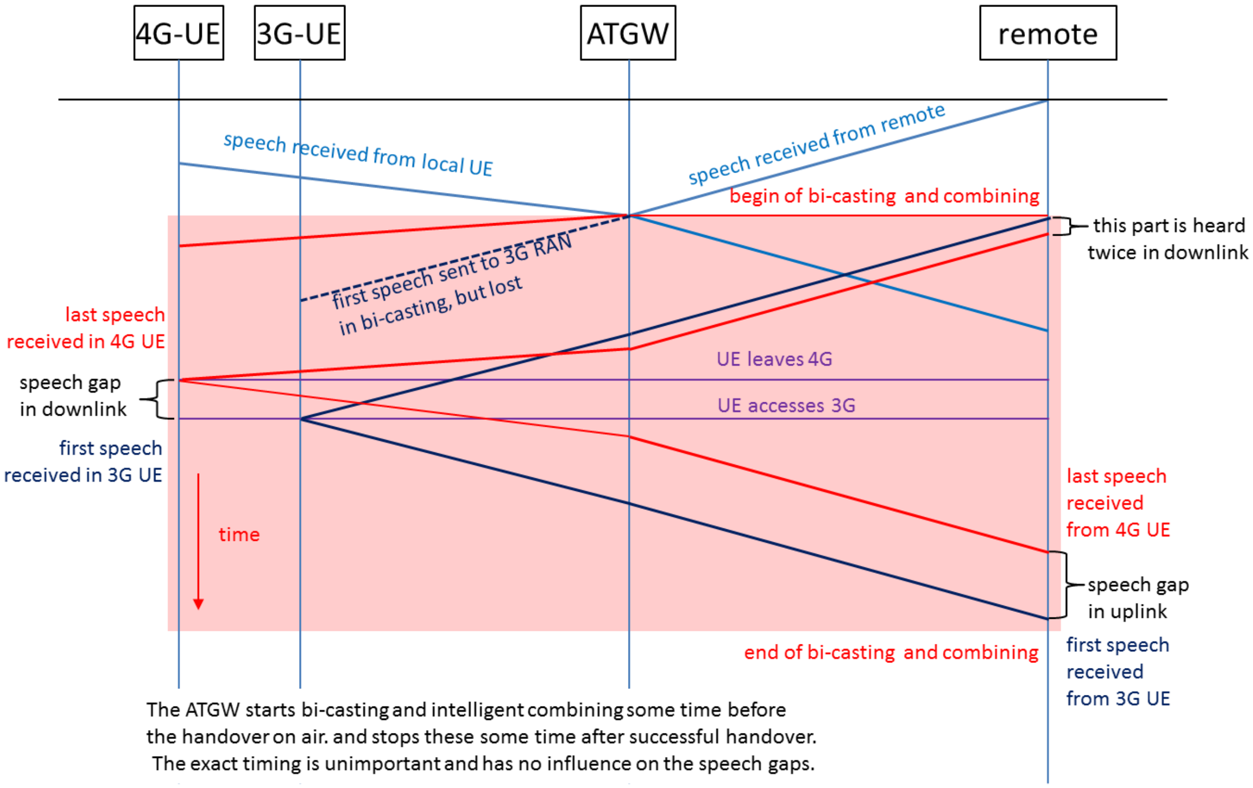

Figure 8.6-1: Ideal eSRVCC handover with bi-casting and intelligent combining

This handling decouples the ATGW operations totally from the handover timing on air. This handling is extremely robust in all kinds of load situations or radio conditions. The perceived speech gap times depend now mainly on the time span the UE "disappears" on both radio accesses.

An interesting artefact: it may happen (theoretically) that the local UE receives and decodes some short part of the remote speech twice, because the 4G-downlink has a substantially shorter speech path delay compared to the 3G-downlink. Some speech frames on 4G-DL "bypass" other speech frames on 3G-DL.

The audible / measurable speech gap in downlink is now only dependent on the implementation skills in the UE, i.e. on the time the UE is not receiving on either access (unfortunately some UEs exist with quite bad performance).

The uplink path has a somewhat longer gap than the downlink path, because here the "disappearing time" and the difference between 3G-UL-delay and 4G-UL-delay add up. Please note that the shown example represents a situation with rather small load in 4G. With higher load, or a marginal uplink radio performance at the edge of the 4G cell, the 4G-UL-delay increases and the Uplink gap gets shorter.

Summary Conclusion: bi-casting in downlink and intelligent-combining in uplink minimize both speech gaps and provide an extremely robust handling in real life networks. This handling in the ATGW allows the SRVCC MSC to wait for the response from the ATCF, indicating that the resources are available in the ATGW and that the ATGW has started this handover handling.

Final note: This presentation here is a bit simplifying, ignoring the jitter buffers in downlink and uplink. It may well be that an optimal implementation in the UE brings the downlink speech gap sometimes close to zero. This is possible in cases of high downlink jitter, because the 4G access may sometimes fill the jitter buffer with frames before the UE changes the access and the UE can decode for a while from this filled jitter buffer. It is, however, not reliable and not predictable what exactly happens in a specific event.

Important is still to note that a good implementation in the UE does not reset the Speech Encoder and Decoder in these many cases, where the LTE Used Codec is identical to the Target RAN Codec. In these cases, eSRVCC may be nearly seamless and inaudible.

9 Codec Mode Control before, during and after SRVCC

9.1 General

The AMR Mode Control procedure was originally designed for Mobile<=>PSTN calls and extended later to cover also Mobile<=>Mobile calls in Transcoding Less Operation (TFO or TrFO).

In the first case, Mobile<=>PSTN, there is typically only one major bottleneck in the voice path: the radio interface, which varies over time and location and requires adaptation of the media (net) bit rate to the channel conditions. These channel conditions may be temporary, as the radio signal strength or the radio interference fluctuates. These channel conditions may also be permanent or semi-permanent, e.g. if GERAN needs handover to a half-rate traffic channel to gain call capacity, or if UTRAN needs handover to a spreading factor SF=256 for the same reason.

In the second case, Mobile<=>Mobile, there are more bottlenecks in the voice path: both radio interfaces may vary over time and location temporarily or semi-permanently. In general there could be even more bottlenecks in the voice path end-to-end, like an overloaded Abis-interface in GERAN, or a satellite link or microwave-links somewhere.

The AMR Mode Control signalling and procedure was designed to cope with multiple bottlenecks in the voice path.

Figure 9.1-1 shows one example of a Mobile<=>Mobile call with an assumed bottleneck in the Core Network (on NboIP).

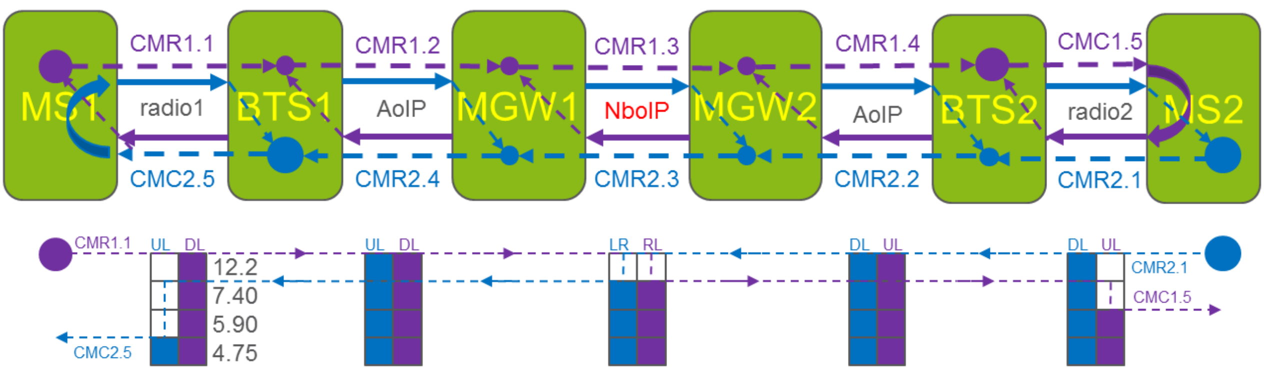

Figure 9.1-1: Mobile<=>Mobile call with two radio interfaces and an assumed bottleneck in the Core Network

The voice media traffic flow is bi-directional, represented in **blue coloured think lines** for the direction "left to right" (i.e. MS1 => MS2) and in **violet coloured think lines** for the direction "right to left" (i.e. MS1 <= MS2). The speech media is transported transcoding-free end-to-end: speech is encoded once in the media-sending mobile (e.g. MS1) and decoded once in the media-receiving mobile (e.g. MS2). This guarantees highest possible voice quality under all given radio conditions, assuming, that the Mode Control feedback keeps the Codec Mode in the optimal range.

The Mode Control signals in form of Codec Mode Requests (CMR) are always sent in the opposite direction (feedback), relative to the media stream. The CMR flow is represented with dashed lines of the same colour of the media stream it controls. The blue and violet columns below the block diagram represent the selected mode-set AMR (0,2,4,7) and the local and temporal rate restrictions, one column for each interface and direction. For example, uplink radio interface 1 has an extreme low maximum rate of only 4.75 kbps, i.e. only the lowest mode is allowed. The Codec Mode Command (CMC2.5) is set by BTS1 to CMC2.5=0. A node receiving media on a specific interface (e.g. MGW2 receiving data from MGW1 via NboIP) estimates the receive-link quality and influences the CMR in the opposite direction accordingly (e.g. CMR2.3).

Prerequisite for end-to-end TLCI is that the media-encoder knows the smallest bottleneck in the media path!

Each media-receiver and media-decoder, e.g. MS1, observes its downlink radio conditions and estimates the maximum mode suitable for these radio conditions. This estimated maximum mode is send backwards, e.g. as CMR1.1. In the case above, radio 1 has no problem in downlink. So CMR1.1=7, i.e. the highest mode with rate=12,2 kbps could be used on this local radio 1 downlink.

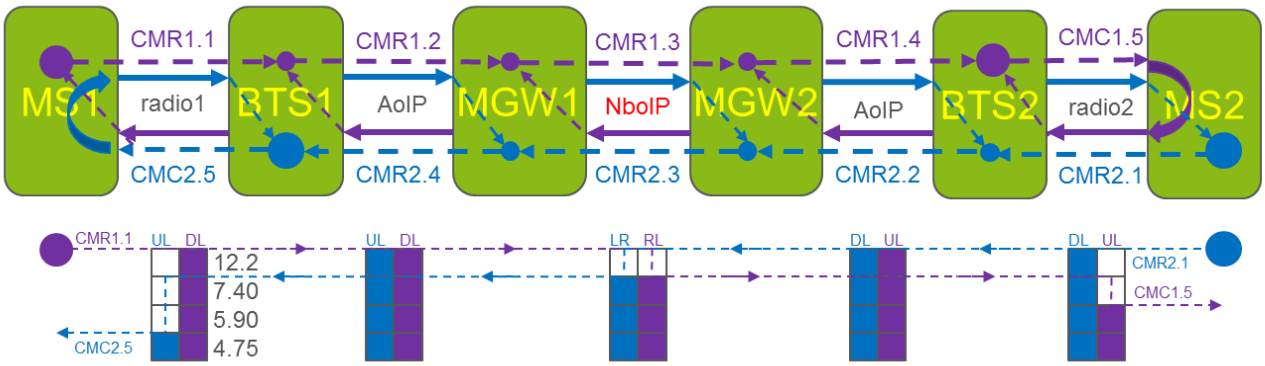

Figure 9.1-2: Reprint of figure 9.1-1: Mobile<=>Mobile call with two radio interfaces …

BTS1 may modify this CMR1.1, e.g. based on load on the incoming AoIP interface, and then send CMR1.2 forward towards the Core Network, i.e. MGW1.

CMR1.2 = MIN (CMR1.1, <local max mode on AoIP1 in downlink>) is the corresponding formula.

In the example here CMR1.2 == CMR1.1 = 7: there is no bottleneck on AoIP1 in downlink.

MGW1 observes the incoming NboIP-link and detects a restriction to rate 7.40, i.e. mode=4. So MGW1 sets CMR1.3 = MIN (CMR1.2, 4) = 4 and forwards (backwards relative to the media-stream) CMR1.3 to MGW2.

The smallest bottleneck in this media flow is in this example in the uplink of radio 2. BTS2 observes this radio link 2 and estimated the maximum mode to 2, i.e. rate 5,90 kbps.
Therefore, finally CMC1.5 = MIN(CMR1.4, 2) = 2 is sent downlink to MS2.

It is mandatory for MS2 to obey this "**C**odec **M**ode **C**ommand" as maximum allowed mode in uplink as soon as possible. CMC1.5 is the minimum of all estimated maximum modes of all bottlenecks, calculated in a distributed manner.

Regardless, where the smallest bottleneck will be: the Distributed Rate Decision always finds it!

Exactly the same procedure, with typically different result, is executed for the opposite media-direction. The Mode Control loop delay is dependent on the position of the bottleneck in the speech path. This control loop delay is always as small as it can be.

9.2 Mode Control commands in the User Plane

The speech path delay is an important factor for a good communication quality for humans. The smaller the speech path delay, the more natural the communication; a long delay causes irritations to the participants. The speech path is optimized. It does not follow the same route as the control signalling and does not pass the same nodes. Therefore, the speech path (User Plane) has typically a (much) lower transport delay than the Control Plane. This is one important reason, why the Mode Control Commands (CMR) are transported in the User Plane.

Another aspect is the tight synchronization between media payload and Codec Mode Request. This allows a fast and timely response in case some bottleneck changes and needs a fast adaptation. The Control Plane could not support this fast reaction.

In GERAN, every SID frame and every second Speech frame (every 40 ms) transports the active CMR, endlessly repeated, even if it does not change. **This endless repetition has to be seen as extreme robust forward error correction code** and allows a fast error recovery. A single lost or disturbed CMR values is quickly healed by the next one. There is no need for an acknowledgement for CMR. CMR is slim signalling.

On AoIP, NboIP and in IMS the AMR and AMR-WB payload is transported in RTP packets, each containing a field for CMR. This CMR field is always present.

3GPP TS 26.114 [5] REL-12 clarifies that the active CMR are to be sent in every RTP packet for AMR and AMR-WB.

9.3 Mode Control Rules for AMR and AMR-WB

Any implementation of AMR or AMR-WB in an MTSI Client has to obey these Mode Control Rules, otherwise end-to-end TLCI is impossible. It is important that IP end-points, maybe not following 3GPP TS 26.114 [5] in all points, do follow the AMR Mode Control Rules, if they offer AMR or AMR-WB in SIP/SDP.

Especially important is that **every media-sender does obey the received CMR as the maximum mode** it is allowed to use for media-encoding. This is true, even if the media-sender itself does not see any restriction in its local access side. None of the involved clients or servers overlooks the total media path. Only the Mode Control feedback provides the overview, how big the smallest bottleneck is.

In a general voice session, it is typically unknown to one end what the other end’s access is and it is any time possible that the conditions on one or the other end change. **It is therefore important that every media-sender follows the received CMR as fast as possible, e.g. within about 40 ms.**

An important example is an handover on the far end, e.g. a GERAN-internal handover from the full-rate channel, AMR (0,2,4,7) to the half-rate channel, AMR (0,2,4). Immediately after the handover (in some implementations already some time BEFORE the handover) the CS network sends CMR=4 and below. If a VoLTE client on the remote end would not obey these CMR-values and continue with mode 7, because it does not see any problem on its local LTE access, then the output on the GERAN terminal will go to muting: mode 7 cannot be transported downlink on a GERAN half-rate channel.

Another important example is the eSRVCC from a VoLTE<=>VoLTE call with AMR-WB () to UTRAN or GERAN. The maximum mode for AMR-WB in UTRAN is AMR-WB (2) or AMR-WB (4) or AMR-WB (8), depending on operator policy and in GERAN is AMR-WB (2). The Target RAN sends Mode Control Commands through the CS-Core - during eSRVCC or after eSRVCC is finished - and they will be received in the ATGW within the RTP packets as CMR=2 (or lower) and then forwarded to the remote VoLTE UE. Important is that the Target MGW or the ATGW obeys the potential difference in AMR-WB configurations and maps the CMR into the common mode-set (potentially needed in case of UTRAN).

Again, it is indispensable that **the remote VoLTE UE does obey these CMR-values as maximum mode for media-encoding; otherwise, the UE on UTRAN or GERAN side goes to muting**.

9.4 Mode Control Rules for EVS

EVS is a new 3GPP Codec with substantially enlarged adaptation capabilities. The principle of "Codec Mode Control" remains the same. The smallest bottleneck in the total voice path end-to-end determines the maximum mode that can be used without transcoding. **Transcoding always brings lower quality.**

The EVS Codec Mode Request (EVS-CMR) comprises commands to restrict the maximum rate, but also for maximum audio bandwidth. In addition, EVS-CMR is used to control the "Variable Bit Rate" mode of EVS (on/off, nb, wb) and the "Channel Aware" mode of EVS (on/off, wb, swb, various options). EVS-CMR controls also the EVS-IO mode of operation and the transitions between the EVS modes of operation.

NOTE: Further clarifications in normative specifications regarding CMR as trigger for transitions between the EVS modes of operation may be required. For transitioning from EVS primary mode to EVS-IO mode due to eSRVCC, further adjustements in standards may become necessary, for instance regarding the EVS AMR-WB IO mode-set, the mode-change-period and the mode-change-neighbor used in the IMS network.

The RTP payload format for EVS is specified in TS 26.445 with several options for EVS-CMR transport. It is allowed to omit the EVS-CMR in RTP. It is allowed to send the EVS-CMR on demand, i.e. only when found necessary. It is allowed to send EVS-CMR in RTCP-APP. It is allowed to send EVS-CMR in every RTP packet: this is the safest option.

For many good reasons it is recommendable to send the active EVS-CMR in every RTP packet, in Speech, SID and CMR-Only packets. Only this permanent repetition allows the fastest possible adaptation, with high error robustness. The Distributed Mode Decision is simplest, if every RTP packet for EVS includes the active EVS-CMR. These considerations are the same as for AMR and AMR-WB.

One important aspect is the TLCI-compatibility between EVS and AMR-WB. EVS includes the EVS AMR-WB IO mode of operation, in short EVS-IO. The EVS-CMR controls also the transition between EVS Primary modes and the EVS-IO modes, together with the maximum bit rate in EVS-IO. Because AMR-WB mandates an active CMR in every RTP Packet, this requirement is passed to the EVS-IO as well.

Same as for AMR, the simplest approach would be starting the EVS-CMR feedback signalling by the media-receiving EVS-client (decoder) and send this EVS-CMR, potentially filtered by the network(s) and potentially modified to a lower maximum rate and/or bandwidth, all the way back to the media-sending EVS-client. Nodes in the path can realize the Distributed Rate Decision fastest and easiest. The network can react to sudden disturbances in the media path, like eSRVCC or handover, in the fastest possible way. Lost or disturbed CMR Commands are corrected with the next received RTP packet.

All other options, a) to c) in the list below, for EVS-CMR transport have disadvantages:

a) It is allowed to omit the EVS-CMR in RTP.
Then either a single-rate / single mode Configuration will be used, or CMR will be sent via RTCP-APP.
In principle, SIP/SDP could be envisaged to change the Codec Mode. This is, however, expensive and too slow.

b) It is allowed to send the EVS-CMR on demand, i.e. only when found necessary.
Lost frames mean a lost CMR. It is often not trivial to detect such a case.
This is already discussed in length for the AMR Rate Control.

c) It is allowed to send EVS-CMR in RTCP-APP.
RTCP-App brings irregular overhead and may interfere on the transport plane with the speech data stream.

d) It is allowed, even mandated in this option, to send EVS-CMR in every RTP packet, in speech pauses even in some extra added CMR-only packets, if an urgent CMR has to be sent.
This is the safest option, as discussed for AMR and EVS above. It is, however, only effective, when in each RTP packet the active CMR is sent, "endless" repeated.

9.5 Call Setup and Initial Codec Mode

Mode Control before, during and after eSRVCC is discussed in the following in examples. The principles hold for all Codecs and call scenarios in modified form, also for PS<=>PS calls. Figure 9.5-1 shows one of many call scenarios, where Mode Control is important.

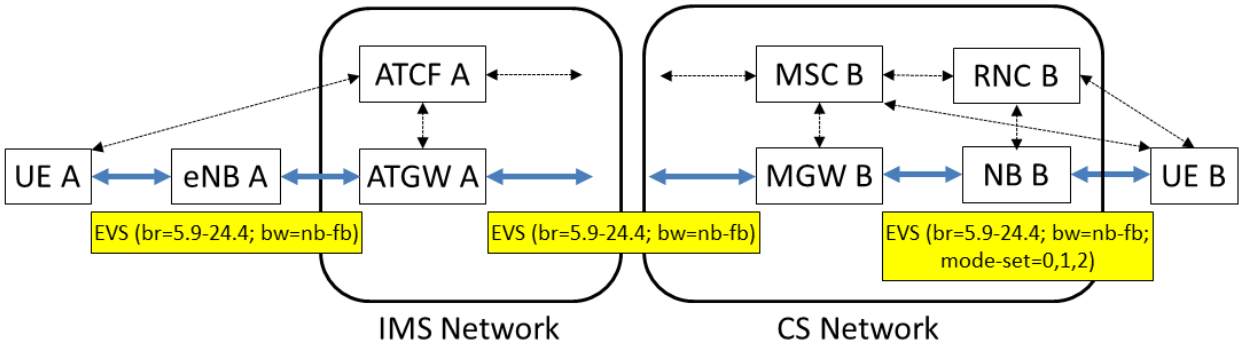

Figure 9.5-1: Mobile<=>Mobile call between 4G and 3G accesses with EVS

This example uses EVS Bottom up Configurations transcoding free all the way between the LTE-UE A and 3G-UE B.

The **LTE Used Codec** (UE A <=> ATGW A) is EVS (br=5.9-24.4; bw=nb-fb), all modes of EVS-IO() included.
The **IMS Selected Codec** (ATGW A <=> MGW B) is EVS (br=5.9-24.4; bw=nb-fb), i.e. the same.
The **UTRAN-Used-Codec (MGW B <=> UE B)** is EVS (br=5.9-24.4; bw=nb-fb; mode-set=0,1,2), i.e. UMTS_EVS (Set 2).

The **EVS Initial Codec Mode** (EVS-ICM) could (in theory) be negotiated and set to EVS (br=24.4; bw=swb) in both directions.

However, this EVS-ICM is not negotiated (according to the current standard), but set by implicit rules. One important input parameter is the smallest EVS Configuration in the path. This, however, is not always known by the endpoints. Other parameters should be the supported audio IO bandwidths in both UEs. The network operator(s) should have influence on the EVS-ICM.

3GPP TS 26.114 [5] defines some implicit rules for the EVS-ICM, these may need review, because they seem to cover not all call scenarios, especially not for the UMTS_EVS.

In this example scenario in figure 9.5.1, RNC B may restrict at call setup the maximum rates in both directions to 13,2 kbps, i.e. the active EVS Configuration would be EVS (br=5.9-13.2; bw=nb-swb; mode-set=0,1,2). This restriction by the RNC would follow the current practise for AMR. The EVS-ICM within UE A should in that case not be higher than EVS (br=13.2; bw=swb), otherwise UE B would perceive muting, until the EVS-CMR signalling after through-connect has corrected the wrong EVS-ICM. The rules in **TS 26.114 do not cover this case, as the EVS-ICM rule for UMTS_EVS is still under discussion**.

In another call scenario, in figure 9.5.2, terminating side B could be a GERAN access with support for AMR-WB(0,1,2).

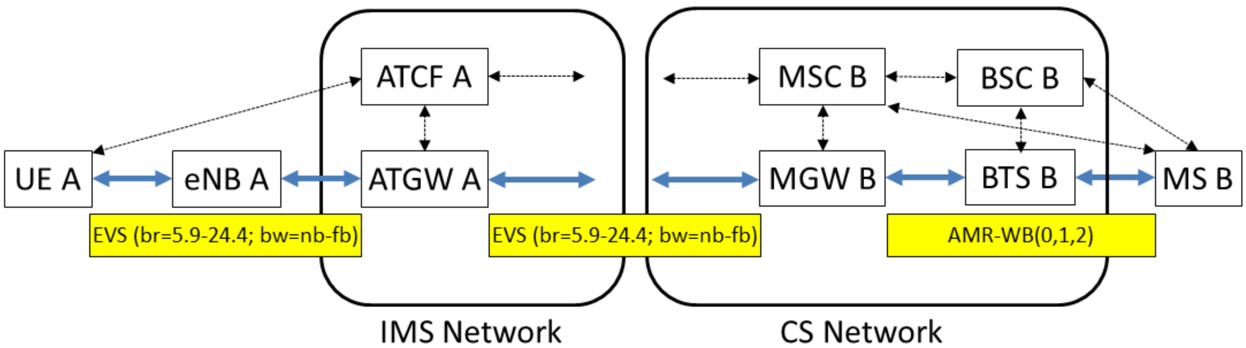

Figure 9.5-2: Mobile<=>Mobile call between 4G and 2G accesses

The **LTE Used Codec** (UE A <=> ATGW A) is EVS (br=5,9-24.4; bw=nb-fb), all modes of EVS-IO() included.
The **IMS Selected Codec** (ATGW A <=> MGW B) is EVS (br=5.9-24.4; bw=nb-fb), i.e. the same.
The **GERAN-Used-Codec (MGW B <=> MS B)** is AMR-WB (0,1,2).

MGW B translates between EVS-packing and AMR-WB packing and between EVS-CMR and AMR-WB-CMR.

**It is important that the EVS-ICM for UE A in that case is equal or lower than AMR-WB (2)!
EVS primary modes are not allowed. The rules in TS 26.114 do not cover this case.**

9.6 Mode Control before eSRVCC

When the call is ongoing, i.e. is in State "Connected", CMR is permanently (preferred) or on demand (not recommended) exchanged in both directions, to control the optimal Codec Modes and EVS modes of operation in both end-points (both media-senders).In what follows, the call scenario in figure 9.5-1 is assumed.

At any time during the call some transport conditions may change, causing a node in the path to change the CMRs.

- RNC B could lower or raise the RNC-Max-Rates in one or both directions due to UTRAN load changes. RNC B would command UE B by an RRC-command and MGW B by a PDU Type 14 Rate Control command. MGW B would send modified CMR towards UE A to reflect that change.
- ATGW A could detect a high uplink frame loss rate and high RTP jitter coming from UE A and may command by CMR a lower Codec Rate in uplink for UE A.
- To combat the high frame loss rate ATGW A could also command UE A to go into the EVS Channel Aware mode, by sending an EVS-CA-CMR command down to UE A. UE B would have to handle this EVS CA mode for decoding. EVS-CA mode of operation is not allowed, if the remote end is using AMR-WB, unless transcoding is inserted in MGW B.
- UE A could detect a high frame loss rate in downlink and send CMR uplink, requesting the EVS CA mode in downlink. ATGW A would have to allow this EVS-CA-CMR to pass through to UE B (or block it), UE B would have to send in EVS Channel Aware mode. MS B, using AMR-WB, would not understand this. Many more examples can be found.

In general, the call may be in any EVS mode of operation, when an eSRVCC is triggered.

Indeed eSRVCC is only one additional reason to trigger Codec Mode Control.

9.7 Mode Control during eSRVCC

Assume the call is ongoing as in figure 9.7-1, with EVS end-to-end and a 3G access at the remote end. The EVS CA mode of operation may be used in both directions.

UE A is roaming and observing its radio environment. It detects that the LTE radio leg is degrading, while a **2G radio leg** is strong, 3G is not detected. UE A sends measurement reports to eNB A and this triggers the eSRVCC to 2G.

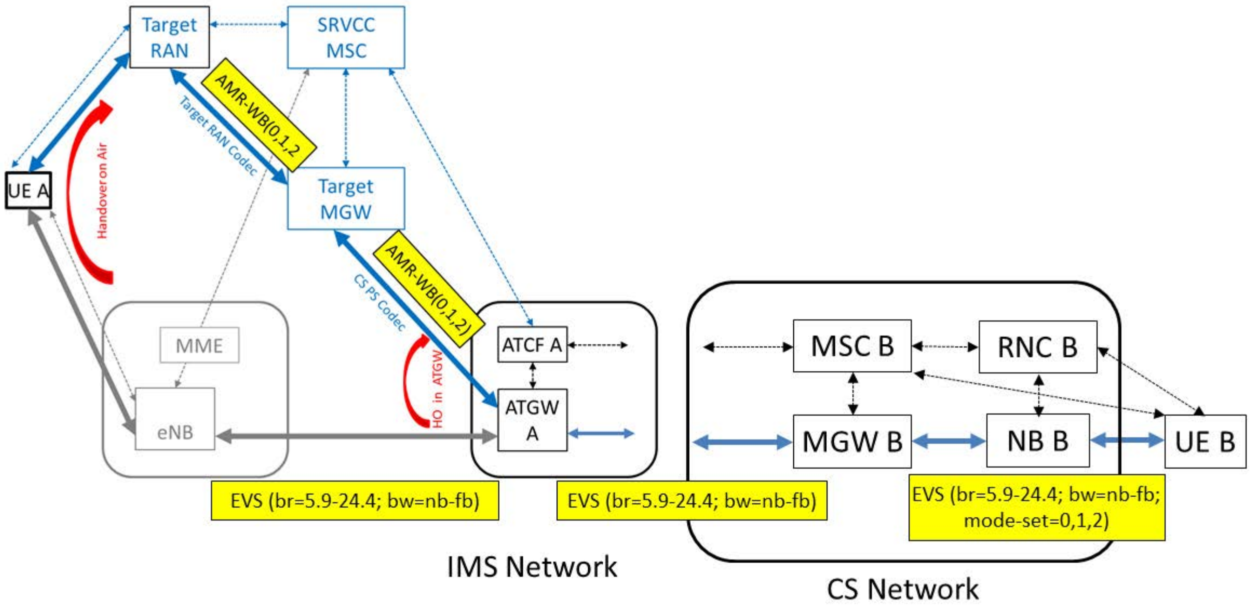

Figure 9.7-1: Call Scenario during eSRVCC to GERAN

The SRVCC MSC selects AMR-WB (0,1,2) as Target RAN Codec and prepares the Target Radio leg and the Target MGW. Then the SRVCC MSC sends message 10a, SIP Invite (MSC Preferred Codec List 2), to ATCF A. This ATCF selects AMR-WB (0,1,2) as CS-PS-Codec and informs the ATGW. While ATGW A prepares the resources for the access transfer, it may already send CMR towards the remote 3G UE B to switch from EVS CA mode to EVS-IO mode of operation.
This Pre-SRVCC Mode Control by the ATGW is not standard agreement and would be implementation dependent.

Then ATGW A returns the connectivity parameters to the ATCF and further to the SRVCC MSC. The Target access leg is prepared. The ATGW switches the User plane sharply from the LTE access to the Target access.

The ATGW A may start sending for a while RTP packets with EVS-CA mode towards the new Target MGW. These packets from the remote UE B are not understood and discarded by the Target MGW. Alternatively, the ATGW may send nothing to the Target MGW, until it receives EVS-IO frames from UE B and repacks them into AMR-WB format.

The Target BTS does still not receive uplink frames and sends nothing in uplink. The Target MGW may start sending CMR-Only RTP packets in AMR-WB payload format with AMR-WB-CMR=0 towards ATGW A to support Pre-SRVCC Mode Control. This Pre-SRVCC Mode Control by the Target MGW is not standard agreement and would be implementation dependent.

The earlier EVS-CMR-IO is sent towards the remote UE B, the better. If it is sent only after the Target BTS received the first speech frames on the new radio leg and these reach the ATGW, then the speech interruption in downlink is extremely long.

The SRVCC MSC sends message 13, PS to CS Response (Target RAN Codec), to the MME, triggering the Handover Command. While the Handover Command is on its way to UE A, Pre-SRVCC CMR-IO could reach UE B and UE B could start sending in EVS-IO mode already before the handover on air happens.

As soon as ATGW A gets these EVS-IO frames from UE B in RTP payload format for EVS, ATGW A repacks them into RTP payload format for AMR-WB and now the Target MGW can understand and forward them to the 2G radio leg. Depending on the remote 3G leg radio conditions, UE B sends CMR between EVS-CMR (br=24.4; bw=fb) and EVS-CMR (br=5.9; bw=wb). After receiving EVS-CMR-IO (br=6.6; bw=wb) from ATGW A, UE B may also start sending between EVS-CMR-IO (br=6.6; bw=wb), reflecting that it is now operating in the EVS-IO mode. The ATGW in the pathwill filter and translate the EVS-CMR coming from UE B into AMR-WB-CMR going to UE A.

Mode Control for the media stream downlink towards the Target RAN is in this scenario most critical.
The earlier this is triggered the better. It is important that EVS-CMR-IO (br=6.6; bw=wb) is fast and reliable received, understood and obeyed by UE B. This is important for a short speech break in local downlink.

Mode Control for the media stream **uplink** from the Target RAN is trivial in this scenario. The UE A starts in any case with the Initial Codec Mode of the Target RAN Codec, here with AMR-WB (0), if the standard is followed. This is always understood by the ATGW A. As soon as ATGW A receives RTP packets in payload format for AMR-WB from the Target MGW, ATGW A repacks them into RTP payload format for EVS and sends them towards UE B. This repacking includes the translation of the CMR commands.

The Target BTS and especially the Target MGW send immediately after eSRVCC AMR-WB-CMR=0 in all RTP packets towards ATGW A. AMR-WB-CMR=0 is translated by ATGW A into EVS-CMR-IO and it would be best to send CMR-EVS-IO in **all** RTP packets towards UE B. In case of a speech pause, CMR-only packets should be sent for a while repeatedly.

The Target BTS sends AMR-WB-CMR=0 downlink on the new radio channel to keep UE A in the Initial Codec Mode for a while. This is done, until the new radio channel is observed and measured long enough to decide the optimal mode.

It is in general much better to use error free frames in a low mode, than to risk lost frames in a high mode.

9.8 Mode Control after eSRVCC

UE A receives the Handover Command via the LTE leg and starts as soon as possible switching to the Target Radio leg. This Handover on air takes a while and is dependent on the radio leg standard and on implementation skills in the UE. Let's say the UE "disappears" from LTE and "appears" on 3G about [100ms] later, to take this just as a "house number".

Because UE A used EVS before eSRVCC, it may use the EVS Codec algorithm also after eSRVCC for encoding and decoding in the EVS-IO mode of operation. In case of eSRVCC from EVS to AMR-WB, there is no need to restart the Codec algorithm. All State-Variables of the Codec algorithm can be used as they are and this helps to combat the speech path interruption.

NOTE: When transitioning from EVS primary mode to EVS-IO mode due to eSRVCC, further adjustements may become necessary, for instance regarding the AMR-WB mode-set, mode-change-period and mode-change-neighbor used in the IMS network, which can necesitate the usage of a re-INVITE or RTCP APP control in IMS, if the parameters in the IMS network are not reasonably chosen. If only mode control is used towards the remote end, unnecessary radio bandwidth for high EVS modes will remain allocated, whereas the far-end network could use a re-INVITE as a trigger for adjusting radio resources at the remote end, as long as the local end uses "only" AMR-WB.

"Reasonable" network configuration is among the most important tasks of every operator and for "resonable" agreements between operators.

This "overprovisioning" problem on one or the other access (or even both) is nothing specific to this scenario. It is inherent to all call scenarios with multi-mode Codecs, also in PS<=>PS calls with AMR (...), AMR-WB (...) or EVS (...).

In any case, it is important that UE A starts/continues after eSRVCC with EVS-IO, sending these EVS-IO coded speech frames in uplink in the format of AMR-WB. This Initial Codec Mode will be kept, until the Target BTS sends AMR-WB-CMR with other, higher values, indicating that the uplink radio leg is good enough. Typically, it takes about [500 ms], until UE A and Target BTS have observed the new radio leg and determined the best codec mode in downlink and uplink. Then the Target BTS will allow CMR up to CMR=2 in downlink and UE A will send CMR up to CMR=2 in uplink and after one more round trip time the call is in the best possible Codec Modes after eSRVCC.

10 SDP Offer-Answer between MSC and ATCF

10.1 General

Clause 5 describes the basic eSRVCC procedure in principle; this clause discusses the communication between SRVCC MSC and ATCF in more detail, considering the current eSRVCC standard.

Figure 5.1-2 shows the simplified message flow for eSRVCC according to Stage 2.

10.2 Message and Information from MSC to ATCF

Message 10a, SIP Invite (MSC Preferred Codec List 2), is the first message from MSC to ATCF in the ongoing eSRVCC procedure. At that point, the Target RAN leg is more or less completely setup and all necessary resources are allocated. Only the link between Target MGW and ATGW is missing. The Target RAN Codec has been selected, "guessed", based on local criteria only. The IP Address and UPD Port (connectivity data) of the Target MGW are allocated. The SRVCC MSC assembled its "MSC Preferred Codec List 2", with the Target RAN Codec at the first place in this ordered list.

Message 10a contains mainly this MSC Preferred Codec List 2 and the connectivity data of the Target MGW, besides the necessary call identifier, allowing the ATCF to find the concerned ongoing call.

10.3 Information in ATCF and ATGW and actions

The ATCF knows the IMS Selected Codec of the ongoing call and all alternative Codecs, which are supported by the ATGW. In principle, the ATCF may also know the "remote Supported Codec List", i.e. the list of all Codec candidates for TLCI between the ATGW and the remote end. The IMS Selected Codec is one Codec of that list.

The ATCF does not know the capabilities of the Target RAN, until Message 10a arrives. This is also the point, when the ATCF gets knowledge that eSRVCC is necessary. Before, no preparation was possible.

The ATCF takes the MSC Preferred Codec List and selects the CS-PS Codec for the link between ATGW and Target MGW. It is not specified, how the ATCF derives this selection.

The selection seems obvious, if IMS Selected Codec and Target RAN Codec are identical or at least TLCI-compatible. If these Codecs do not match, then transcoding will be inserted and the choice for the CS-PS Codec is less obvious. With the selection of the CS-PS Codec the ATCF has the power to decide, where transcoding has to be inserted, if needed.

Then ATCF sends Message 10b, Session Transfer (CS-PS Codec), to the ATGW. Message 10b contains also the connectivity data of the Target MGW. The ATGW allocates the necessary resources, determines the IP address and UPD Port in the ATGW (connectivity data), and returns these to the ATCF. In that moment, the ATGW switched the User Plane from the old LTE leg to the new Target RAN leg sharply.

Message 10b is the first point, when the ATGW gets informed about eSRVCC. The ATGW decides, whether transcoding between the IMS Selected Codec and the CS-PS-Codec is necessary.

The ATGW may also detect and decide, if it is important that Mode Control commands are sent to the remote end, in order to bring the Remote Used Codec into the mode of operation and rate- and bandwidth-range, necessary to match the CS-PS Codec. That is not standardized and left for implementation. In the simplest case, the ATGW just allocates the resources and switches the User Plane from the old LTE leg to the new Target leg. If the Remote Used Codec is TLCI-compatible to the CS-PS-Codec, but currently operating in a non-compatible mode, then speech data from the remote end cannot be understood in the Target RAN. This causes a muting period in the local downlink, until the Remote Codec is in the right mode of operation.

10.4 Message from ATCF to MSC, MGW actions

Message 11b, SIP Response (CS-PS Codec), contains the CS-PS Codec for the link between ATGW and Target MGW and the connectivity data of the ATGW. When the ATCF sends Message 11b to the SRVCC MSC, then the IP link between ATGW and Target MGW can be closed. Now data may already flow between these MGWs.

In principle, the Target MGW can now send Mode Control commands (CMR) towards the ATGW, hoping the ATGW would send them further towards the remote end. This is not standardized, too. If successful, it shortens at least the time until the remote end is in the right mode, although these Mode Control commands are already rather late.

Since the local UE has so far most likely not landed on the new radio leg, there are no Speech or SID frames arriving in the Target MGW in uplink. The Target MGW may, however, send CMR-Only packets towards the ATGW to initiate this Mode Control. The 3GPP standards for AMR and AMR-WB and the RTP payload format for these allow and recommend these CMR-Only packets (often called "No_Data" packets). In order to combat packet loss these CMR-Only packets should be repeated.

10.5 Message from MSC to MME and LTE UE

Figure 5.1-2 shows that the SRVCC MSC may send message 13, PS to CS Handover Response (Target RAN Codec), immediately after the Target RAN Codec is selected, even before the ATCF is involved by Message 10b.

Message 13 triggers the Handover Command towards the LTE UE. Sending Message 13 early accelerates the Handover on air, but it bears the risk that the resources in the ATGW are not ready, when the UE accesses the new radio leg. Sending Message 13 later, e.g. after Message 11b has been received from the ATCF, bears the risk that the handover on air is too long after the ATGW has switched the User Plane sharply. Whatever the MSC does, it seems insufficient for an optimal eSRVCC handover switching in real life networks with load and radio errors.

11 Codec Compatibility

11.1 Digital Mobile Communication

In all digital communication system the analogue voice signal (**Mic**rophone signal) is in one of the very first processing steps **A/D**-converted into a digital signal representation. The used sampling frequency (sf) has to be at least twice as high as the highest frequency of the voice band that is to be transmitted. The resolution of the signal amplitude has to be sufficiently high in order to not loose quality in this first step.

Not negligible is the limitation in the voice bandwidth: **n**arrow-**b**and (e.g. 300-3 400 Hz), **w**ide-**b**and (e.g. 100-8 000 Hz), **s**uper-**w**ide-**b**and (e.g. 50-16 000 Hz) or even **f**ull-**b**and (e.g. 20-20 000 Hz).

Some further typical steps in digital voice processing are Acoustic Echo Cancellation (**AEC**), Automatic Gain Control (**AGC**), Noise Reduction (**NR**) and maybe more, just to mention some of these, often proprietary algorithms. The resulting digital signal is still in linear PCM representation and has still a very high bit rate: too high for a commercially viable transmission in most wireless systems.

Therefore a very important step for interworking follows: the reduction of the bit rate with as little as achievable loss in signal quality. This step is called "Encoding" (**ENC**) and results in a substantially reduced bit rate. This is now much better suited for transmission over long distances and especially over wireless connections.

At the receiving side the counterpart, the "Decoding" (**DEC**) has to take place, typically followed by Gain Control (**GC**) - and more - and finally the **D/A**-conversion back into an analogue signal, which feeds the loudspeaker (**Lsp**).

Figure 11.1-1 shows the principle of this typical voice processing within two terminals A and B.

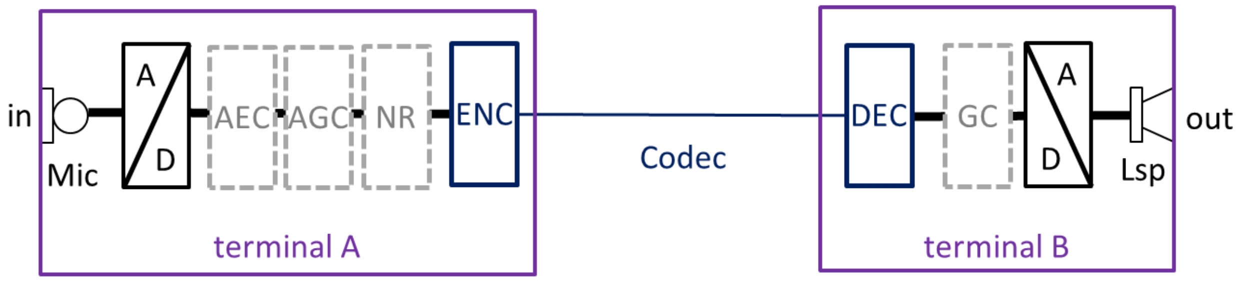

Figure 11.1-1: Principle of voice processing within two terminals A and B

In the present document a specific transmission link is named with the used **Codec** for that link. It is obviously indispensable that En**co**der and **Dec**oder on both ends of the **coDec**-link have to use identical or compatible codec algorithms.

Every Encoding - Decoding step causes degradation in speech quality. In fact, the main bottlenecks for voice quality are nowadays not in the Codec, but in the audio input/output of the terminals.

11.2 Transcoding

Figure 11.1-1 simplifies the connection between the terminals dramatically. In reality this connection is quite complex and often both terminals do not support the same Codec, therefore "Trans-Coding" has to take place. Transcoding is the "translation" from one Codec-language into another Codec-language. This Transcoding is performed within "Media GateWays" (MGW), see Figure 11.2-1.

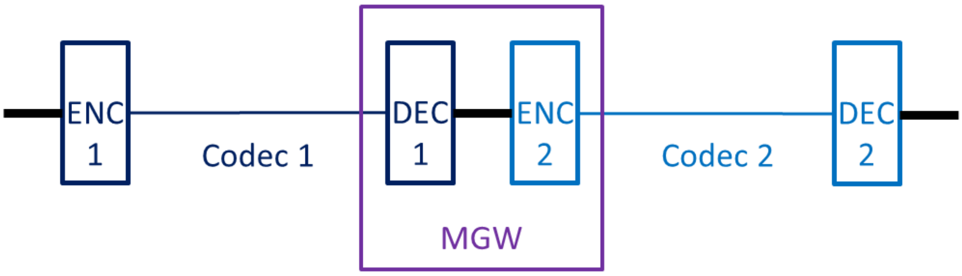

Figure 11.2-1: Principle of Transcoding

The typical Transcoding is a cascade of the Decoding of the signal on the incoming link back into the linear presentation and then the Encoding for the outgoing link. This second Encoding step causes another voice quality degradation. These two Codecs, Codec 1 and Codec 2, are called here to be "in tandem". Tandem Free Operation (TFO) was the first attempt to avoid this quality loss for the call cases, where both Codecs, "right" and "left" of the MGW, or right and left of a PCM-coded link, were TFO-compatible.

11.3 EVS configurations

11.3.1 General

The SDP media parameters and the RTP payload format of the EVS codec are specified in 3GPP TS 26.445 [8], annex A. The EVS Codec includes EVS Primary modes and EVS AMR-WB IO modes.

For EVS Primary modes, the specification of the RTP payload format of the EVS codec includes media parameters in SDP to specify/negotiate bit rates (symmetric or asymmetric) and audio bandwidths (symmetric or asymmetric). For simplicity of the discussion, only the symmetric SDP parameters 'br' and 'bw' are considered in the following. An excerpt of the definitions of 'br' and 'bw' is provided below.

Begin of cite from 3GPP TS 26.445 [8]:

br: Specifies the range of source codec bit-rate(s) for EVS Primary mode (...) to be used in the session, in kilobits per second, for the send and the receive directions. The parameter can either have: a single bit-rate (br1); or a hyphen-separated pair of two bit-rates (br1-br2). If a single value is included, this bit-rate, br1, is used. If a hyphen-separated pair of two bit-rates is included, br1 and br2 are used as the minimum bit-rate and the maximum bit-rate respectively. br1 shall be smaller than br2. br1 and br2 have a value from the set: 5,9, 7.2, 8, 9,6, 13,2, 16,4, 24,4, 32, 48, 64, 96, and 128. 5,9 represents the average bit-rate of source controlled variable bit rate (SC-VBR) coding, and 7.2, …, 128 represent the bit-rates of constant bit-rate source coding. Only bit-rates supporting at least one of the allowed audio bandwidth(s) shall be used in the session.

bw: Specifies the audio bandwidth for EVS Primary mode (...) to be used in the session for the send and the receive directions. bw has a value from the set: nb, wb, swb, fb, nb-wb, nb-swb, and nb-fb. nb, wb, swb, and fb represent narrowband, wideband, super-wideband, and fullband respectively, and nb-wb, nb-swb, and nb-fb represent all bandwidths from narrowband to wideband, super-wideband, and fullband respectively.

End of cite from 3GPP TS 26.445 [8].

For EVS AMR-WB IO modes, the specification of the RTP payload format of the EVS codec includes the same 'mode-set' media parameter in SDP as AMR-WB (IETF RFC 4867 [9]) for compatibility reasons.

A comprehensive list of EVS operating points is illustrated in table 11.3-1, together with AMR and AMR-WB operating points.

Table 11.3-1: The complete set of EVS bit rates and audio bandwidths (with AMR and AMR-WB modes for comparison). Numbers are in kbps

AMR		4.75	5.15	5.9		6.7		7.4	7.95				10.2	12.2															
AMR-WB					6.6						8.85				12.65		14.25	15.85		18.25	19.85	23.05	23.85						
EVS-NB	2.8 [1]						7.2 [1]			8 [1]		9.6				13.2			16.4					24.4					
EVS-WB	2.8 [1]						7.2 [1]			8 [1]		9.6				13.2 [2]			16.4					24.4	32	48	64	96	128
EVS-SWB												9.6				13.2 [2]			16.4					24.4	32	48	64	96	128
EVS-FB																			16.4					24.4	32	48	64	96	128
EVS AMR-WB IO					6.6						8.85				12.65		14.25	15.85		18.25	19.85	23.05	23.85						

NOTE 1: The EVS-VBR mode of operation has an expected average bit rate of 5,9 kbps in active periods of speech. This 5,9 kbps does not correspond to any physical bit rate. Instead it is composed of three constant bit rate modes of 2,8, 7,2 and 8,0 kbps, as marked in this table.

NOTE 2: At 13,2 kbps there are two EVS modes of operation, the "normal" constant bit rate mode and the "channel-aware" mode (EVS-CA), with dynamically varying "primary" and "secondary" (redundant) parts inside the constant net bit rate.

NOTE 3: The present document does mainly consider the parameters for bit rates and audio bandwidths for EVS Primary and EVS AMR-WB IO modes. It should be noted that for overall interworking all media type parameters are taken into consideration.

In the following two types of EVS configurations are discussed: **Bottom-up configurations** and **single-audio bandwidth configurations**, which belong to the biggest class of all, the "**Punctured Configurations**".

11.3.2 The EVS Bottom-up Configurations

EVS "**Bottom-up Configurations**" are characterized by two important facts:

a) they always include all lower bit rates, from 5,9 kbps up to their maximum bit rate;

b) they always include all lower audio bandwidths, from nb up to their maximum audio bandwidth.

Table 11.3.2-1 shows the complete table of EVS Primary modes that constitute also all possible **Bottom-up Configurations**. The mandatory EVS AMR-WB IO modes are not shown in this table.

Table 11.3.2-1: The complete table of EVS (Primary) modes

Rate (kbps)	Narrow-Band bw=nb	Wide-Band bw=wb	Super-Wide-Band bw=swb	Full-Band bw=fb
br=128		(br=128; bw=wb)	(br=128; bw=swb)	(br=128; bw=fb)
br=96		(br=96; bw=wb)	(br=96; bw=swb)	(br=96; bw=fb)
br=64		(br=64; bw=wb)	(br=64; bw=swb)	(br=64; bw=fb)
br=48		(br=48; bw=wb)	(br=48; bw=swb)	(br=48; bw=fb)
br=32		(br=32; bw=wb)	(br=32; bw=swb)	(br=32; bw=fb)
br=24.4	(br=24.4; bw=nb)	(br=24.4; bw=wb)	(br=24.4; bw=swb)	(br=24.4; bw=fb)
br=16.4	(br=16.4; bw=nb)	(br=16.4; bw=wb)	(br=16.4; bw=swb)	(br=16.4; bw=fb)
br=13.2	(br=13.2; bw=nb)	(br=13.2; bw=wb)	(br=13.2; bw=swb)	
br=9.6	(br=9.6; bw=nb)	(br=9.6; bw=wb)	(br=9.6; bw=swb)	
br=8.0	(br=8; bw=nb)	(br=8; bw=wb)		
br=7.2	(br=7.2; bw=nb)	(br=7.2; bw=wb)		
br=5.9 (VBR)	(br=5.9; bw=nb)	(br=5.9; bw=wb)		

Each of these 35 EVS Primary modes is an "upper right corner" of one of the 35 EVS "**Bottom-up Configurations**". **EVS** (br=128; bw=fb) is the biggest EVS mode and the upper right corner of the biggest EVS **Bottom-up Configuration**.

The SDP notation for EVS specifies "**EVS** (br=128; bw=fb)" by: encoding name=EVS; br=128; bw=fb.

The biggest **EVS Bottom-up Configuration** is EVS (br=5.9-128; bw=nb-fb) and it contains all 35 EVS Primary modes.
Important other **EVS Bottom up Configurations** are:

- EVS (br=5.9-8; bw=nb -wb);
- EVS (br=5.9-13,2; bw=nb-swb); and
- EVS (br=5.9-24.4; bw=nb-fb).

These are candidates for EVS over Circuit Switched networks and important for MTSI clients, too.

EVS Bottom up Configurations have important properties:

1) All possible EVS Bottom up Configurations are TLCI-compatible to each other.
2) The intersection of EVS Bottom up Configurations leads always to an EVS Bottom up Configuration.
3) Transcoding free interworking between two or more different EVS Bottom up Configurations may use all common EVS Modes, i.e. the intersection of all EVS Bottom up Configurations in the call path.

The resulting EVS Bottom up Configuration at call setup (or after Handover, or after Codec Renegotiation), negotiated by SIP/SDP- or CS-Signalling, is named the "**Framework Bottom up Configuration**" for this call. The Framework Bottom up Configuration can only be changed by Codec Renegotiation, typically resulting in a speech path interruption.

EVS Mode Control by EVS-CMR may shrink or expand the "**active Bottom up Configuration**", but never expand the active Bottom up Configuration beyond the boundaries of the Framework Bottom up Configuration. EVS Mode Control by EVS-CMR does not cause speech path interruptions.

For completeness, each EVS Bottom up Configuration has a complementing EVS AMR-WB IO Configuration.

By definition in the present document, the EVS Channel-Aware (EVS-CA) mode is always included in the Framework Bottom up Configuration, as far as maximum bit rate is at least 13,2 kbit/s and bandwidth of the Framework Bottom up Configuration is at least wb.

The EVS-VBR mode of operation is always included in all Framework Bottom up Configurations.

The name of the biggest EVS mode in a requested active Bottom up Configuration determines also the name for the corresponding EVS-CMR command. Example: "EVS-CMR (br=24.4; bw=swb)" commands the remote partner(s) to shrink or expand the active Bottom up Configuration to EVS (br=5.9-24.4; bw=nb-swb).

The **Maximum Mode Control** signalling is typically started by the media-receiver, which sends in its EVS-CMR the highest possible EVS mode in is prepared to receive. Each node in the speech path (e.g. MGW) may modify the EVS-CMR on the fly according to the Maximum Mode Control principle, i.e. it may shrink the requested active Bottom up Configuration, but never expand it. In this way the (modified) EVS-CMR, which is finally received by the media-sender, indicates the biggest possible active Bottom-up Configuration in that very moment, for the whole path from media-sender to media-receiver.

11.3.3 The EVS Punctured Configurations

11.3.3.1 General

If one or more lower audio bandwidths than the maximum negotiated bandwidth or one or more lower bit rates than the maximum negotiated bitrate are not included in an EVS Configuration, as negotiated/selected by SIP/SDP signalling or CS-signalling, then this EVS Configuration is not a Bottom up Configuration and such an EVS Configuration is not TLCI-compatible to any of the EVS Bottom up Configurations. Such an EVS Configuration is called a "Punctured Configuration". Punctured Configurations are typically not TLCI-compatible to most other EVS Configurations.

EXAMPLE 1: EVS (br=7.2-24.4; bw=nb-swb) in client 1 is NOT-TLCI compatible to the EVS Bottom-up Configuration
EVS (br=5.9-24.4; bw=nb-swb) in client 2, because the EVS-CMR has been defined to indicate only the maximum mode (maximum bit rate and maximum audio bandwidth). EVS-CMR (br=7.2; bw=swb) from client 1 to client 2 would not disallow EVS (br=5.9; bw=wb) to be used by the media-sender in client 2, although client 1 is not allowed to use it.
Vice versa, EVS-CMR (br=5.9; bw=wb) from client 2 would not be followed by client 1.

EXAMPLE 2: EVS (br=13.2; bw=nb-swb; CA=on) in client 1 is NOT TLCI-compatible to the EVS Bottom-up Configuration
EVS (br=5.9-13.2; bw=nb-swb; CA=on) in client 2 due to the Maximum Mode Control, although EVS-CA mode of operation is included in both Configurations and EVS AMR-WB IO is also included in both.

11.3.3.2 EVS Configurations with single audio bandwidth

It is possible to use EVS at a single audio bandwidth by specifying a single bandwidth value (e.g. "bw=swb"). These single audio bandwidth Configurations form specific classes of Punctured Configurations.

The EVS SDP parameters and the RTP Payload Format (3GPP TS 26.445 [8]) and the profiling in 3GPP TS 26.114 [5] allow many Punctured Configurations and many single-audio bandwidth Configurations in SIP/SDP. The "biggest" single audio bandwidth Configurations are as shown in table 11.3.2-1:

- EVS-NB (br=5.9-24.4; bw=nb);
- EVS-WB (br=5.9-128; bw=wb);
- EVS-SWB (br=9.6-128; bw=swb);
- EVS-FB (br=16.4-128; bw=fb).

The advantage of such single audio bandwidth configurations is that they can guarantee that the specified single audio bandwidth is used, as far as the audio input signal provides it. They can allow testing the EVS codec in a well-defined operation point and simplify the usage of the EVS codec. The disadvantage is that lower bit rates are not always allowed, potentially compromising the radio error performance in marginal radio conditions and requesting higher cell capacity in case of network overload.

Only the set of single audio bandwidth Configurations with (br=5.9-brmax; bw=**nb**) contains Bottom-up Configurations. All other single-audio bandwidth Configurations are Punctured Configurations and not TLCI-compatible to Bottom-up Configurations and not to each other.

Important property:

Two single audio bandwidth Configurations are only TLCI-compatible, if they share the same audio bandwidth and the same lowest bit rates.

EXAMPLE 1: EVS (br=9.6-24.4; bw=swb) and EVS (br=9.6-13.2; bw=swb) are TLCI-compatible.
The latter is candidate for EVS over CS networks (called "Set 3").

EXAMPLE 2: EVS (br=9.6-13.2; bw=swb) and EVS (br=9.6-128; bw=swb) are also TLCI compatible.

Punctured single audio bandwidth Configurations consist themselves of many TLCI-compatible punctured Configurations of the same single audio bandwidth. EVS (br=9.6-128; bw=swb) has in total 9 TLCI-compatible Configurations, see table 11.3.2-1. In some sense, these single audio bandwidth Configurations represent Codecs like AMR or AMR-WB, which have only one bandwidth, but a set of bit rates.
Note that EVS (br=13.2-128; bw=swb) includes another set of Configurations, which are not compatible to the one above.

NOTE 1: Interworking between any EVS Bottom up Configuration and a punctured single audio bandwidth Configuration requires always transcoding and this leads always to lower speech quality. Even the fall back to the EVS AMR-WB IO mode of operation reduces the resulting speech quality, although TLCI is possible.

NOTE 2: An EVS Bottom up Configuration allows operating a call end-to-end in a single selected audio bandwidth, e.g. in swb, as long as all nodes in the path allow the necessary bandwidth and bit rates. The Bottom up Configuration EVS (br=5.9-128; bw=nb-swb) includes the single-audio bandwidth Configuration EVS (br=9.6-128; bw=swb) in that sense. Nevertheless, transcoding is inserted between these two at call setup (or handover). Adaptation by EVS-CMR between these two Configurations is not possible, because EVS-CMR specifies and changes always only the maximum allowed mode, but does not exclude lower rates or lower bandwidths.

NOTE 3: The media-sender in client 1, setup with the punctured single audio bandwidth Configuration EVS (br=9.6-24.4; bw=swb; mode-set=0,1,2) may indeed use all EVS primary modes of the following punctured Configuration: EVS (br=9.6-24.4; bw=nb-swb; mode-set=0,1,2), if the audio input signal is classified by the EVS encoder as nb or wb or swb. This is inherent to the EVS Codec algorithm. However, rule is that client 1 sends EVS-CMR within the local Configuration, i.e. use only bw=swb and br between 9,6 and 24,4. All MGWs in the path do not modify the EVS-CMR to command a mode below/outside the selected Configuration. The sent and/or received media stream may, however, contain speech packets with EVS (br=9.6 ... 24.4; bw=nb ... swb) and client 1 accepts and decodes these.

11.4 Transcoding Less Operation

11.4.1 General

Transcoding Less Operation (TLCI) is of key importance to many voice service aspects. High Definition Voice services (HD Voice) is an important example (although - strictly speaking - transcoding may occur also in some HD Voice calls, see below). It is important that the Codecs used at both ends of the communication are TLCI-compatible to achieve best possible quality, as transcoding always degrades quality.

In its simplest form Codec 1, left of the MGW and Codec 2, right of the MGW, are identical. The MGW detects this and connects both links without transcoding. It is, however, not strictly required that both Codecs are identical to avoid Transcoding. It is sufficient that both Codecs are TLCI-compatible. Table 11.4-1 lists TLCI-compatible 3GPP Codecs.

Table 11.4-1: Examples of TLCI-compatible 3GPP Codecs.

Codec 2 ==> Codec 1	GSM EFR	AMR (7)	AMR (0,2,4,7)	AMR-WB (0,1,2) and AMR-WB()	AMR-WB or EVS-IO Bottom up Configurations	EVS Bottom up Configurations	EVS single audio bandwidth Configuration (nb / wb / swb / fb)
GSM EFR	TLCI	SID-Con					
AMR(7)	SID-Con	TLCI					
AMR (0,2,4,7)			TLCI				
AMR-WB (0,1,2) and AMR-WB()				TLCI	TLCI by Mode-ctrl	TLCI by Mode-ctrl	(TLCI via AMR-WB IO)
AMR-WB or EVS-IO Bottom up Configurations				TLCI by Mode-ctrl	TLCI	Mode-ctrl CMR-IO ≤ 8	(TLCI via AMR-WB IO)
EVS Bottom up Configurations				TLCI by Mode-ctrl	TLCI by Mode-ctrl	TLCI	
EVS single audio bandwidth Configuration (nb / wb / swb / fb)				(TLCI via AMR-WB IO)	(TLCI via AMR-WB IO)		TLCI only to itself (NB<=>NB etc.) and via AMR-WB IO

The diagonal "upper-left to lower-right" of Table 11.4-1 shows "**TLCI**" in all squares: Codec 1 and Codec 2 are identical or TLCI-compatible by Mode Control. Empty squares indicate: transcoding is required.

The green marked area indicates the TLCI-compatibility section of the EVS Bottom up Configurations in combination with all the EVS AMR-WB IO Configurations that include all lower modes up to their maximum. If TLCI is put in brackets (TLCI) in some squares, then additional conditions exist for compatibility. For example, an EVS single-audio bandwidth Configuration may include the EVS AMR-WB IO () Configuration and then this is TLCI-compatible to AMR-WB (0,1,2) or any other EVS AMR-WB IO Configuration that include all lower modes up to their maximum or any AMR-WB Configuration that include all lower modes up to their maximum.

The AMR-WB (0,1,2) is explicitly listed, although it belongs to the class of AMR-WB Configurations that include all lower modes up to their maximum, because it is an important and well-known CS-Codec.

11.4.2 GSM_EFR and AMR (mode-set=7)

11.4.2.1 General

GSM_EFR and AMR (mode-set=7) are "nearly" TLCI-compatible: the Speech frames are compatible, i.e. a GSM_EFR encoded frame can be decoded by AMR and an AMR (7) encoded frame can be decoded by GSM_EFR. The SID frames of both are, however, different and a "SID Conversion" (SID-Con) is needed. The term "SID Transcoding" is not used here, as the conversion is done without full decoding/encoding. SID frames describe the background noise in speech pauses and a small deviation in background noise is typically not perceivable by end-users, so GSM_EFR and AMR (7) are called TLCI-compatible. GSM_EFR and AMR (7) play still an important, although decreasing role in many GERAN and UTRAN networks.

11.4.2.2 Additional Conditions for TLCI-Compatibility for GSM_EFR

GSM_EFR and AMR (7) are single-rate-single-band Codecs. They are specified in all parameters in a way that they are always TLCI-compatible. DTX may be switched ON/OFF in the encoding side in 2G networks, but the decoding side and all network elements in the path are able to handle DTX. In 3G networks, DTX in uplink is always mandatory for AMR (7).

11.4.3 AMR

11.4.3.1 General

Important Codecs are AMR(0,2,4,7), AMR(0,2,4), AMR(0,2) and AMR(0). Not all of these Codecs are listed explicitly in Table 11.4-1 to keep the table readable. Please note that these four Codecs should be kept formally as four different Codecs: same Codec Type, but different Codec Configurations. They are all TLCI-compatible under the important assumption that the Rate Control rules are strictly followed by all terminals and all nodes in the voice path! For details, see clause 9.

EXAMPLE: Codec 1 == HR_AMR(0,2,4) ----- Codec 2 == AMR(0,2,4,7) ----- Codec 3 == UMTS_AMR2(0,2)/SF=256.

This cascade of a GERAN----Core----UTRAN call is transcoding free for the two AMR-modes 0 (4.75) and 2 (5,90). Rate Control end-to-end (CMR $\leq$ 2) ensures that the maximum Rate is 5,90, i.e. mode=2. If one of the partners would not comply to AMR Rate Control rules, then transcoding would have to be included with lower voice quality than AMR(5,90) end-to-end. Otherwise one side of the call could end in "silence", e.g. if the GERAN side sends with AMR(4) the UTRAN side could not receive this and would go muting. Even worse: the AMR-SID frames, sent in speech pauses, would be able to pass and be decoded: the UTRAN side would not be totally silent, but background noise and some speech clips could be heard.

The term "SF=256" denotes here the WCDMA Spreading Factor 256 and SF=128 the WCDMA Spreading Factor 128.

Another important AMR Configuration is AMR (mode-set=0,2,5,7). This is not TLCI-compatible to AMR (mode-set=0,2,4,7), but has otherwise similar properties.

11.4.3.2 Additional Conditions for TLCI-Compatibility for AMR

The most used, recommended AMR Configuration, AMR (mode-set=0,2,4,7), is a multi-rate-single-band Codec. It is deployed in GERAN and UTRAN and in MTSI with different options for the transport of mode-control-commands and the rate switching. Not all these options are TLCI-compatible. It is therefore important to obey some rules. It is not recommended to deploy the AMR in another way, as there is no obvious advantage. These rules and additional conditions are the following.

GERAN allows changing the rate in media-sending and media-receiving direction every second speech frame, i.e. every 40ms, but not in between. Changing the rate in between, at the wrong phase, causes a severe decoding error and a substantial, potentially catastrophic quality loss. It is indispensable that every remote partner that wants to be TLCI-compatible with GERAN obeys this additional condition. Typically, rate changes occur far less often than every 40ms and this additional condition is de facto no disadvantage, but not obeying it causes either the need for transcoding or severe quality degradations.

The Codec Type UMTS_AMR allows the encoder changing the rate every frame. UMTS_AMR is therefore not TLCI-compatible to GERAN and not recommended for any use today.

The Codec Type UMTS_AMR2 obeys this additional condition and is therefore the recommended multi-rate Codec Type in UTRAN.
MTSI deploys the AMR, but allows a multitude of options for rate switching. The SDP Parameters "mode-change-capability" and "mode-change-period" (see 3GPP TS 26.114 [5] and IETF RFC 4867 [9]) allow negotiating this GERANspecific condition. Because it is typically unknown at call setup (or handover, or re-negotiation), where a call is routed to and which access is used at the remote end, it is recommended to always set mode-change-capability=2 in the SDP Offer. TS 26.114 table 6.1 mandates this. If that is not included in the SDP Offer or SDP Answer towards a CS Network with GERAN (and UTRAN), then it is unavoidable to insert transcoding. The safest way it to include mode-change-capability=2 also in every SDP Answer.

GERAN mandates an AMR Encoder switching the rate only one step up or down. This second additional condition was intended for optimal channel decoding at the radio receiver side (most likelihood decoding in case of bad radio channels). Every GERAN mobile obeys this rule in uplink. In downlink, however, it is necessary to accept also other changes, because handover or other events may change the rate unpredictably. In good radio conditions, this is no problem and therefore this second additional condition is less stringent. Nevertheless, it is recommended to obey it by every AMR Encoder. The SDP Parameter "mode-change-neighbor" (see IETF RFC 4867 [9]) allows negotiating this additional condition. If this is not achieved, then the call may continue without Transcoding: the degradation to be expected is less severe than transcoding and far less severe than mode switching in the wrong phase.

CS Networks are sending one speech frame in one RTP packet or one Iu PDU Type 0, typically one every 20ms. Several alternative transport solutions exist due to history and development of the standard, like AoTDM and AoIP, but these are simply selected depending on the version of the control protocol. This guarantees minimal transport delay and simple interworking.

IETF RCF 4867 [9], however, allows a multitude of packing options, e.g. packing of multiple speech frames into one RTP packet in order to reduce the number of packets per second and to reduce the packet overhead. This increases the speech path delay. IETF RCF 4867 [9] allows also sending a speech frame redundantly several times in several consecutive RTP packets in order to reduce the rate of lost frames. This increases the speech path delay, too. Other options are octet-aligned or bandwidth-efficient packing, or inclusion of CRC, or robust sorting.
These different transport conditions are, however, no problem for TLCI-compatibility! If different packing methods are deployed along a speech path, then MGWs are inserted to re-pack and, if necessary deploy buffering, but they do not need transcoding, as long as both codecs at both sides of the MGW are TLCI-compatible with respect to the other rules.

GERAN transports the Codec Mode Request (CMR) endlessly repeated in every second speech frame, i.e. every 40ms. Always the "**active** Codec Mode Restriction" is sent; there is no "neutral" CMR-value defined. This little overhead (2 bit every 40 ms) guarantees that the CMR-status is always clear and transmission errors are quickly healed, rate changes are achieved as fast as possible.
In UTRAN and the CS Core Network, Rate Control Commands are transported totally different: only on demand, i.e. only when a Codec Mode Restriction changes. It is important that the receiver of such a Rate Control Command on demand remembers always the latest received one. It needs a Rate Control Command Status-memory. MGWs in the CS Core Network terminating GERAN translate these different signalling means. That is no severe problem as long as no transmission errors occur. In case of transmission errors (e.g. loss of an on-demand Rate Control Command), it takes quite a while, until the error is detected and corrected. This is, however, not judged as a TLCI-compatibility problem.

In MTSI every RTP packet includes a CMR-bit-field. It was for long time not clearly mandated that this CMR-bit-field contains always the active Codec Mode Restriction, as mandated in GERAN. A CMR-code-point "CMR=15" was defined with ambiguous meaning, leading to severe interworking issues. Now (since 2015) this is clarified: CMR=15 has exactly the same meaning as "the active Codec Mode Restriction is equal to the maximum mode of the selected mode-set". An important consequence of this clarification has to be obeyed for the case that two different AMR mode-sets are selected at both sides of a MGW, e.g. AMR (mode-set=0,2,4) <==> MGW <==> AMR (mode-set=0,2,4,7). Such a situation can occur during a call, e.g. due to eSRVCC or handover. If the MGW receives CMR=15 on the left side, then it has to (!) translate this to CMR=4 on the right side. If the MGW does not perform this correctly, then the call may go into muting on the left side! If the MGW receives CMR=15 on the right side, equivalent to CMR=7, then it would be good (!) to translate this to CMR=4 on the left side. CMR=7 is - strictly speaking - outside the selected mode-set on the left side, but it's clear and tolerable for many receivers. Some receivers, however, ignore any CMR outside their mode-set, because IETF RFC 4867 [9] recommends this. This may cause interworking issues. CMR=15 on the left side is possible, but not recommendable due to this ambiguity and existing legacy equipment using CMR=15 in different ways.

This Mode-Control additional condition is not a severe TLCI-compatibility problem, but MGWs have to obey it, otherwise transcoding would be required, or the call fails (muting), with no gain and only higher costs and quality degradation.

Some implementations have been observed in the past that did not obey a received CMR or Rate Control Command from the remote side, because they did not observe any local reason to restrict the rate. This behaviour is clearly outside the AMR standard and causes call failure (muting).

The 2G network always supports DTX in uplink and downlink, but may enable/disable DTX on the encoding side. In most 2G networks, DTX is enabled in uplink and supported in downlink, but may be disabled in downlink inside network-located transcoders. DTX works well end-to-end in both directions in case of TLCI. 3G networks always enable DTX in the encoding side (at least in UEs). The decoding side and all network elements in the path are required to be able to handle DTX, i.e. transport and decode SID frames.

11.4.4 AMR-WB and EVS-IO

11.4.4.1 General

The AMR-WB and the EVS AMR-WB IO are compatible Codec Types with 9 modes and bit rates each. The EVS AMR-WB IO is an integral part of the EVS Codec.

AMR-WB(0,1,2) is deployed world-wide in UTRAN as UMTS_AMR-WB(0,1,2)/SF=128 and in GERAN as FR_AMR-WB(0,1,2).

In VoLTE (MTSI) the higher modes of AMR-WB are deployed, too, notably the highest mode 8 (23.85). AMR-WB (mode-set=8) is not TLCI-compatible to any other AMR-WB Configuration, because it does not fulfil the Maximum Rate Control principle. In order to allow TLCI-Interworking between GERAN, UTRAN and VoLTE, an AMR-WB Configuration has to be used, that includes all lower modes up to its maximum, i.e. at least mode-set=0 or better, i.e. mode-set=0,1,2 **is to be** included in all Codecs in the path. It is recommended to deploy AMR-WB(), i.e. the AMR-WB with all 9 modes in VoLTE (MTSI), i.e. the full AMR-WB Configuration. All other AMR-WB Configurations, which include all lower modes up to their maximum, could also be used and TLCI would always be guaranteed.

A VoLTE<=>VoLTE call may use all 9 modes of AMR-WB () or EVS AMR-WB IO ().

A VoLTE <=> CS call with AMR-WB () <=> AMR-WB (0,1,2) may use the three lower modes, disallowing the higher modes by Maximum Rate Control: end-to-end Rate Control takes care that no mode higher than 2 is allowed: CMR≤2. Essential is, that the VoLTE-UE (any MTSI-client) follows the Rate Control commands strictly and as fast as possible. An important rule for Codec Negotiation is set in RFC 4867: "If an SDP Offer is received without a mode-set, then the Selected Codec may contain any mode-set, or no mode-set". The SDP Offer without mode-set is called "Open Offer" in 3GPP TS 26.114 [5]. The SDP Answer without mode-set is called "Open Answer".

AMR-WB(0,1,2) in end-to-end TLCI is better than AMR-WB(0,1,2) plus transcoding to AMR-WB(8).

Interworking between any AMR-WB mode-set and any EVS-IO mode-set is always transcoding free, if the mode-set are compatible, i.e. include all lower modes in common up to a certain maximum common mode.

11.4.4.2 Additional Conditions for TLCI-Compatibility for AMR-WB and EVS-IO

The same additional conditions as for AMR apply also for AMR-WB and EVS AMR-WB IO.

11.4.5 EVS

11.4.5.1 General

The most recent 3GPP Codec is the Codec for Enhanced Voice Services (EVS). EVS supports four different audio bandwidths (NB, WB, SWB and FB) and a wide range of bit rates (5,6-VBR, 7,2 up to 128 kbps). The AMR-WB is included within the EVS as "EVS AMR-WB IO", in short EVS-IO in the present document. EVS-CMR supports seamless transitions between all EVS Primary modes and between EVS Primary and EVS-IO modes during the call, as well as commanding EVS-VBR and EVS-CA modes of operation. Again, as for AMR and AMR-WB, all Codecs in the speech path have to follow the EVS-CMR rules strictly.

As stated above: All EVS Bottom up Configurations are TLCI-compatible to each other and using EVS Bottom up Configurations in SIP/SDP negotiation guarantees always best possible interworking.

A call may be setup with different, TLCI-compatible **EVS Bottom up Configurations** in the path, reflecting different preferences or limitation of the involved UEs and different interworking operators. An important example is EVS (br=5.9-13.2; bw=nb-swb) in CS-UTRAN, EVS (br=5.9-24.4; bw=nb-fb) in CS-CN and EVS (br=5.9-128; bw=nb-fb) in IMS. The common "Framework Bottom up Configuration" in this example is EVS (br=5.9-13.2; bw=nb-swb) and guarantees end-to-end TLCI-compatibility. Mode Control by EVS-CMR during the call may adapt the active EVS Configurations in both directions of the speech path to changing transport conditions by limiting maximum bit rate (and possibly audio bandwidth).

- In case CS-UTRAN is (temporarily or locally) overloaded, it could downgrade to EVS (br=5.9-8; bw=nb-wb).
- In case CS-UTRAN is free of load, it could upgrade to EVS (br=5.9-24.4; bw=nb-fb) for best possible quality. In VoLTE<=>VoLTE calls the full bit rate and bandwidth could be used.

A call may also be setup with different, TLCI-compatible **single audio bandwidth EVS Configurations** in the path, reflecting different preferences or limitation of the involved UEs and different interworking operators. An important example is EVS (br=9.6-13.2; bw=swb) in CS and EVS (br=9.6-128; bw=swb) in IMS. The resulting common "Framework Configuration" is EVS (br=9.6-13.2; bw=swb) and guarantees end-to-end TLCI-compatibility. Mode Control by EVS-CMR during the call may adapt the active EVS Configurations in both directions of the speech path to changing transport conditions by limiting maximum bit rate (keeping the maximum audio bandwidth).

- In case CS-UTRAN is (temporarily or locally) overloaded, rate limitation to 9,6 is allowed.).
- In case CS-UTRAN is free of load, the rate limitation is lifted to 13,2.
- In VoLTE<=>VoLTE calls the full bit rate and swb could be used.

Every EVS implementation includes the EVS AMR-WB IO. Every offered and selected EVS Configuration has to include a parallel EVS-IO Configuration. EVS Primary mode Configurations are strictly speaking not TLCI-compatible to EVS-IO Configurations. However, the EVS Codec allows by design a seamless (i.e. inaudible) transition between both modes of operation. This is important for interworking between EVS and AMR-WB, especially when mid-call modifications occur, like eSRVCC, or other handovers or when mid-call services are invoked. A call may be setup end-to-end with any EVS Configuration and a seamless transition to EVS AMR-WB IO (0,1,2) allows continuation, without transcoding, after a remote eSRVCC to AMR-WB (0,1,2).

11.4.5.2 Additional Conditions for TLCI-Compatibility for EVS

All EVS-Bottom up Configurations are TLCI-compatible to each other. An Initial SIP Offer with an EVS Bottom up Configuration should always find a remote partner that can accept it, unless the other partner explicitly reject EVS Bottom up Configurations. The SDP response may include a smaller EVS Bottom up Configuration. Within such an offered and selected EVS Bottom up Configuration, some single-audio bandwidth EVS Configurations may be "emulated", as long as all additional conditions allow this. Including an EVS-IO Bottom up Configuration in the SIP/SDP negotiation guarantees also TLCI-compatibility to AMR-WB (0,1,2), provided the additional conditions as for AMR-WB are obeyed.

All single audio bandwidth Configurations sharing the same lowest bit rate limit are TLCI-compatible to each other. Similarly to EVS-Bottom up configurations, TLCI-compatibility to AMR-WB(0,1,2) is guaranteed, provided the additional conditions as for AMR-WB are obeyed.

As long as Bottom up Configurations are not widely accepted, an Initial SIP/SDP Offer should not only include punctured EVS Configurations and not only Bottom up Configurations, because the answerer may not be able to accept it. As for AMR and AMR-WB, the RTP packing rules allow a multitude of options. If different RTP packing options are used on both sides of a MGW, then the MGW repacks, but transcoding is not required. In CS Networks, it is mandatory to send one speech frame in one packet.

11.5 Transcoding Less Operation at call setup

Codec Negotiation at call setup tries to ensure that all nodes in the path, including the end terminals, agree on the optimal combination along the voice path, ideally a TLCI-compatible combination of Codecs. As said: these Codecs need not be identical, but it is important that they are TLCI-compatible. This task is no trivial, especially when the call is setup between different networks and these operators follow different strategies or have different historical background and/or different access technologies.

Some overview and discussion is provided in 3GPP S4-150326 "Discussion Paper on Offer-Answer for AMR and AMR-WB". The considerations hold as well for EVS, see also S4-150858 "On Interworking Guidelines for EVS".

11.6 Transcoding Less Operation after Handover

As important as call setup (maybe more) is to consider subsequent handover cases!

Many calls undergo handover in frequencies like one handover in 10 seconds. Often the handovers change also the radio access technology, GERAN<=>UTRAN, LTE<=>WiFi, LTE<=>UTRAN and so on. Especially during network-migration phases it might happen that a new Codec is inserted into the ongoing voice path and this Codec is sometimes not TLCI-compatible to another Codec already in use.

Very often these handover aspects are ignored or forgotten during network design. The current eSRVCC procedure is such an example. Important is also to consider that e.g. after a eSRVCC from LTE to UTRAN a subsequent handover may follow from UTRAN to GERAN or any other combination or sequence. To guarantee end-to-end TLCI in all these (practically infinite) call scenarios requires strict rules for network design and inter-operator and inter-vendor agreements.

12 Enhancements for media and quality aspects

12.1 General

Clause 12 is refers to clause 7, which has been drafted, but is not yet included in this version of the present document.

The identified problems in clause 7 and the discussion in the other clauses lead to the following proposals to achieve significant enhancements for media transport, voice, and communication quality.

12.2 Early Information exchange between MSC and ATCF

12.2.1 Proposed Requirement

Clause 7.2 states: "***Without knowledge about the IMS Selected Codec, the Target RAN Codec cannot be selected optimally***".

Non optimal Target RAN Codec often means transcoding, with noticeable quality loss for the whole duration of the call after eSRVCC. Alternatively, it requires a mid-call modification of the just selected Target RAN Codec immediately after the eSRVCC handover. This would interrupt the voice path a second time, immediately after eSRVCC, unnecessarily.

This leads to the conclusion that the MSC needs to retrieve somehow the necessary information from the ATCF, before the Target RAN Codec is selected.

12.2.2 Proposed Solution 1: CS/IMS Bi-directional Codec List Exchange

12.2.2.1 Overview

Here "solution 1" from 3GPP TR 23.717 [6] is reported in a shortened and modified form:

- Proposal 1:

  "Immediately after it received Message 5, PS to CS Handover Request, the updated SRVCC MSC should send a new message, called "**PS to CS Handover Preparation Request**" to the ATCF. This new message should contain the necessary call identifier, allowing the ATCF finding the concerned call and the wanted call-specific IMS Selected Codec.

The updated ATCF should send a new message, called "**PS to CS Handover Preparation Response**" back to the SRVCC MSC, containing the wanted **IMS Selected Codec**. Now the updated MSC could select the Target RAN Codec in an optimal way and could then continue in the eSRVCC procedure as standardized."

Figure 12.2.2.1-1 shows the essential message flow, where the two new Handover Preparation Messages 5a and 5b are just inserted between Message 5 and Message 7a.

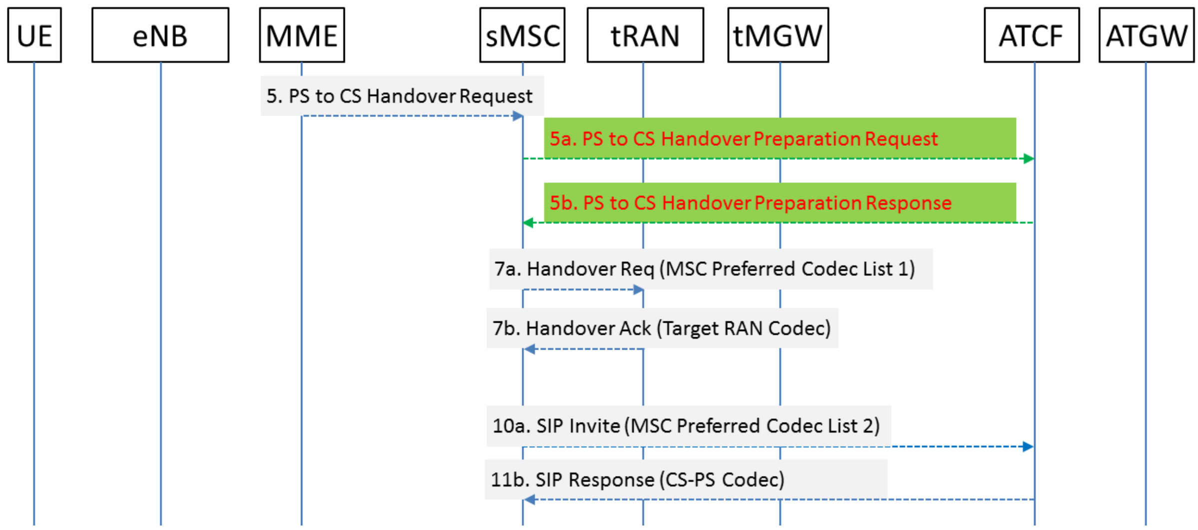

Figure 12.2.2.1-1: Two Handover Preparation messages inserted

These two additional messages between MSC and ATCF would delay the eSRVCC procedure by a minimal, insignificant time span, which would not have any negative influence on the speech path interruption time and no significant effect on the handover success rate.

The message type for this information exchange may be discussed (Stage 2 and Stage 3 work). One simple solution could be to use a tailor made SIP MESSAGE in both directions. The coding of the IMS Selected Codec could follow the SDP description as used in SIP INVITE.

12.2.2.2 Information in Handover Preparation Response

Clause 7.2.2 states ***"If the IMS Selected Codec is better than the Target RAN capabilities, then the SRVCC MSCneeds to be informed about the full IMS Preferred Codec List."***

This leads to a small extension of the **PS to CS Handover Preparation Response**.

- Proposal 2:

    "The ATCF should include the IMS Selected Codec and **alternative Codec candidates** in the so called "**IMS Preferred Codec List**". The usual SDP description as for SIP INVITE could be used. The additional implementation effort would be minimal. The IMS Selected Codec should be on first place in this IMS Preferred Codec List."

If the MSC finds a Target RAN Codec, which is TLCI-compatible to the IMS Selected Codec, then the eSRVCC is optimally prepared and can be completed fast.

EXAMPLE 1: IMS Selected Codec = EVS (br=5.9-24.4; bw=nb-fb)
IMS Preferred Codec List = {EVS (br=5.9-24.4; bw=nb-fb); AMR-WB(), G.722, AMR(0,2,4,7), G.711}.
MSC Supported Codec List = {AMR-WB(0,1,2), AMR(0,2,4,7), G.711, EFR}
==> Target RAN Codec = AMR-WB(0,1,2)
IMS Selected Codec after CMR= EVS-IO(0,1,2), which is TLCI-compatible to AMR-WB-2.

If the MSC does not find a Target RAN Codec, which is TLCI-compatible to the IMS Selected Codec, then transcoding is unavoidable, at least temporarily. The alternative Codecs in the IMS Preferred Codec List would allow the MSC to select the best possible Target RAN Codec that has a TLCI-compatible counterpart in this IMS Preferred Codec List.

EXAMPLE 2: IMS Selected Codec = EVS (br=5.9-24.4; bw=nb-fb)
IMS Preferred Codec List = { EVS (br=5.9-24.4; bw=nb-fb); AMR-WB(), G.722, AMR(0,2,4,7), G.711}.
MSC Supported Codec List = {AMR(7), AMR(0,2,5,7), AMR(0,2,4,7), G.711, EFR}
==> Target RAN Codec = AMR(0,2,4,7).

In example 2, eSRVCC is also completed fast, but leads to Transcoding within the ATGW, with CS-PS Codec = AMR(0,2,4,7). After the call is safely landed in the Target RAN, the ATCF may start a SIP Re-Invite to change IMS Selected Codec and Remote Used Codec to AMR(0,2,4,7). By this SIP Re-Invite TLCI is regained for the rest of the call. This SIP Re-Invite to modify the IMS Selected Codec, better to say: the subsequent modification of the User Plane, may interrupt the voice path as any other handover. This interruption is implementation dependent and it depends on the remote access. Without this small interruption, the call would have to stay in transcoding.

Note that the MSC in example 2 does not know the EVS Codec at all. Sending the IMS Selected Codec alone would not help. The MSC would prefer AMR(7), where no TLCI-compatible counterpart exists on the IMS side.

12.2.2.3 Information in Handover Preparation Request

Clause 7.3 states ***"Pre-SRVCC Mode Control is necessary for the optimal eSRVCC."***

Now, with the new "**PS to CS Handover Preparation Request**" message the ATCF gets early information that eSRVCC is coming. If this new message would include information about the candidates for the Target RAN Codec, then the ATCF could decide, if TLCI would be possible, with which Codec and whether or not Pre-SRVCC Mode Control is required. Therefore:

- Proposal 3:

    "The "**PS to CS Handover Preparation Request**" should contain the full "**MSC Supported Codec List**", meaning the list, from which the Target RAN Codec will be selected. The usual SDP description as for SIP INVITE could be used."

When the ATCF gets this MSC Supported Codec List and compares it with its own IMS Preferred Codec List, then the ATCF could anticipate the Target RAN Codec, before it is selected and allocated by the MSC.

This early knowledge about the Target RAN Codec would allow the ATCF to decide, whether Pre-SRVCC Mode Control should be initiated and which CMR command should be sent to the remote end. The ATCF would then have to inform the ATGW, too. This would not be a command to transfer the access, but just to modify the CMR flow coming from the local LTE UE towards the remote end, preparing the Remote Used Codec for the coming eSRVCC.

In example 1 of the previous clause 12.2.1, this CMR command for EVS-io mode 0 would switch the EVS Codec from the EVS Primary mode into the EVS-IO mode of operation, with the default Initial Codec Mode of AMR-WB-2.

Summary so far:

By simply introducing two new optional messages into the standardized eSRVCC message flow, the selection of the Target RAN Codec could be optimized for all call scenarios. In addition, the ATCF could prepare the Pre-SRVCC Mode Control Command and could trigger the ATGW to send it within the CMR stream towards the remote end. These two new messages between MSC and ATCF would not trigger any resource allocation and not the access transfer.

12.2.3 Proposed Solution 2: MSS initiated codec inquiry

12.2.3.1 Overview

Here "solution 6" from 3GPP TR 23.717 [6] is reported and detailed. It is called in the present document "solution 2".

In this solution 2, the SRVCC MSC queries the codec information from the ATCF, as in solution 1, but with a different message and contents. The ATCF responds with the codec that is currently in use with the ongoing IMS session (i.e. the IMS Selected Codec). The SRVCC MSC can then proceed with the rest of the eSRVCC procedures by reserving the same codec or a compatible one, if such exist, from the Target RAN, in order to achieve e2e TLCI after eSRVCC.

Codec A in the following figure 12.2.3-1 is synonym to the IMS Selected Codec. Note that in this example figure the IMS Selected Codec and the LTE Used Codec are (by coincidence) identical. In general, that is not always the case.

12.2.3.2 Procedures

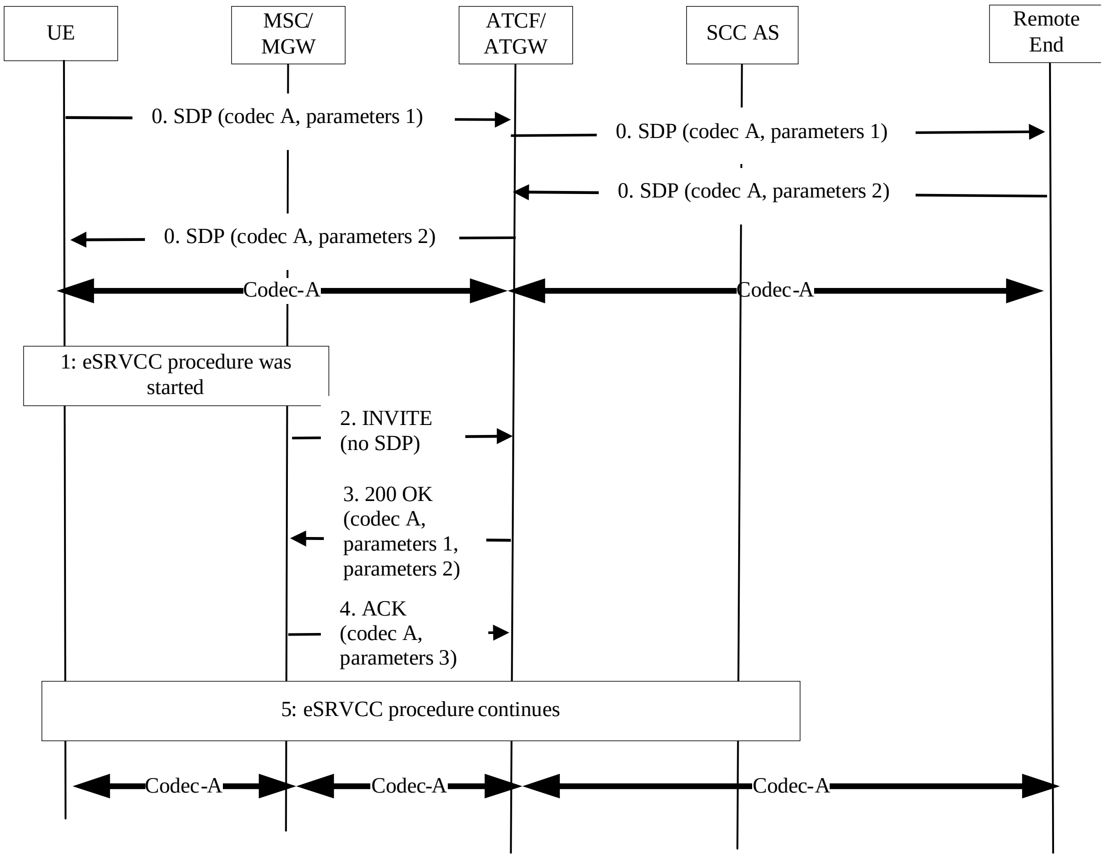

Figure 12.2.3-1: Codec inquiry by MSC server

0. The ongoing IMS session uses codec A as IMS Selected Codec. Parameters for codec A (e.g. the RTP payload type, packetization time, bandwidth information and codec specific parameters, like Codec Configuration, mode-set, etc.) have been negotiated between UE, ATCF and Remote end via a previous SDP offer-answer exchange. Figure 12.2.3.1-1 assumes that Codec A is used as IMS Selected Codec and as LTE Used Codec. Parameters 1 describe codec parameters of Codec A that apply to packets send in the downlink direction. Parameters 2 describe codec parameters of Codec A that apply to packets send in the uplink direction.

1. eSRVCC is started. The SRVCC MSC receives the SRVCC PS to CS request from MME as defined in 3GPP TS 23.216 [3].

2. The MSC server sends a codec query to ATCF. It uses a SIP INVITE without SDP.

NOTE 1: A SIP OPTIONS could possibly also be used for this purpose, but the protocol details are up to CT WG1.

3 ATCF responds with the IMS Selected Codec details that are currently in use with the ongoing IMS session. For instance, it replies to the SIP INVITE without SDP with a SIP 200 OK including an SDP offer with Codec A and related parameters 1 (for downlink direction) in normal SDP and related parameters 2 (for uplink direction) encapsulated within a new SDP attribute.
The ATCF preferably also adds additional codecs it supports for transcoding, or additional codecs it knows to be supported by the remote peer (e.g. not selected codecs that have been received from the remote peer in an SDP offer) as less preferred options into the SDP offer.

4. The SRVCC MSC and the Target RAN support codec A or a TLCI-compatible one. It will send the payload according to parameters 2 and uses parameters 1 to select the payload format and encoding it will expect to receive. If a SIP INVITE without SDP was used in step 2, the SRVCC MSC replies to the SIP 200 OK with a SIP ACK with an SDP answer including parameters 3 that are equivalent to parameters 2.

5. The eSRVCC procedure continues as in TS 23.216 [5]. The SRVCC MSC can use the received codec from ATCF towards the Target RAN in order to reserve the same codec or a compatible one.

NOTE 2: This "solution 2" helps a lot, if the SRVCC MSC knows and supports the IMS Selected Codec or a TLCI-compatible one. In that case, solution 2 is equivalent to solution 1.
Therefore, "solution 2" covers a part of solution 1, but nothing beyond solution 1.

12.2.3.3 Impact on Existing Entities and Interfaces

The MSC server needs to be modified to:

- support the new procedure for codec query towards the ATCF;
- take the received codec information into account when deciding the codec to be used towards the Target RAN.

The ATCF needs to be modified to:

- support the new procedure for codec query from MSC server.

NOTE: These impacts in solution 2 on SRVCC MSC and ATCF are the same as in solution 1.
Beyond that, solution 1 has the option to inform the ATGW in an early stage for minimizing the interruption time; this, however, has also impact on the vertical interface between ATCF and ATGW and on the ATGW of course.
The other aspects, like bi-casting in DL and intelligent combining in UL, are identical options in both solutions.

12.3 Access Transfer and Handover Command

Clause 7.4 states: ***"Prerequisite for minimal speech path interruption during eSRVCC is a successful bi-casting in downlink and intelligent combining in uplink."***

The ATGW may insert the bi-casting in downlink and intelligent combining in uplink immediately, when triggered by Message 10b, Session Transfer (CS-PS Codec).

This could be implemented already today without mandating it in the eSRVCC standard. On the other hand, it would not have its full effect, if the MSC would send Message 13, PS to CS handover Response, too early.

Therefore the following

- Proposal 4:

  "The updated ATGW inserts the bi-casting in downlink and intelligent combining in uplink immediately, when triggered by Message 10b, Session Transfer (CS-PS Codec).
  Due to backward compatibility, it is not required that all ATGWs do this.

The ATCF indicates this updated ATGW-capability already in the **PS to CS Handover Preparation Response** to the MSC.

If the MSC is informed about this updated ATGW-capability, then the MSC sends Message 13, PS to CS handover Response, after the ATCF has send Message 11b, SIP Response (CS-PS Codec), back to the MSC."

In this way, the MSC could rely on this ATGW-capability and the timing of the Handover Command is no longer critical. A small shift in time would just delay the handover on air, but would not have any effect on the speech path interruption time. As long as this shift in time is not too extensive, the handover success rate would not be degraded.

The timing of the handover on air and the handover in the ATGW would be decoupled. **The speech path interruption times, both in uplink and in downlink, would be always minimal** due to the improved ATGW handover handling. Load on network links or in network nodes, as well as radio transmission errors, could still delay the execution of certain actions, but this would not have any influence on the speech break.

Note: sending message 13 later without the proposed updated ATGW handling has not the full effect.

12.4 Unblock the Target MGW in Uplink

Clause 7.5 states that ***"The uplink path in the Target MGW is blocked (is set to one-way downlink-only), until the MSC has received a "Handover Complete" message from the UE."***

This control of the Target MGW is unusual and not necessary. It blocks the uplink speech path in the Target RAN too long and causes an unnecessary uplink interruption. The target base stations have strong error detection mechanisms, allowing differentiating good speech frames in uplink from garbage quite well. These base stations send only valid speech frames uplink and the Target MGW should let them pass immediately. The "Handover Complete" message from the UE is just the confirmation that the handover was successful. After that, the old radio leg can be shut down.

- Proposal 5:

  "Unblock the Target MGW immediately at resource allocation".

12.5 Clarify that it is indispensable to follow CMR commands

Clause 7.6 reports that some UEs are observed not following CMR commands at all or only delayed. This CMR problem is not only an eSRVCC problem; it is a serious misbehaviour in many situations.

- Proposal 6:

  "Clarify in 3GPP TS 26.114 [5] and in IR 92 (and where else it seems appropriate, e.g. in terminal specifications) that it is indispensable that every received CMR is followed as soon as possible, for AMR, AMR-WB and EVS."

NOTE: This is meanwhile clarified in 3GPP TS 26.114 [5] for AMR and AMR-WB.

12.6 Updated Message flow according to proposed solution 1

Figure 12.6-1 shows the essential parts of the updated message flow for eSRVCC (Stage 2 level) with the new actions in ATCF and ATGW.

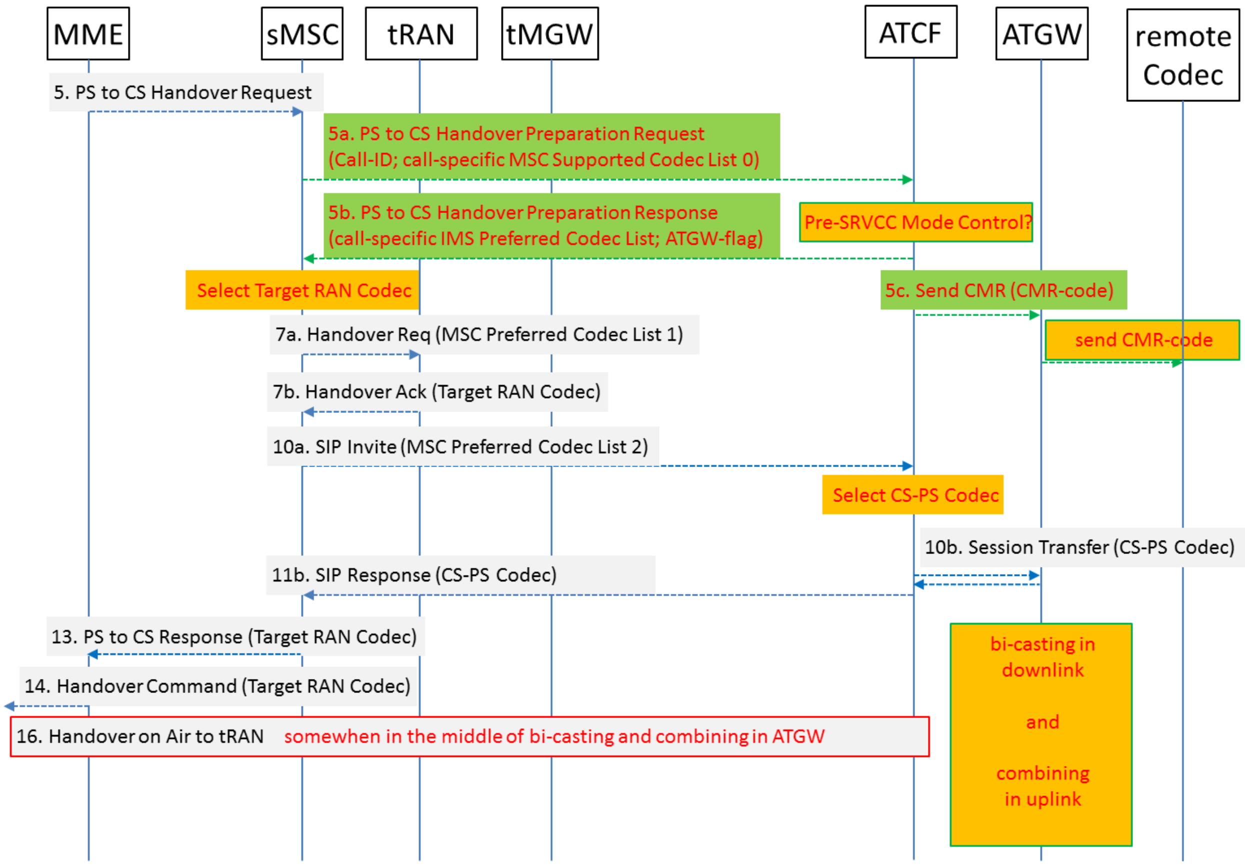

Figure 12.6-1: Essential parts of the updated message flow and new actions in ATCF and ATGW

The essential flow with these new elements is (summary):

1. The MSC informs with Message 5a the ATCF/ATGW at the earliest possible stage about the coming eSRVCC and the candidates for the Target RAN Codec.
2. The ATCF decides, whether Pre-SRVCC Mode Control is requested and triggers the ATGW to send the necessary CMR Command towards the remote end.
3. The ATGW sends these CMR Commands at the earliest possible stage to the remote end to get speech encoded with the new Codec Mode as soon as possible from the remote end; it does not matter, when this new Codec Mode is received at the ATGW and local LTE UE before the handover on air happened.
4. The ATCF sends the complete, call-specific IMS Preferred Codec List to the MSC, indicating, whether the ATGW supports bi-casting.
5. The MSC selects the optimal Target RAN Codec, based on the IMS Preferred Codec List and allocates the Target RAN Resources as usual.
6. Then the MSC sends the SIP INVITE with an updated MSC Preferred Codec List, with the Target RAN Codec on first place, to trigger the access transfer in ATCF and ATGW.
7. The ATCF selects the optimal CS-PS Codec (typically identical or TLCI-compatible to the Target RAN Codec) and allocates the necessary resources in the ATGW.

8. The ATGW starts bi-casting the speech data, coming from the remote end, downlink to the old and new radio access legs and starts intelligent combining of speech data, coming from the old or the new radio access leg in uplink, to forward the result towards the remote end; if necessary transcoding is inserted towards the new radio leg.
9. The ATCF returns the CS-PS Codec to the MSC together with the connectivity data of the ATGW.
10. The MSC closes the link between Target MGW and ATGW.
11. The MSC sends latest now the PS to CS Handover Response to the MME, including the Target RAN Codec, triggering the handover on air.
12. The Target RAN is prepared and the Target MGW sends speech downlink and uplink as available without any blocking.
13. The UE performs the handover on air, while the ATGW is sending and receiving from both radio legs.
No speech frame can be lost, except due to the handover-inherent interruption and the potentially different speech path delays before and after eSRVCC.
14. After the UE has safely landed in the Target RAN, it sends "Handover Complete" to the MSC.
15. The MSC informs the ATCF about the eSRVCC completion.
16. The ATCF shuts down the old radio leg.
17. The ATGW detects autonomously that no more speech is coming in uplink from the old radio leg and speech is only received on the new radio leg and after a certain time out, the ATGW stops bi-casting and combining; alternatively, the ATCF could inform the ATGW.
18. If found appropriate the ATCF may start a SIP Re-Invite towards the remote end to modify the IMS Selected Codec and the Remote Used Codec.

13 Proposals for Stage 2 and Stage 3

The WID on Media and Quality Aspects of SRVCC Enhancements (FS_SETA_S4) states as one objective

- Support SA2 SETA work by SA4 expertise in speech quality and media handling.

Clause 12 lists a (draft) series of enhancement proposals on high level. Comments and additional proposals are invited. SA4 welcomes a close cooperation with SA2 and CT groups to progress this work specifying the details for Stage 2 and Stage 3 specifications.

Annex A: Change history

Change history							
Date	**TSG #**	**TSG Doc.**	**CR**	**Rev**	**Cat**	**Subject/Comment**	**New version**
2015-12	SA#70	SP-150666	-	-		Presented to TSG SA#70 for information	1.0.0
2016-03	SA#71	SP-160080				Presented to TSG SA#71 for approval	2.0.0
2016-03						Approved at TSG SA#71	14.0.0
2016-06	SA#72	SP-160270	0001	4	F	Alignment to approved EVSoCS Specification	14.1.0
2016-09	SA#73	SP-160602	0002	1	F	Naming Replacement for EVSoCS	14.2.0

History

Document history		
V14.2.0	May 2017	Publication