sysmocom

sysmocom - s.f.m.c. GmbH

AMR-RTP in Combination with DTX

by Harald Welte and Max Suraev

DRAFT

Copyright © 2016-2021 sysmocom - s.f.m.c. GmbH

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled "GNU Free Documentation License".

The Asciidoc source code of this manual can be found at http://git.osmocom.org/osmo-gsm-manuals/

HISTORY			
NUMBER	DATE	DESCRIPTION	NAME
DRAFT 1.8.0-4-ga8800	2024-Sep-30		

Contents

1 Full-Rate (TCH/AFS) 1

1.1 TCH/AFS Uplink (MS to Network) . . . 1

1.1.1 TCH/AFS Uplink: Initial Assumptions . . . 1

1.1.2 TCH/AFS Uplink: During Talk-Spurt . . . 1

1.1.3 TCH/AFS Uplink: End of Voice; Start of Silence . . . 1

1.1.4 TCH/AFS Uplink: During Silence . . . 2

1.1.5 TCH/AFS Uplink: End of Silence; Start of Voice . . . 4

1.1.6 TCH/AFS Uplink: Speech Frame Following a SID_FIRST frame . . . 5

1.1.7 TCH/AFS Uplink: FACCH/F Frame During DTX Operation . . . 5

1.2 TCH/AFS Downlink (Network to MS) . . . 6

1.2.1 TCH/AFS Downlink: During Talk-Spurt . . . 6

1.2.2 TCH/AFS Downlink: End of Voice; Start of Silence . . . 7

1.2.3 TCH/AFS Downlink: During Silence . . . 8

1.2.4 TCH/AFS Downlink: End of Silence; Start of Voice . . . 9

1.2.5 TCH/AFS Downlink: Inhibiting a SID_FIRST frame . . . 9

1.2.6 TCH/AFS Downlink: FACCH/F During DTX Operation . . . 10

2 Half-Rate (TCH/AHS) 11

2.1 TCH/AHS Uplink (MS to Network) . . . 11

2.1.1 TCH/AHS Uplink: During Talk-Spurt . . . 11

2.1.2 TCH AHS Uplink: End of Voice; Start of Silence . . . 11

2.1.3 TCH/AFS Uplink: During Silence . . . 12

2.1.4 TCH/AHS Uplink: End of Silence; Start of Voice . . . 13

2.1.5 TCH/AHS Uplink: Inhibited SID_FIRST . . . 14

2.1.6 TCH/AHS Uplink: Inhibited SID_UPDATE . . . 15

2.1.7 TCH/AHS Uplink: FACCH/H During DTX operation . . . 15

2.2 TCH/AHS Downlink (Network to MS) . . . 16

2.2.1 TCH/AHS Downlink: During Talk-Spurt . . . 16

2.2.2 TCH/AHS Downlink: End of Voice; Start of Silence . . . 16

2.2.3 TCH/AHS Downlink: During Silence . . . 17

2.2.4 TCH/AHS Downlink: End of Silence; Start of Voice . . . 18

2.2.5 TCH/AHS Downlink: Inhibited SID_FIRST_P1 . . . 18

2.2.6 TCH/AHS Downlink: Inhibited SID_UPDATE . . . 19

2.2.7 TCH/AHS Downlink: FACCH/H During DTX Operation . . . 19

3 Implementation details 20

The purpose of this document is to describe the sometimes quite intricate interactions between a MS, the BTS-PHY and the BTS software in case of AMR (Adaptive Multi Rate) codec and DTX (Discontinuous Transmission).

It is written with the OsmoBTS implementation and the Nutaq GSM PHY API in mind, but should more or less be applicable to any GSM BTS PHY or any BTS software implementation, assuming it uses RTP on the back-haul towards the MGW.

1 Full-Rate (TCH/AFS)

1.1 TCH/AFS Uplink (MS to Network)

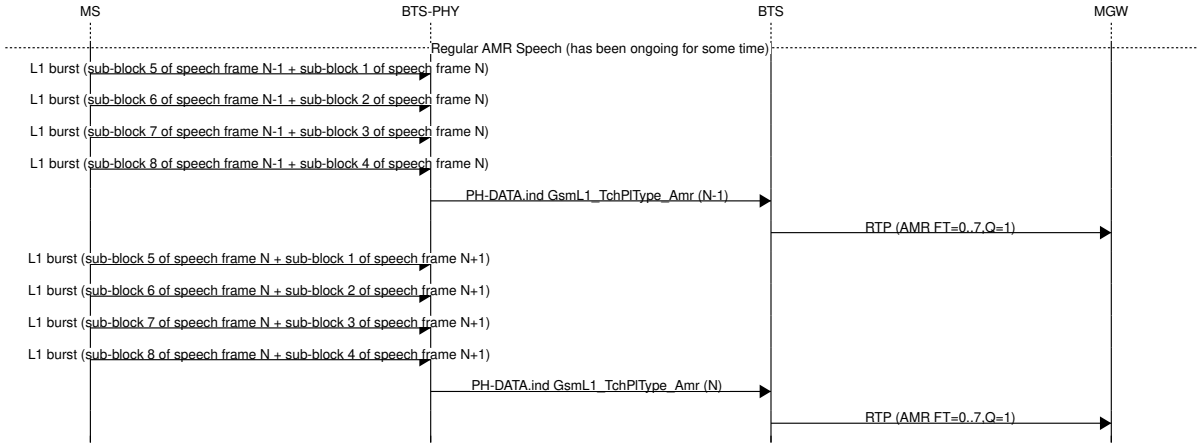
1.1.1 TCH/AFS Uplink: Initial Assumptions

When a call is established, the BTS expects PH-DATA.ind (TCH) events to be triggered by BTS-PHY at regular intervals and none of them being lost, even if the MS is not transmitting at that time due to DTX. This requirement is important as upper layers rely on this timing to update the RTP source clock at a correct pace. Given an event from BTS-PHY is lost, the RTP source clock will drift and this may be seen as steady increase of delay over time from the receiver point of view.

In the case of Nutaq GSM PHY and API, option fBFILevel is set to a ridiculously low value -200 to ensure all indications are notified to upper layers, even if no meaningful payload is available at that time. In this case, an empty payload is delivered from PHY-BTS to BTS, which updates the clock without sending any RTP packet.

1.1.2 TCH/AFS Uplink: During Talk-Spurt

During a talk-spurt, the system behaves identical to a system without DTX enabled: Every four radio bursts, the BTS-PHY has one AMR frame ready and hands it up to the BTS process, which creates an RTP AMR frame from it and sends that to the MGW.

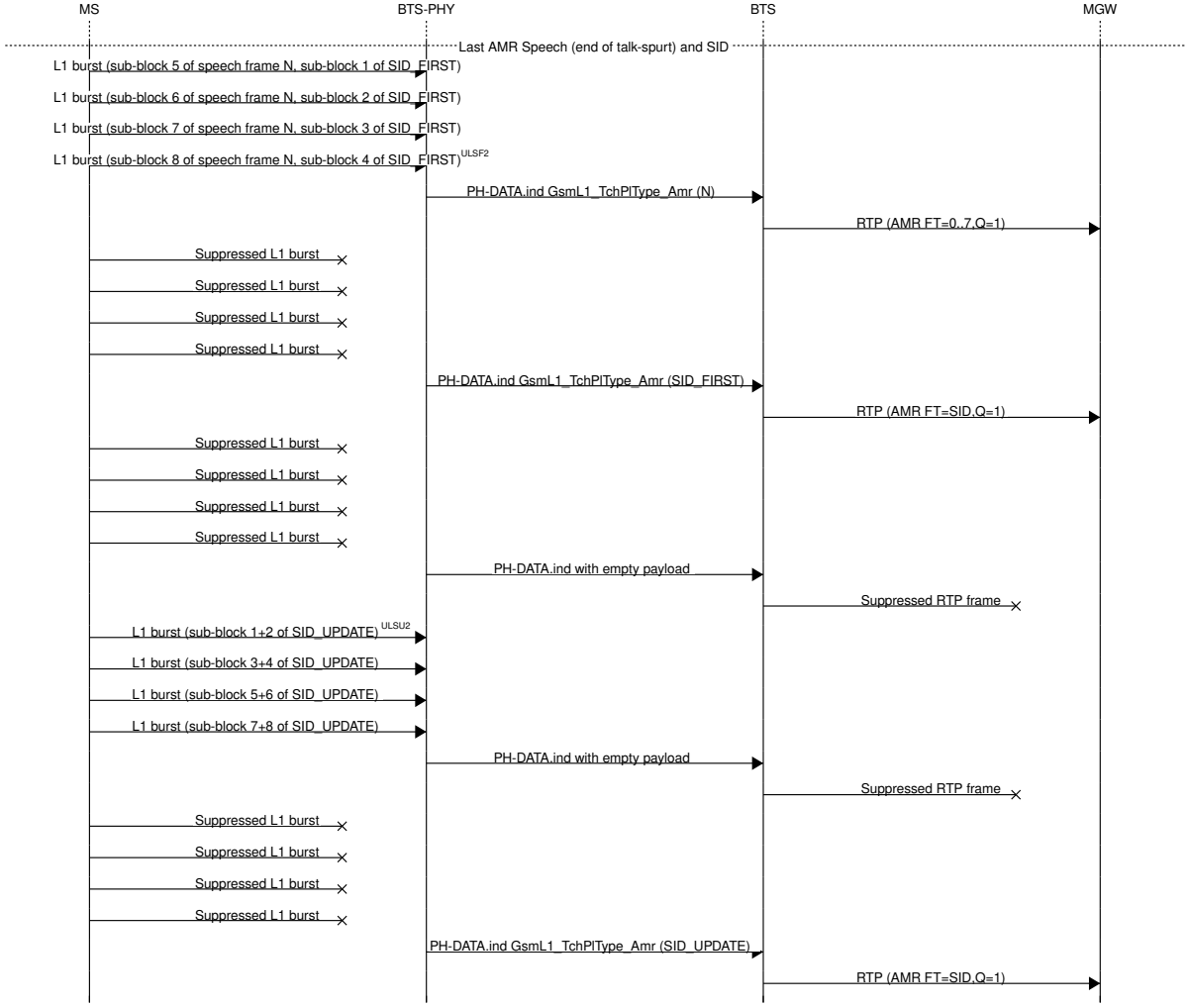


1.1.3 TCH/AFS Uplink: End of Voice; Start of Silence

When the voice encoder in the MS detects no voice activity anymore, it signals towards the MS-PHY that SID_FIRST shall be transmitted.

The BTS-PHY reports the following primitives to the BTS after all four related bursts have been received:

The BTS sends an RTP frame with AMR Frame Type SID, in which the STI is set to indicate a SID_FIRST message.



ULSF2

As per 3GPP TS 05.03 section 3.9.2.4 The last 4 bursts shall not be transmitted unless the SID_FIRST frame is immediately followed by a speech frame. It has been observed that some phone does not transmit the last 4 bursts even if it is not followed by a speech frame.

ULSU2

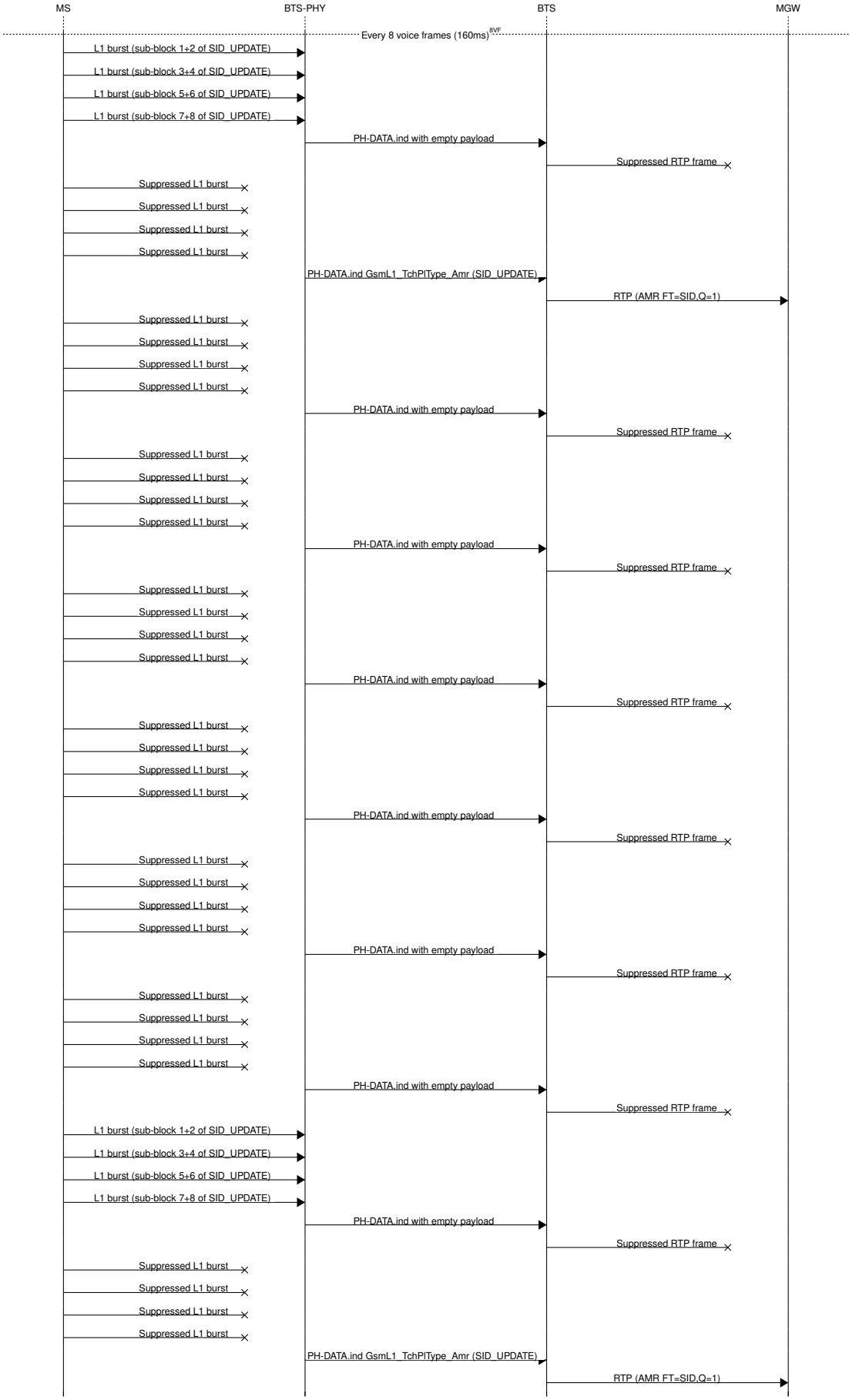
There must be exactly two suppressed voice frames between the SID_FIRST and the SID_UPDATE, i.e. there's 60ms between SID_FIRST and SID_UPDATE.

1.1.4 TCH/AFS Uplink: During Silence

While the period of silence is ongoing, the MS pauses all transmissions, except the periodic scheduling of SID_UPDATE every 8 voice frames (160ms).

NOTE

Silence can also be interrupted at any time by ONSET, see Section 1.1.5.


MS
BTS-PHY
BTS
MGW
Every 8 voice frames (160ms) 8VF
L1 burst (sub-block 1+2 of SID_UPDATE)
L1 burst (sub-block 3+4 of SID_UPDATE)
L1 burst (sub-block 5+6 of SID_UPDATE)
L1 burst (sub-block 7+8 of SID_UPDATE)
PH-DATA.ind with empty payload
Suppressed RTP frame
Suppressed L1 burst
PH-DATA.ind GsmL1_TchPlType_Amr (SID_UPDATE)
RTP (AMR FT=SID,Q=1)

8VF
This happens every 8 **voice frames** (160ms), not every 8 GSM TDMA frames!

1.1.5 TCH/AFS Uplink: End of Silence; Start of Voice

Once the voice encoder in the MS detects voice activity again, it asks its transmitter to perform transmission of SID_ONSET, which is a special frame whose information is encoded only in sub-blocks 3+4, and sub-blocks 1+2 are discarded before transmission.

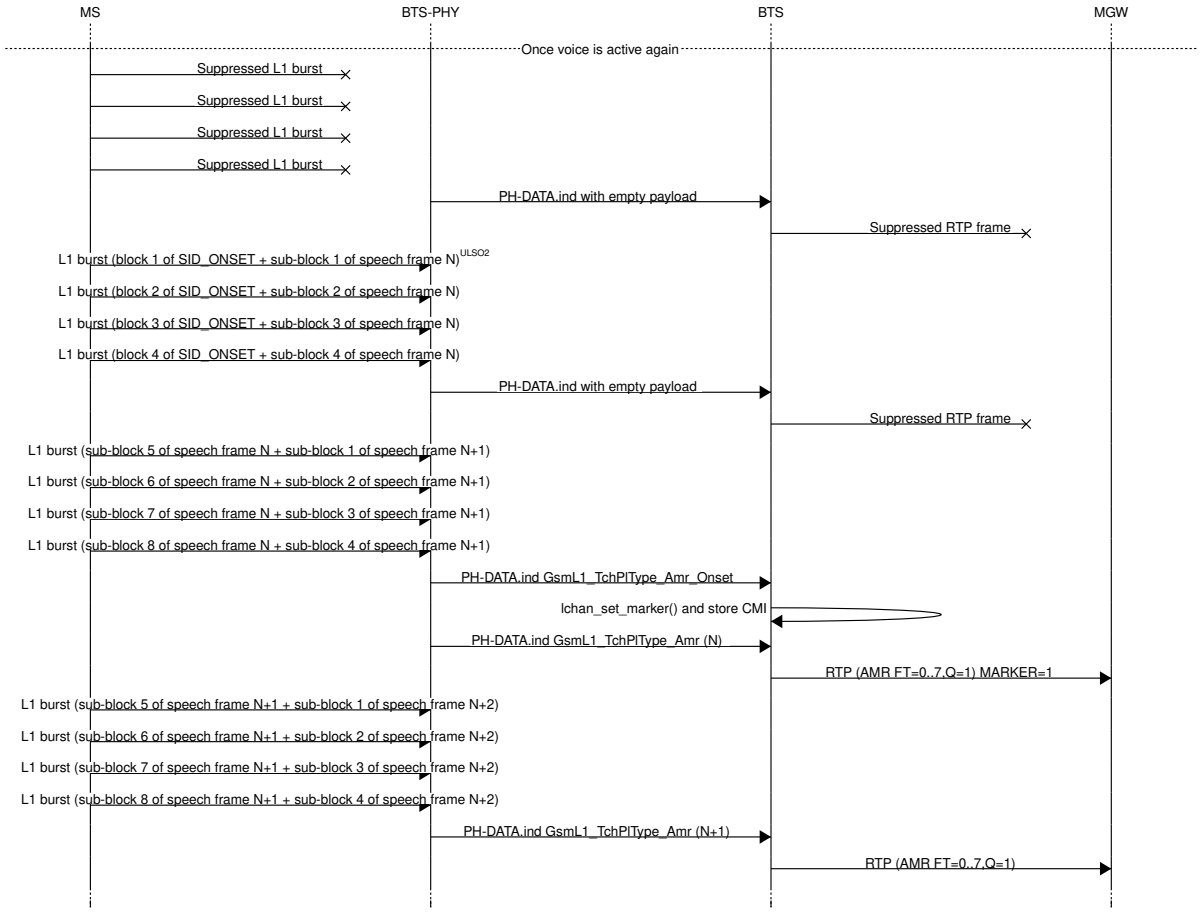
A set of four radio bursts is sent, containing

- the only four transmitted sub-blocks of the SID_ONSET frame
- all four sub-blocks of the first voice codec frame
- the first two blocks of the second voice codec frame

The BTS-PHY informs the BTS using two primitives:

- PH-DATA.ind GsmL1_TchPlType_Amr_Onset indicates the presence of SID_ONSET, including the Channel Mode Indication (irrespective of CMI Phase)
- PH-DATA.ind GsmL1_TchPlType_Amr indicates the first voice frame

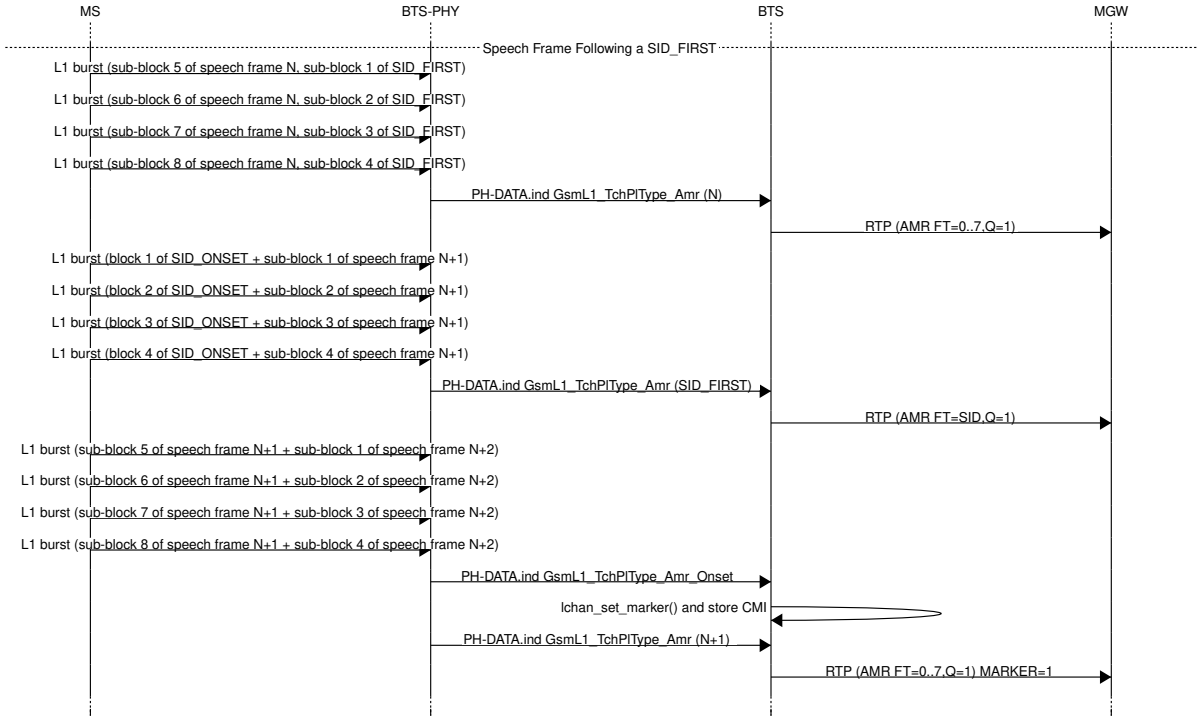
The BTS transmits an RTP frame with AMR payload of the corresponding speech frame type, and sets the RTP MARKER bit to indicate the ONSET condition.



ULSO2
sub-blocks 1..4 of SID_ONSET are never transmitted as all information is contained in blocks 5..8.

1.1.6 TCH/AFS Uplink: Speech Frame Following a SID_FIRST frame

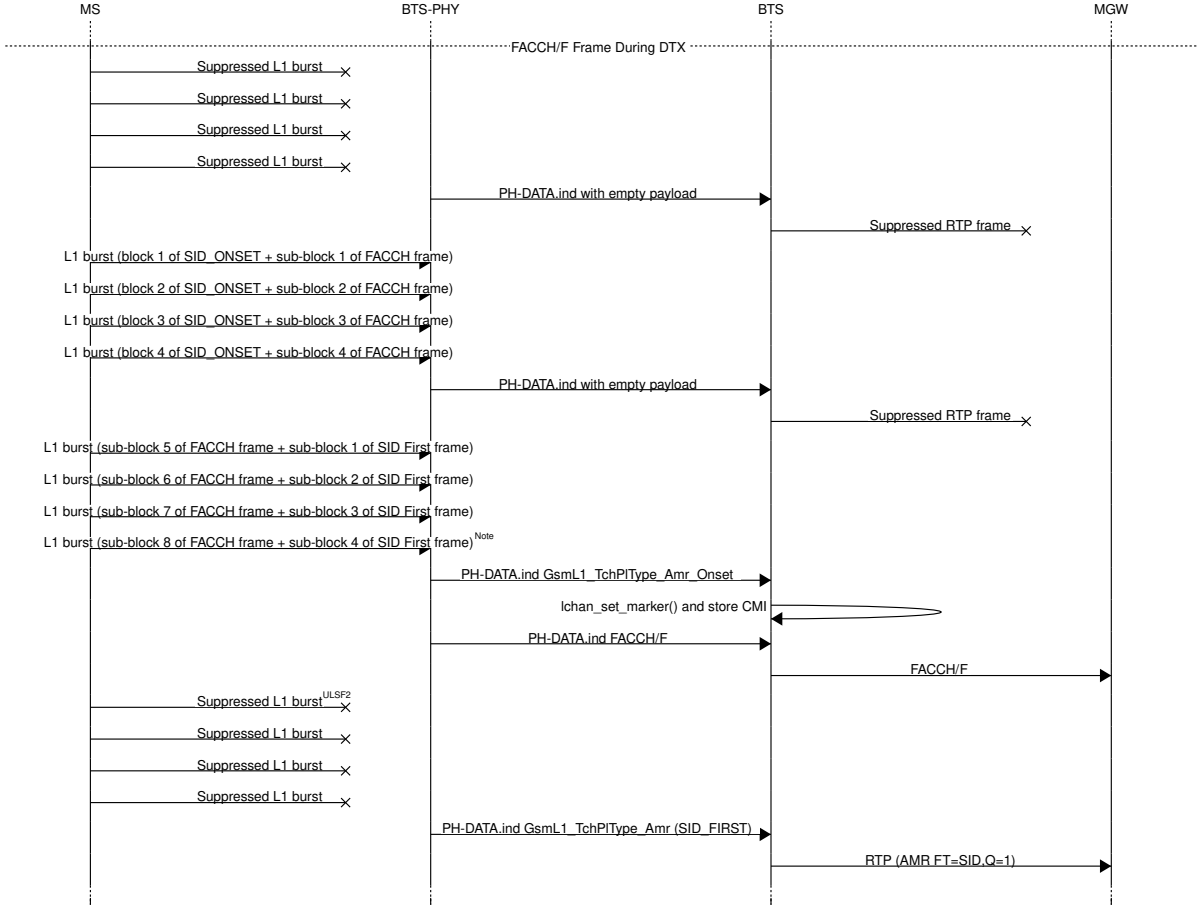
The four last bursts of a SID_FIRST frame can be replaced by an ONSET frame in order to quickly resume speech.



1.1.7 TCH/AFS Uplink: FACCH/F Frame During DTX Operation

As mentioned in section A.5.1.2.1 of 3GPP TS 26.093 :

- If the frame preceding the FACCH frame is not of TX_TYPE=*SPEECH_GOOD*, then an ONSET frame shall be signalled to the CHE, followed by the FACCH frame(s).
- If the frame following the FACCH frame is not of TX_TYPE=*SPEECH_GOOD*, then a SID_FIRST shall be signalled to the CHE.



ULSF2
The sub-blocks 5-8 of SID_FIRST are not transmitted, as all information bits are contained in sub-blocks 1-4 only

Note
It has been observed with some phones that the SID_FIRST is not sent following the FACCH/F frame. If this case occurs the No Data Frame and SID_UPDATE order resumes.

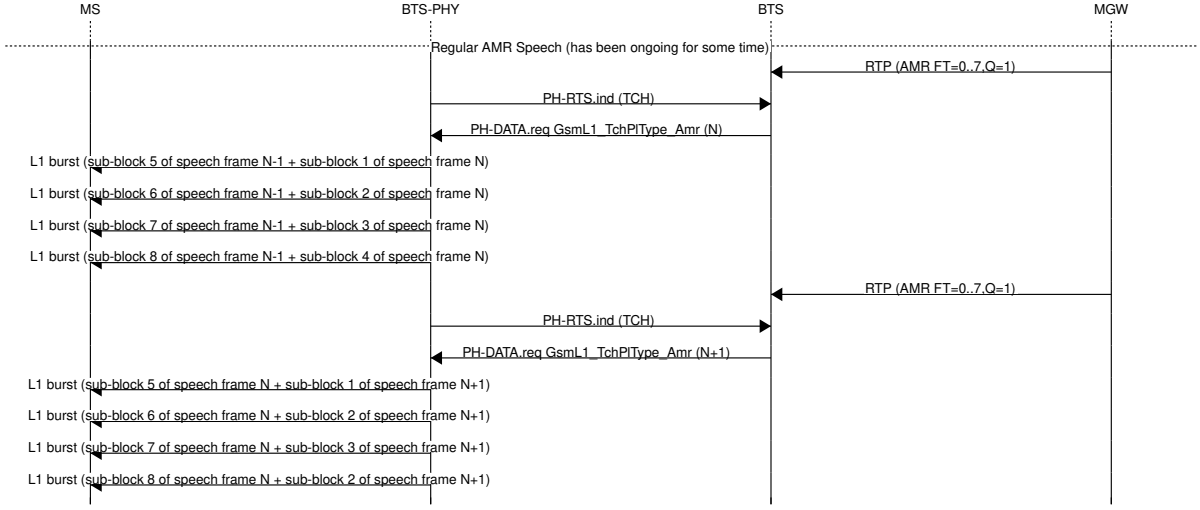
1.2 TCH/AFS Downlink (Network to MS)

1.2.1 TCH/AFS Downlink: During Talk-Spurt

During a talk-spurt, the system behaves identical to a system without DTX enabled: an RTP frame arrives every 20ms.

The PHY sends a PH-RTS.ind in similar intervals, to which the BTS responds with a PH-DATA.req containing GsmL1_TchPlType_Amr.

The BTS-PHY then encodes and interleaves the codec frame into eight sub-blocks. Due to the interleaving, one new PH-RTS.ind is issued every four bursts.



1.2.2 TCH/AFS Downlink: End of Voice; Start of Silence

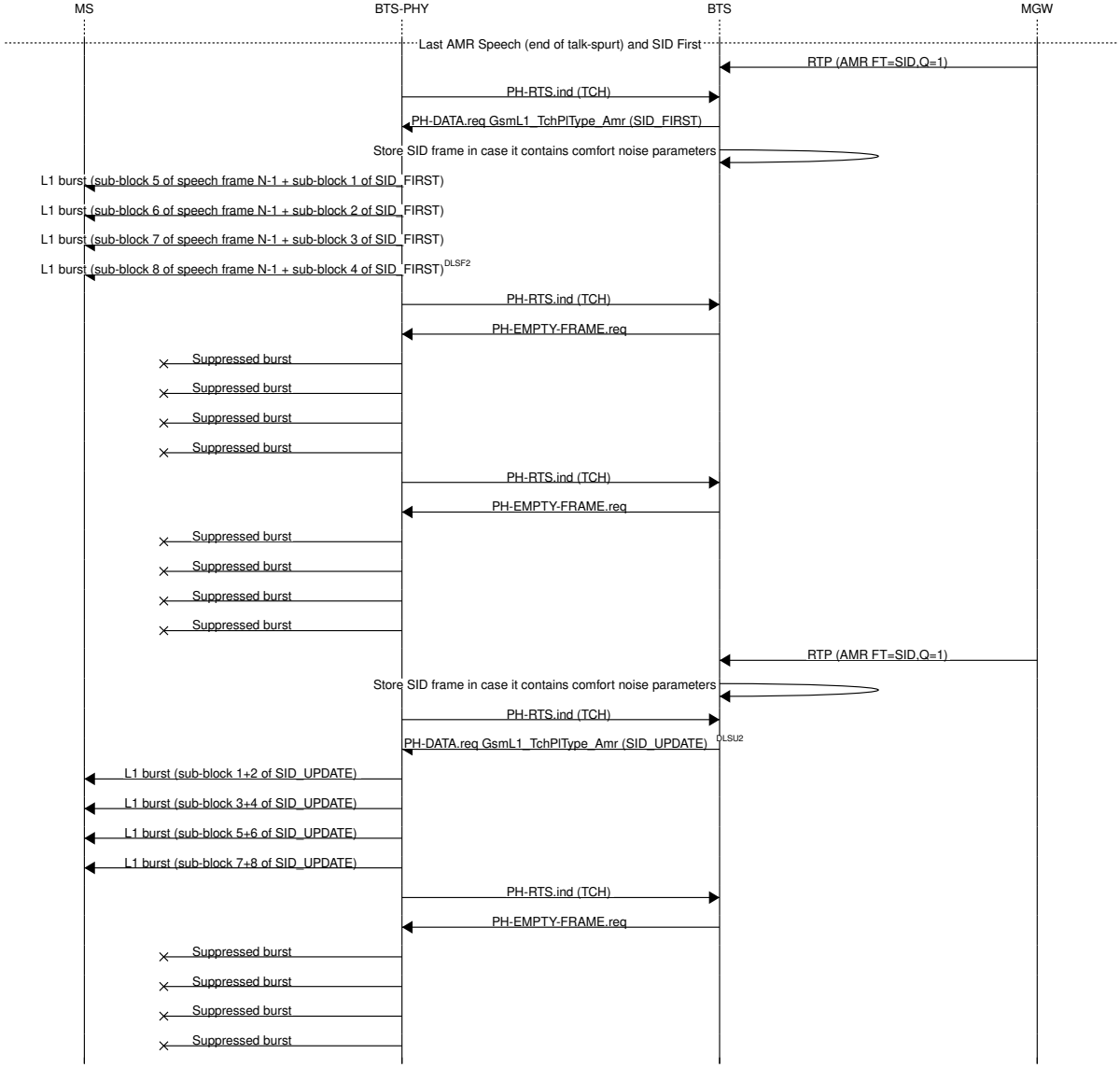
When the BTS receives the first RTP frame with Frame Type SID, it will generate a SID_FIRST AMR frame. The AMR frame is interleaved in a way that all information is contained in the first four sub-blocks, with the latter four sub-blocks being dropped and not transmitted.

Three codec frames (60ms) later, the BTS needs to transmit a SID_UPDATE AMR frame, which should consist of comfort noise parameters received in either the first AMR SID frame, or a subsequent AMR SID frame received meanwhile.

In between SID_FIRST and SID_UPDATE, and after the SID_UPDATE, the BTS sends PH-EMPTY-FRAME.req to all PH-RTS.ind, causing the BTS-PHY to cease transmission in those periods.

NOTE

At any time, the BTS can end the silence period by issuing ONSET (in case of a new downlink talk-spurt or a FACCH downlink frame). See Section 1.2.4.



DLSF2
sub-frames 5..8 of SID_FIRST are not transmitted, as all information is contained in sub-frames 1..4

DLSU2
The SID_UPDATE must be sent exactly three voice frames (60ms) after the SID_FIRST, resulting in two suppressed voice frame periods of empty bursts in-between.

1.2.3 TCH/AFS Downlink: During Silence

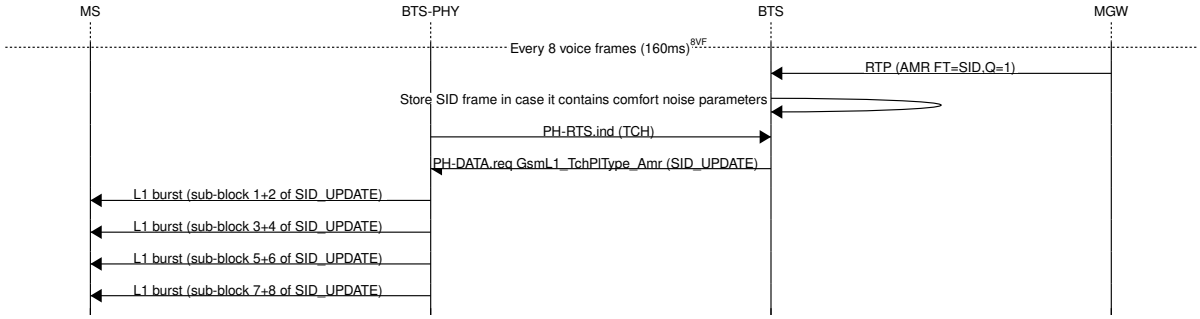
During Silence periods, the BTS may at any time receive RTP AMR SID frames, and keep a copy of the last received frame around.

Every eight voice frames (160ms), the BTS shall respond to the PH-RTS.ind with a PH-DATA.req containing a GsmL1_TchPlType_Amr with SID_UPDATE frame.

At all other times, the BTS sends PH-EMPTY-FRAME.req to any received PH-RTS.ind, causing the BTS-PHY to cease transmission in those periods.

NOTE

At any time, the BTS can end the silence period by issuing ONSET (in case of a new downlink talk-spurt or a FACCH downlink frame). See Section 1.2.4.



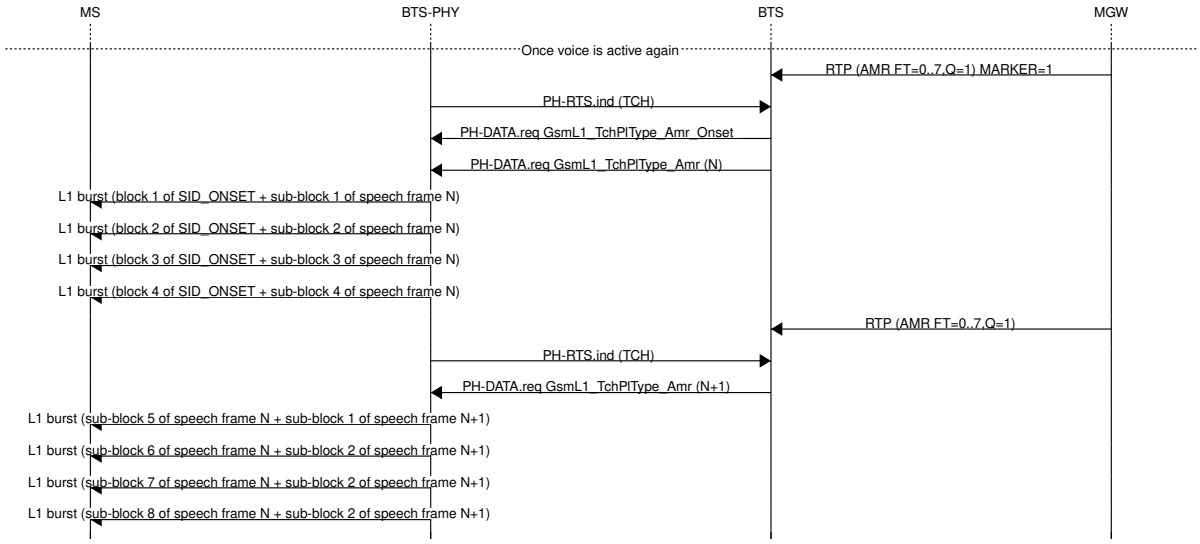
8VF

This happens every 8 **voice frames** (160ms), not every 8 GSM TDMA frames!

1.2.4 TCH/AFS Downlink: End of Silence; Start of Voice

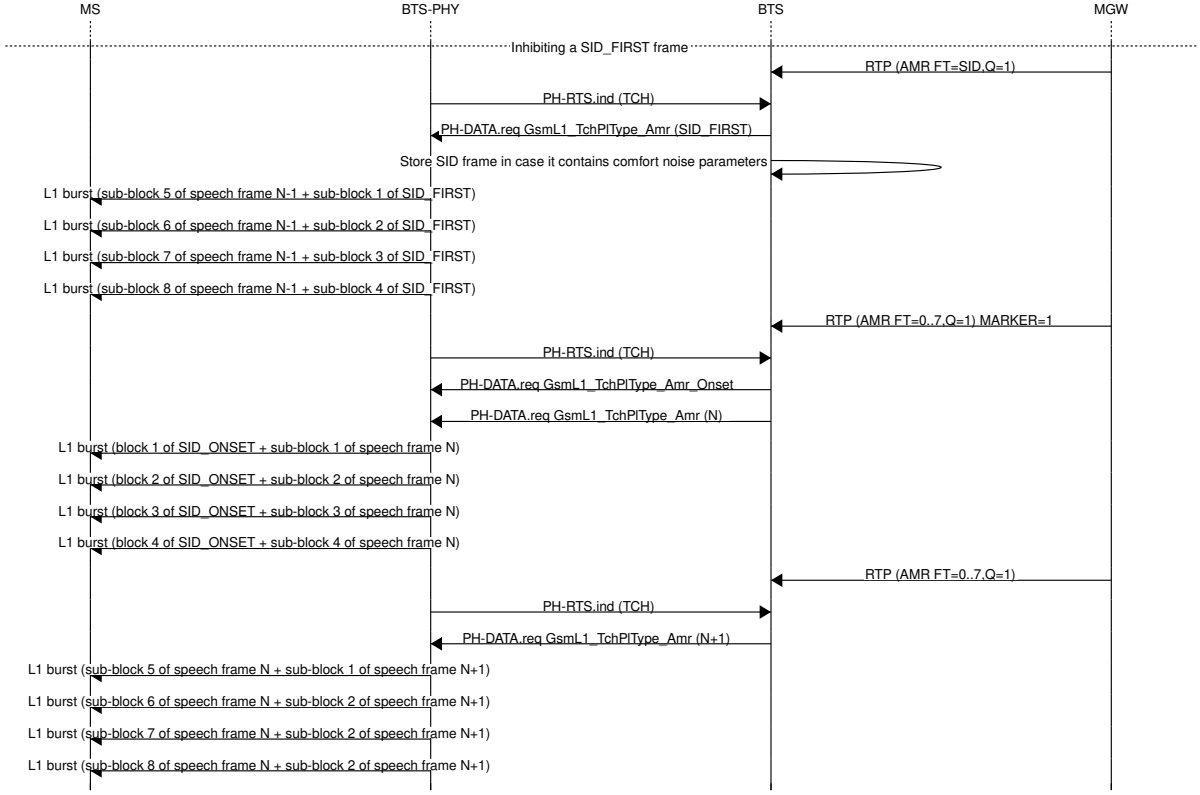
Once the BTS receives a non-SID AMR RTP frame (which should also have the MARKER bit set to 1, but let's not take that for granted), the contained AMR voice data is passed to the BTS-PHY in the next PH-DATA.req (GsmL1_TchPlType_Amr_Onset).

From that point onwards, regular transmission resumes, see Section 1.2.1



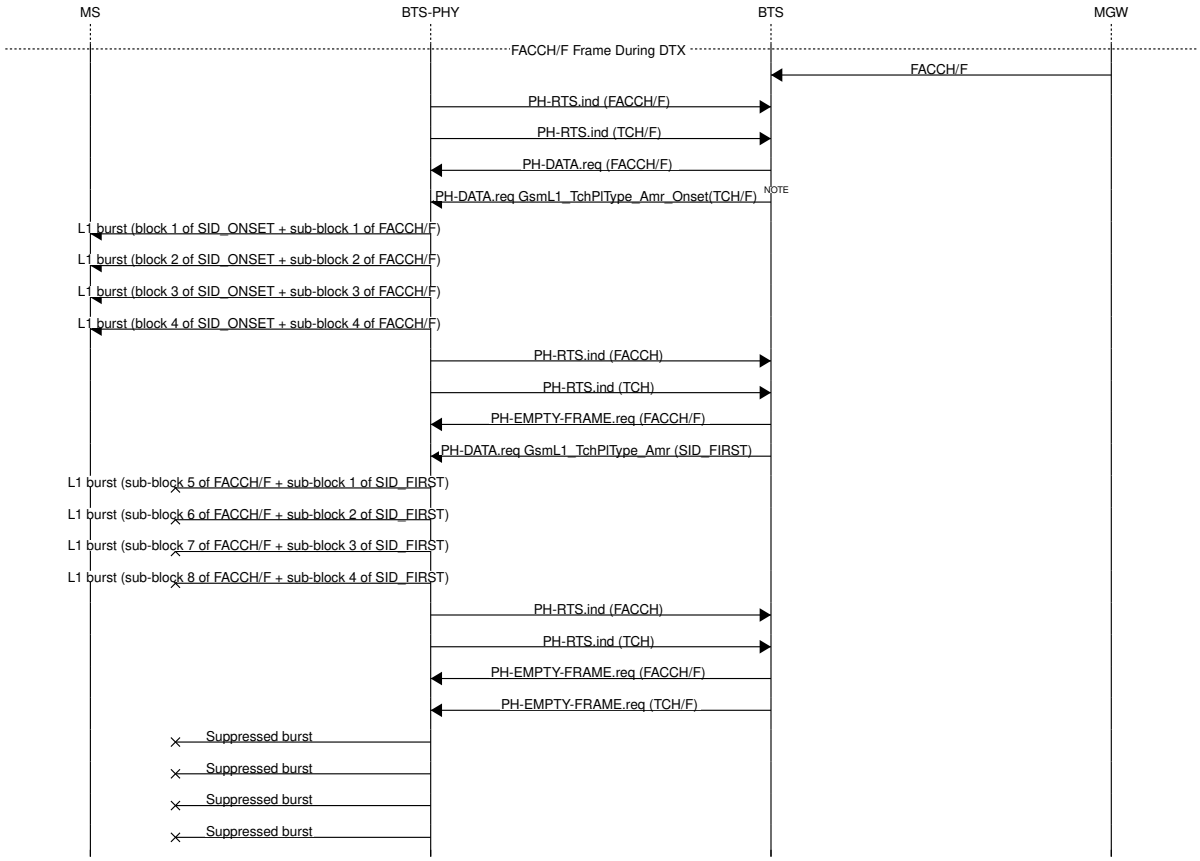
1.2.5 TCH/AFS Downlink: Inhibiting a SID_FIRST frame

Here is the procedure to inhibit a SID_FIRST frame in order to quickly resume speech.



1.2.6 TCH/AFS Downlink: FACCH/F During DTX Operation

The following procedure must be observed if a FACCH/F frame must be transmitted during DTX operation.



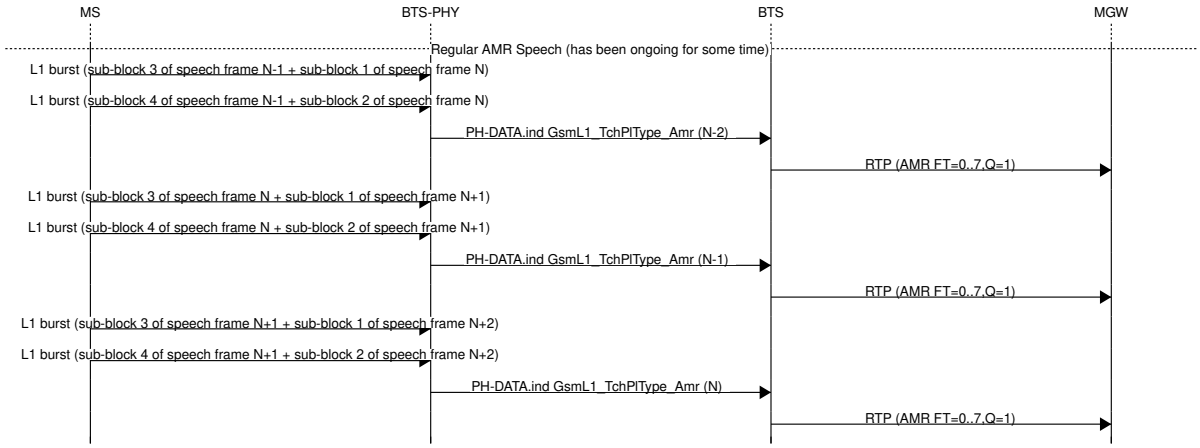
NOTE
The ONSET and the FACCH/F PH-DATA requests must both be sent to the PHY.

2 Half-Rate (TCH/AHS)

2.1 TCH/AHS Uplink (MS to Network)

2.1.1 TCH/AHS Uplink: During Talk-Spurt

During a talk-spurt, the system behaves identical to a system without DTX enabled: Every two radio bursts, the BTS-PHY has one AMR frame ready and hands it up to the BTS process, which creates an RTP AMR frame from it and sends that to the MGW.



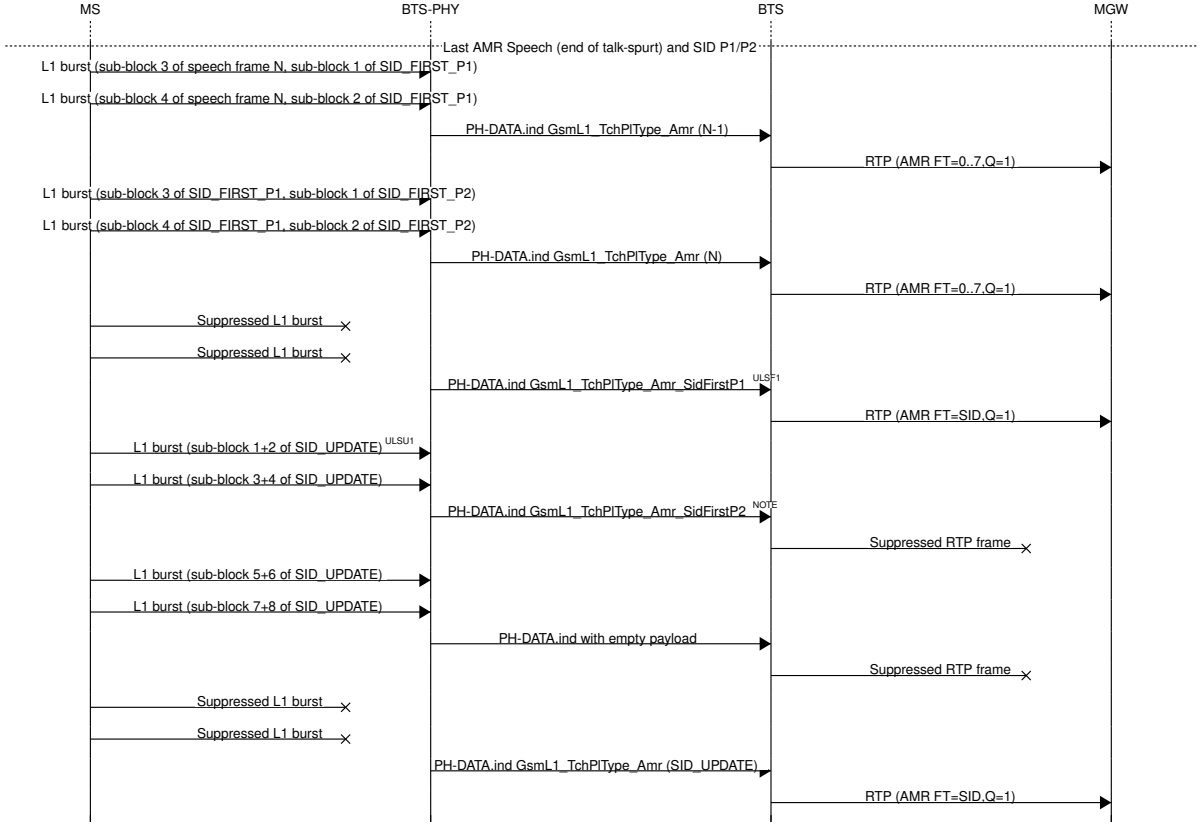
2.1.2 TCH AHS Uplink: End of Voice; Start of Silence

When the voice encoder in the MS detects no voice activity anymore, it signals towards the MS-PHY that SID_FIRST_P1 and SID_FIRST_P2 shall be transmitted.

The BTS-PHY reports the following primitives to the BTS after all four related bursts have been received:

- PH-DATA.ind GsmL1_TchPlType_Amr_SidFirstP1
- PH-DATA.ind GsmL1_TchPlType_Amr_SidFirstP2

The BTS sends an RTP frame with AMR Frame Type SID, in which the STI is set to indicate a SID_FIRST message.



ULSF1
Only SID_FIRST_P1 contains information so it must be the only one transmitted over RTP.

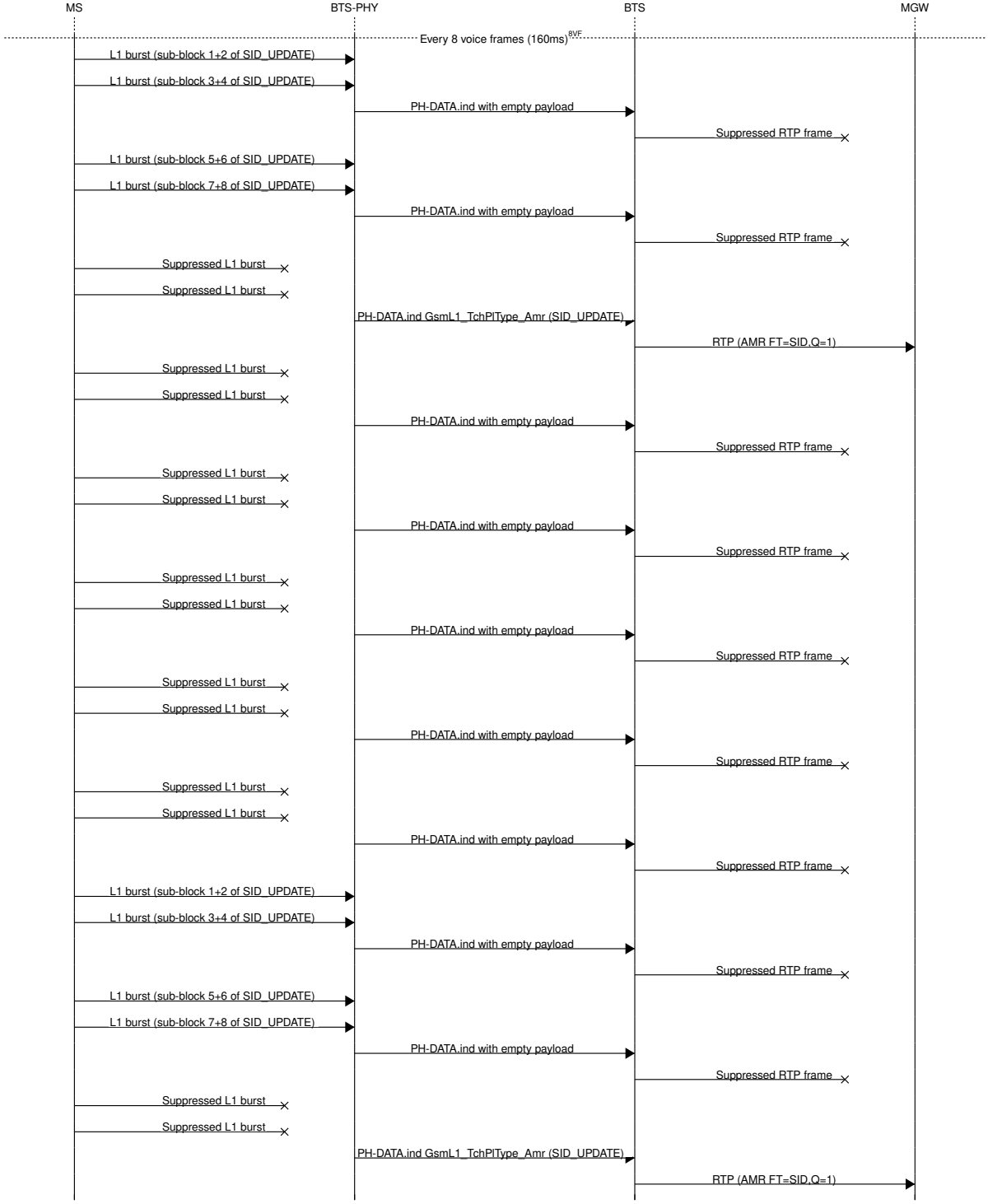
NOTE
It has been observed that not all phones transmit SID_FIRST_P2 so the PH-DATA.ind GsmL1_TchPlType_Amr_SidFirstP2 is not guaranteed to be sent to the BTS.

ULSU1
There must be exactly two suppressed voice frames between the SID_FIRST and the SID_UPDATE, i.e. there's 60ms between SID_FIRST and SID_UPDATE.

2.1.3 TCH/AFS Uplink: During Silence

While the period of silence is ongoing, the MS pauses all transmissions, except the periodic scheduling of SID_UPDATE every 8 voice frames (160ms).

NOTE
Silence can also be interrupted at any time by ONSET, see Section 2.1.4.



8VF

This happens every 8 **voice frames** (160ms), not every 8 GSM TDMA frames!

2.1.4 TCH/AHS Uplink: End of Silence; Start of Voice

Once the voice encoder in the MS detects voice activity again, it asks its transmitter to perform transmission of SID_ONSET, which is a special frame which has information encoded only in sub-blocks 3+4, and sub-blocks 1+2 are discarded before transmission.

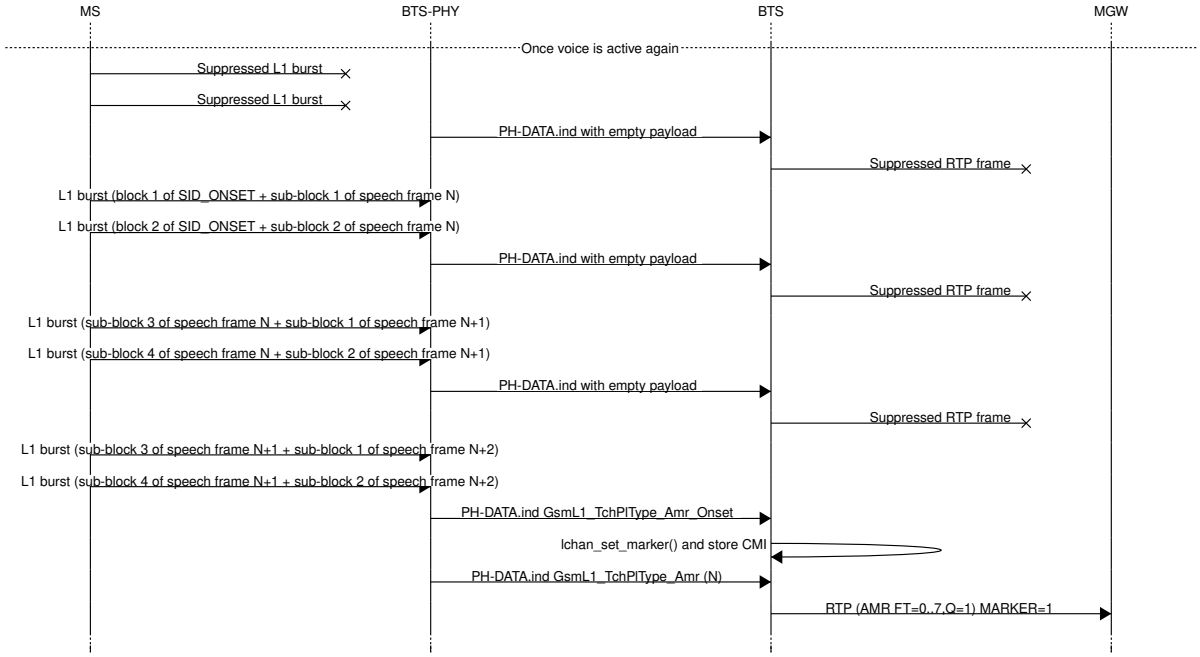
A set of four radio bursts is sent, containing

- the only two transmitted sub-blocks of the SID_ONSET frame
- all four sub-blocks of the first voice codec frame
- the first two blocks of the second voice codec frame

The BTS-PHY informs the BTS using two primitives:

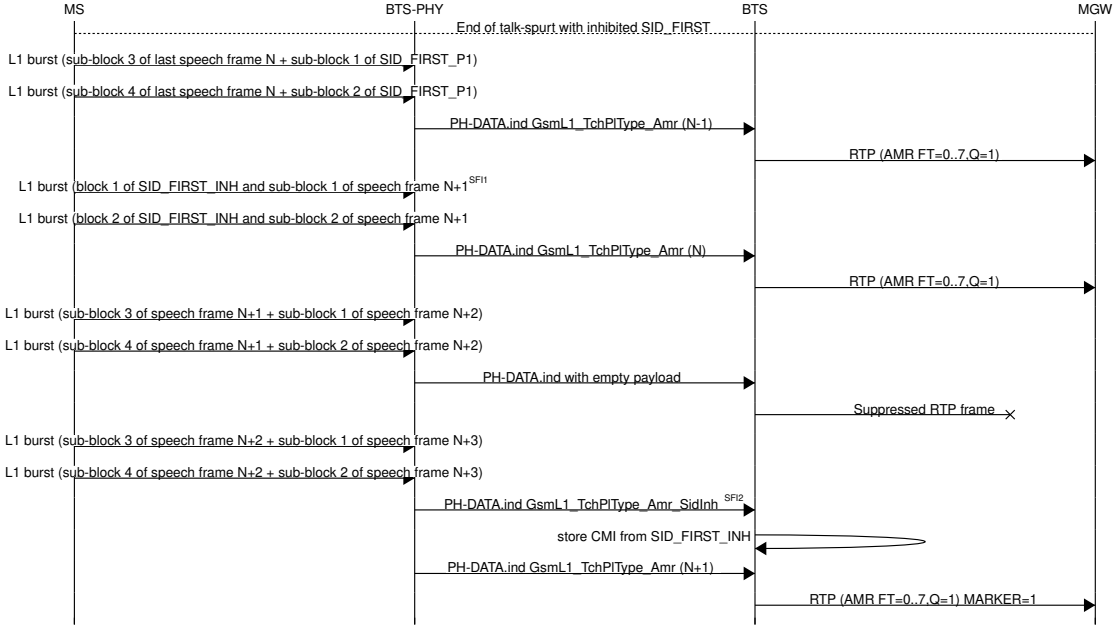
- PH-DATA.ind GsmL1_TchPlType_Amr_Onset indicates the presence of SID_ONSET, including the Channel Mode Indication (irrespective of CMI Phase)
- PH-DATA.ind GsmL1_TchPlType_Amr indicates the first voice frame

The BTS transmits a RTP frame with AMR payload of the corresponding speech frame type, and sets the RTP MARKER bit to indicate the ONSET condition.



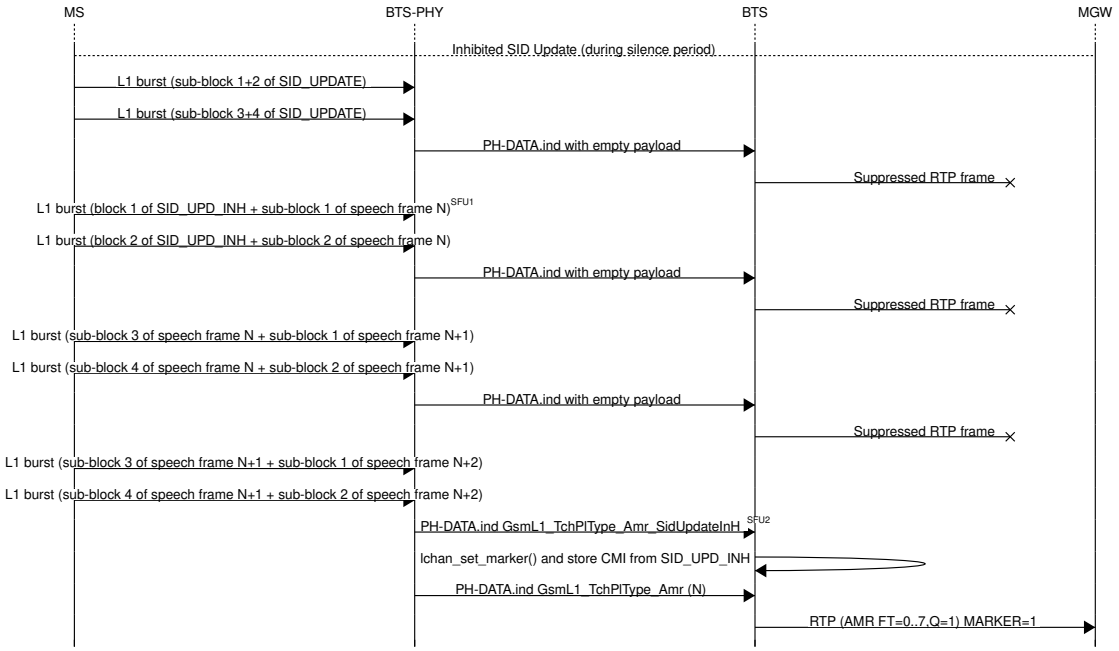
2.1.5 TCH/AHS Uplink: Inhibited SID_FIRST

In case voice activity is detected again while the SID_FIRST_P1 transmission is ongoing or completed, but the SID_FIRST_P2 has not been transmitted yet, SID_FIRST can be inhibited by means of a SID_FIRST_INH frame. This allows the first voice frame to be transmitted with minimal delay, compared to first completing the regular SID_FIRTS_P2 and SID_ONSET procedure.



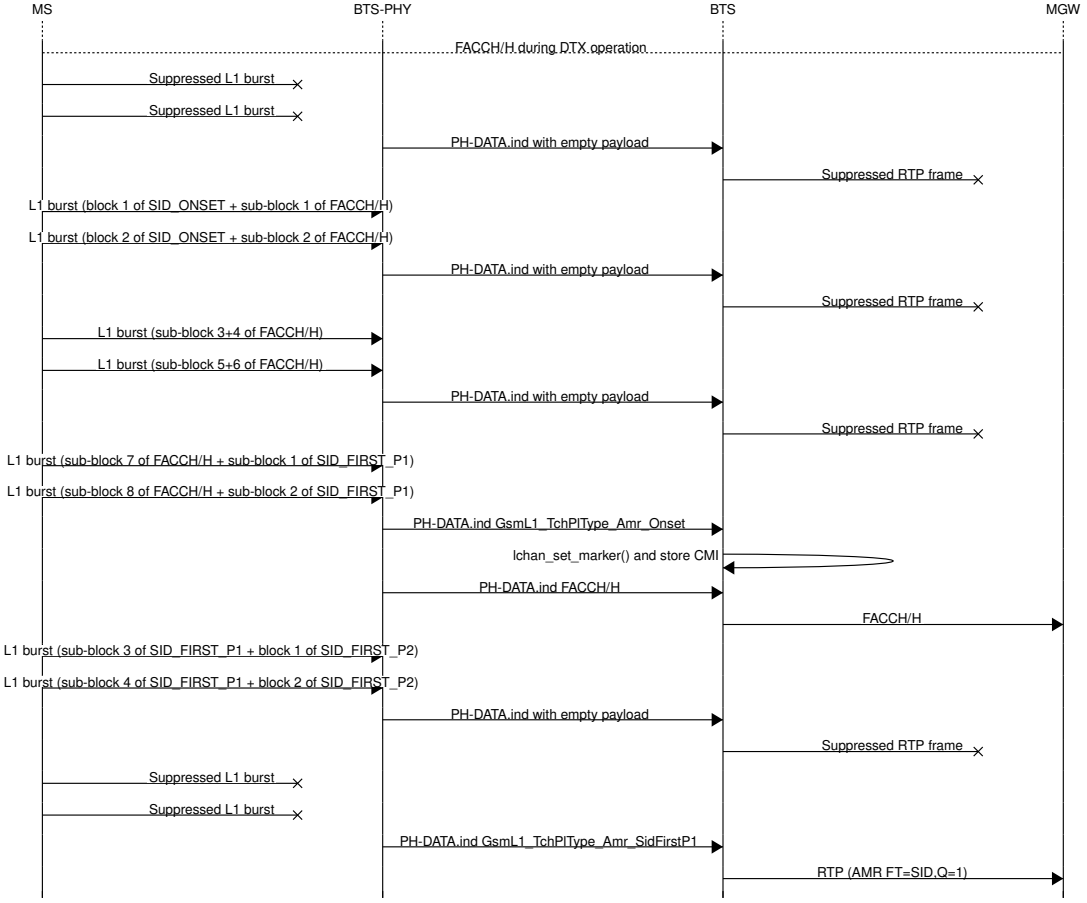
2.1.6 TCH/AHS Uplink: Inhibited SID_UPDATE

In case voice activity is detected again while the SID_UPDATE transmission of the first two sub-blocks is ongoing or completed, but the second two sub-blocks have not been transmitted yet, SID_UPDATE can be inhibited by means of a SID_UPDATE_INH frame. This allows the first voice frame to be transmitted with minimal delay, compared to first completing the regular SID_UPDATE and SID_ONSET procedure.



2.1.7 TCH/AHS Uplink: FACCH/H During DTX operation

The following procedure must be observed if a FACCH/H frame must be transmitted during DTX operation.



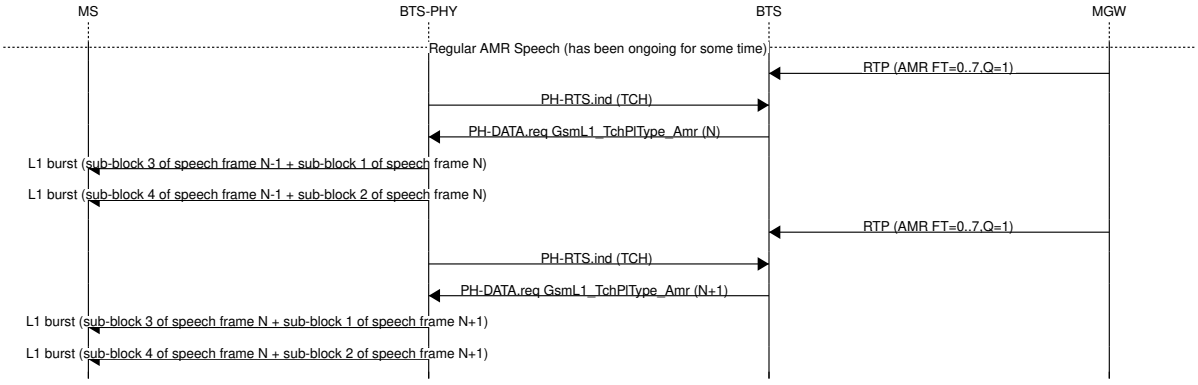
2.2 TCH/AHS Downlink (Network to MS)

2.2.1 TCH/AHS Downlink: During Talk-Spurt

During a talk-spurt, the system behaves identically to a system without DTX enabled: an RTP frame arrives every 20ms.

The PHY sends a PH-RTS.ind in similar intervals, to which the BTS responds with a PH-DATA.req containing GsmL1_TchPlType_Amr.

The BTS-PHY then encodes and interleaves the codec frame into four sub-blocks. Due to the interleaving, one new PH-RTS.ind is issued every two bursts.



2.2.2 TCH/AHS Downlink: End of Voice; Start of Silence

When the BTS receives the first RTP frame with Frame Type SID, it will first issue a GsmL1_TchPlType_Amr_SidFirstP1 primitive towards the BTS-PHY, followed by a GsmL1_TchPlType_Amr_SidFirstP2 primitive.

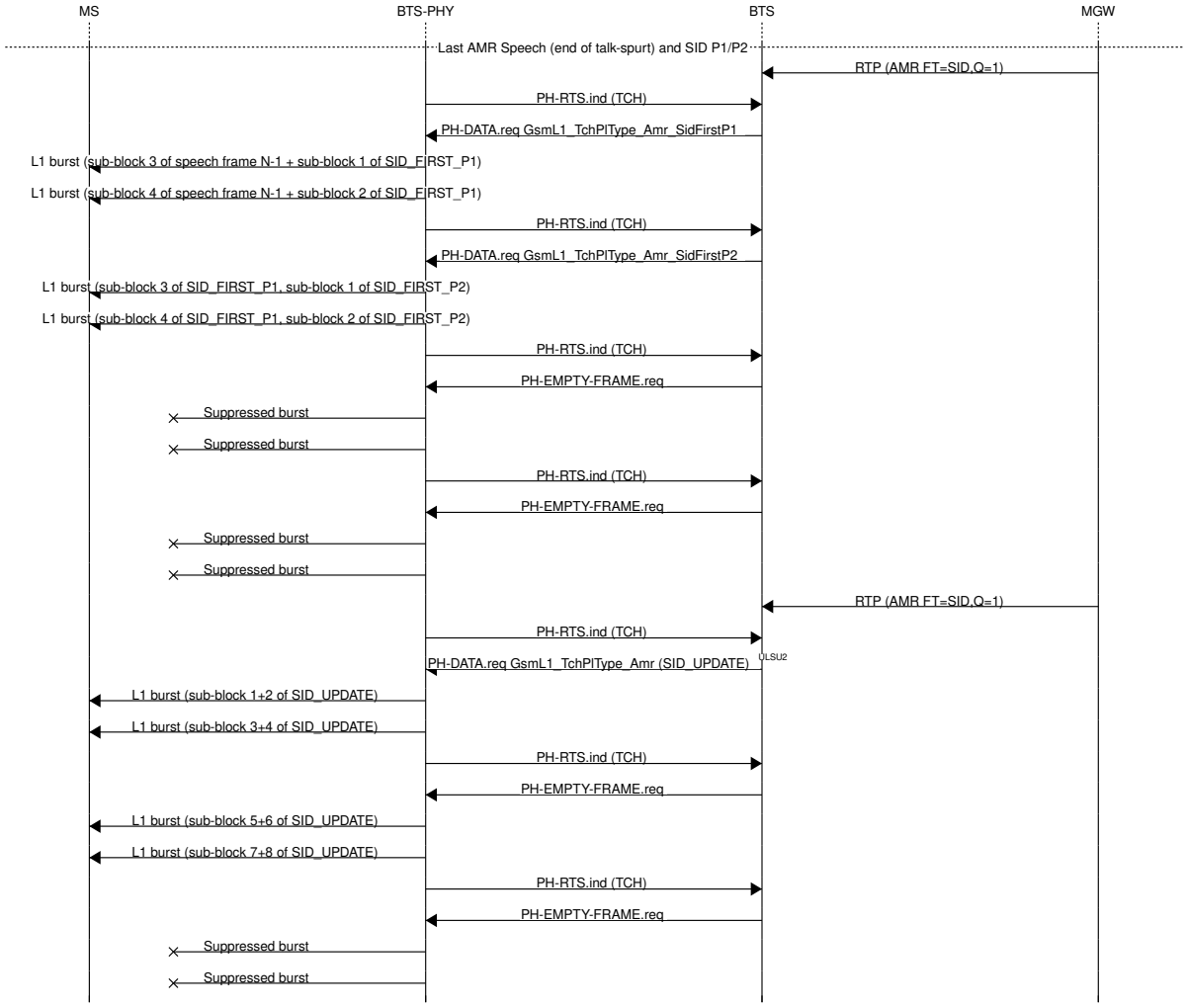
The SID_FIRST_P2 is interleaved in a way that all information is contained in the first two sub-blocks, with the latter two sub-blocks being dropped and not transmitted.

Three codec frames (60ms) later, the BTS needs to transmit a SID_UPDATE AMR frame, which should consist of comfort noise parameters received in either the first RTP AMR SID frame, or a subsequent RTP AMR SID frame received meanwhile.

In between SID_FIRST_P2 and SID_UPDATE, and after the SID_UPDATE, the BTS sends PH-EMPTY-FRAME.req to all PH-RTS.ind, causing the BTS-PHY to cease transmission in those periods.

NOTE

At any time, the BTS can end the silence period by issuing ONSET (in case of a new downlink talk-spurt or a FACCH downlink frame). See Section Section 2.2.4.



ULSU2

The SID_UPDATE must be sent exactly three voice frames (60ms) after the SID_FIRST, resulting in two suppressed voice frame periods of empty bursts in between.

2.2.3 TCH/AHS Downlink: During Silence

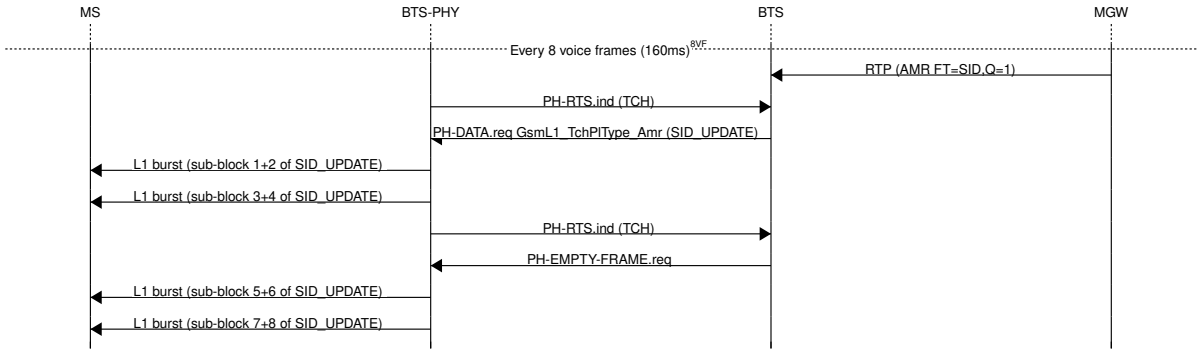
During Silence periods, the BTS may at any time receive RTP AMR SID frames and keep a copy of the last received frame around.

Every eight voice frames (160ms), the BTS shall respond to the PH-RTS.ind with a PH-DATA.req containing a GsmL1_TchPlType_Amr with SID_UPDATE frame.

At all other times, the BTS sends PH-EMPTY-FRAME.req to any received PH-RTS.ind, causing the BTS-PHY to cease transmission in those periods.

NOTE

At any time, the BTS can end the silence period by issuing ONSET (in case of a new downlink talk-spurt or a FACCH downlink frame). See Section Section 2.2.4.



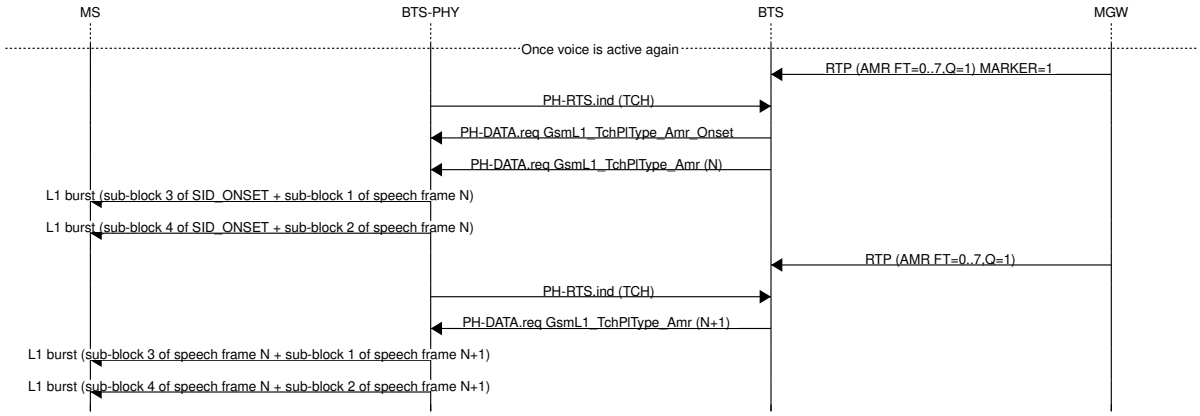
8VF

This happens every 8 **voice frames** (160ms), not every 8 GSM TDMA frames!

2.2.4 TCH/AHS Downlink: End of Silence; Start of Voice

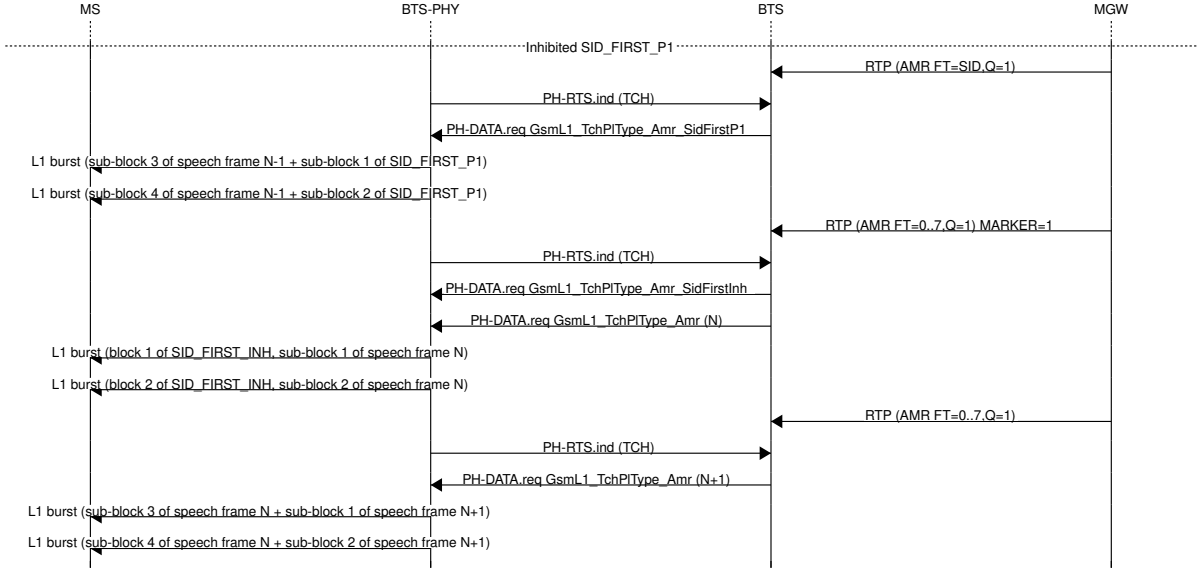
Once the BTS receives a non-SID AMR RTP frame (which should also have the MARKER bit set to 1, but let's not take that for granted), the contained AMR voice data is passed to the BTS-PHY in the next PH-DATA.req (GsmL1_TchPlType_Amr_Onset).

From that point onwards, regular transmission resumes, see Section 2.2.1.



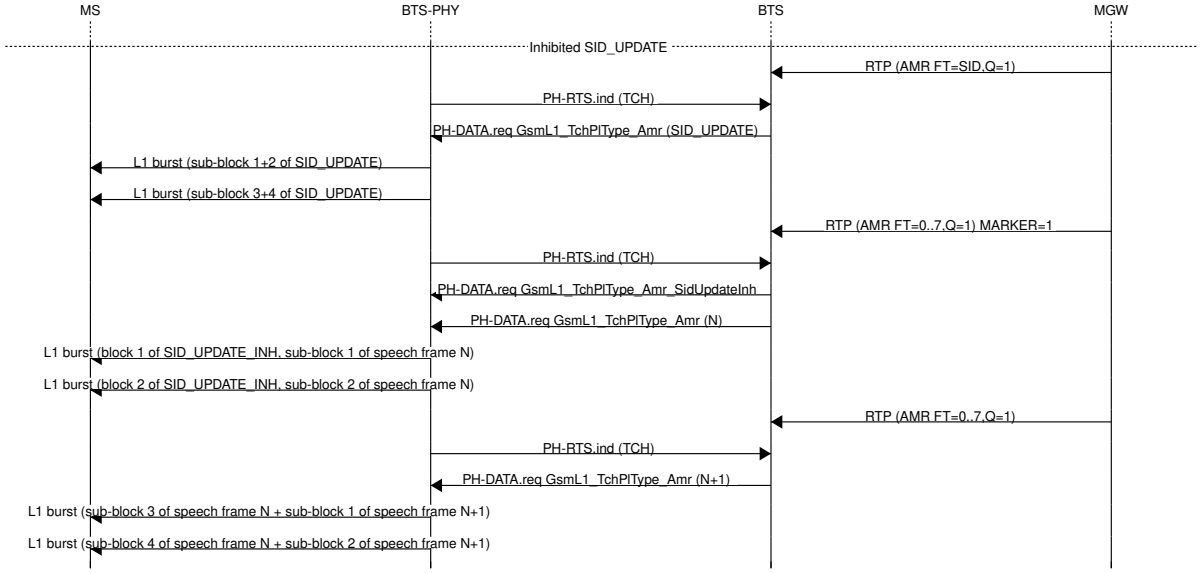
2.2.5 TCH/AHS Downlink: Inhibited SID_FIRST_P1

The following procedure must be observed in case of a SID_FIRST must be inhibited.



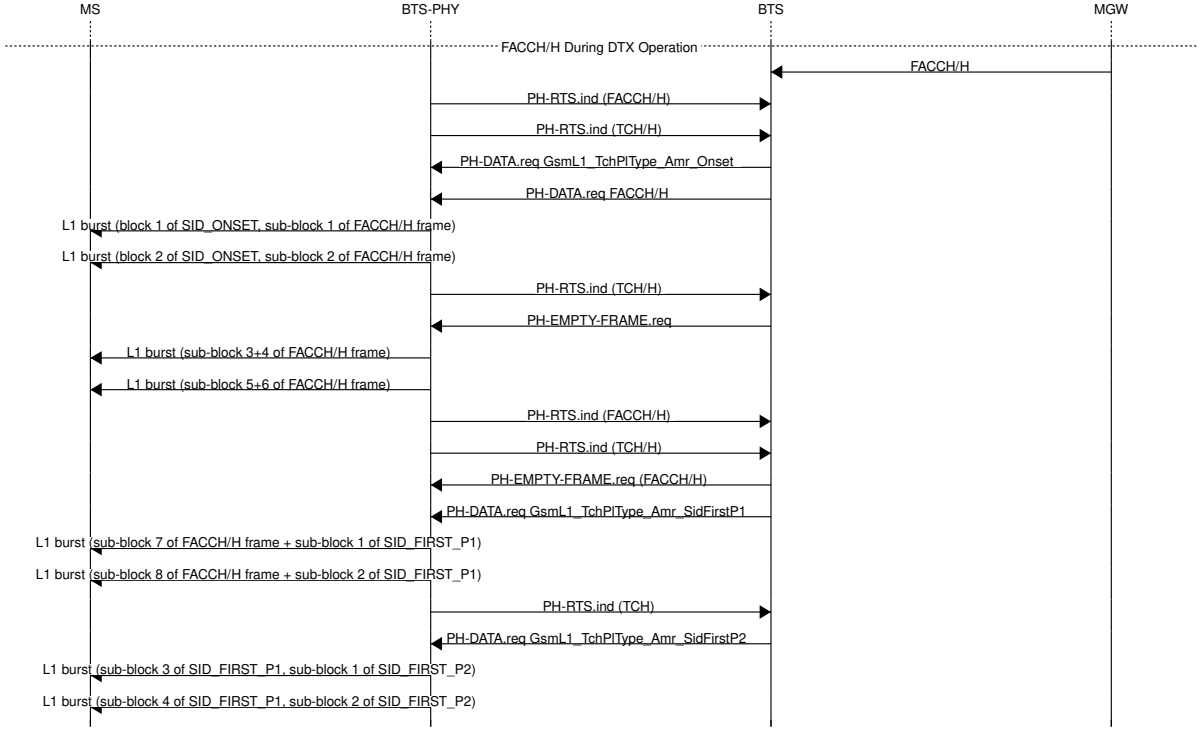
2.2.6 TCH/AHS Downlink: Inhibited SID_UPDATE

The following procedure must be observed in case of a SID_UPDATE must be inhibited.



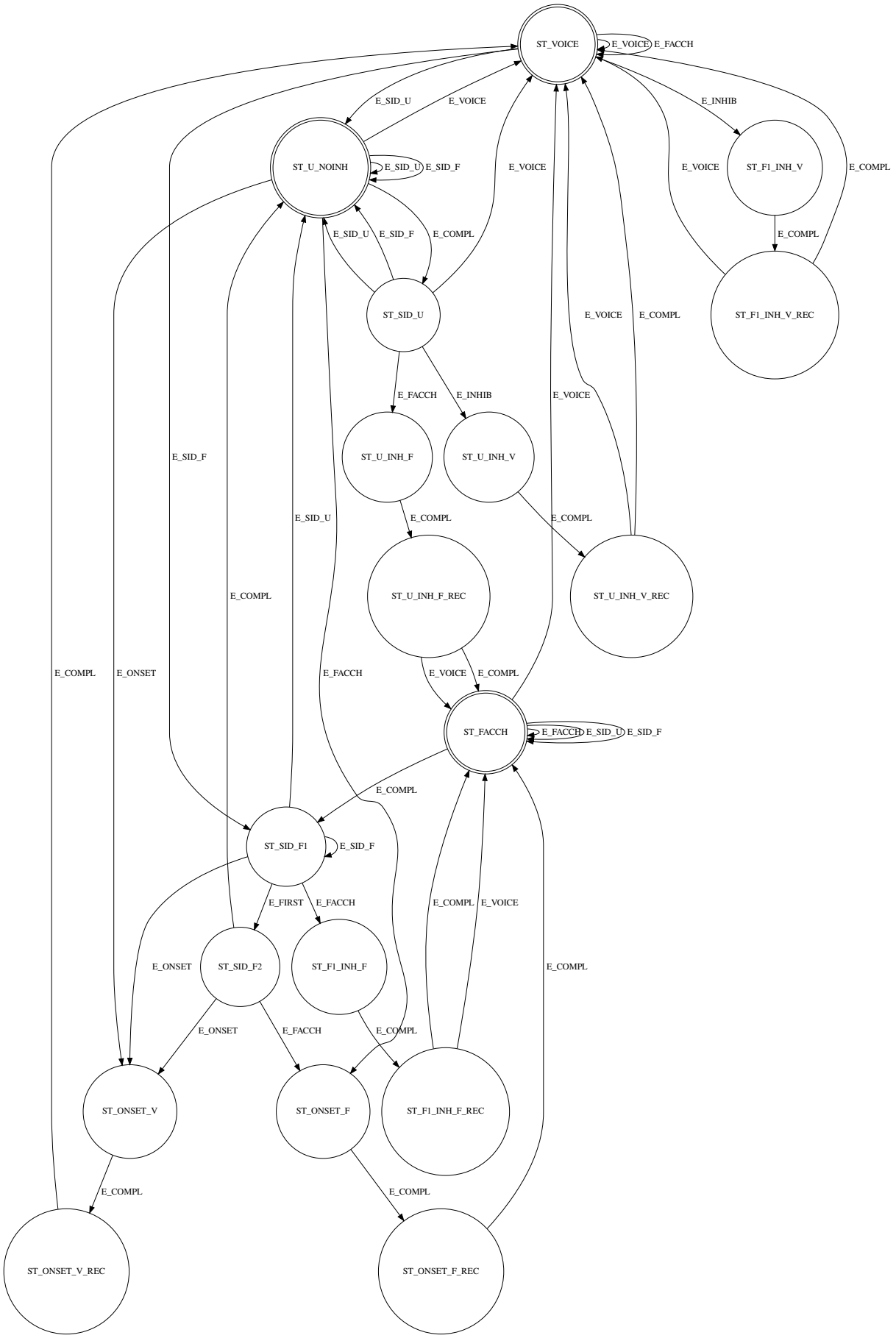
2.2.7 TCH/AHS Downlink: FACCH/H During DTX Operation

The following procedure must be observed in case of a FACCH/H frame must be inserted during DTX operation.



3 Implementation details

There is FSM implementing all the necessary states and transitions for DTX DL.


ST_VOICE
E_VOICE
E_FACCH
E_SID_U
E_VOICE
E_INHIB
ST_U_NOINH
E_SID_U
E_SID_F
E_VOICE
E_VOICE
ST_F1_INH_V
E_COMPL
E_SID_U
E_SID_F
E_COMPL
E_COMPL
ST_SID_U
E_VOICE
E_COMPL
ST_F1_INH_V_REC
E_FACCH
E_INHIB
E_VOICE
ST_U_INH_F
ST_U_INH_V
E_SID_F
E_SID_U
E_COMPL
E_COMPL
ST_U_INH_F_REC
ST_U_INH_V_REC
E_COMPL
E_COMPL
E_ONSET
E_FACCH
E_VOICE
E_COMPL
ST_FACCH
E_FACCH
E_SID_U
E_SID_F
E_COMPL
ST_SID_F1
E_SID_F
E_FIRST
E_FACCH
E_COMPL
E_VOICE
E_ONSET
ST_SID_F2
ST_F1_INH_F
E_COMPL
E_ONSET
E_FACCH
E_COMPL
ST_ONSET_V
ST_ONSET_F
ST_F1_INH_F_REC
E_COMPL
E_COMPL
ST_ONSET_V_REC
ST_ONSET_F_REC

The idea is that each state corresponds to the particular message type which have to be send to L1 next while state transition happens on incoming events like FACCH or Voice frames. There are 3 different classes of of events driving this FSM:

- Voice frame types: E_VOICE, E_SID_U, E_SID_F
- Incoming FACCH: E_FACCH
- Internal: E_ONSET, E_INHIB, E_COMPL

They represent different types of incoming RTP frames (Voice, SID UPDATE and SID FIRST correspondingly), incoming FACCH events or important events internal to DTX operations. The latter are Onset (interruption of silence period), Inhibition (of currently transmitted SID FIRST or UPDATE) and Completion (of silence initiation).

The double-circled states are "stationary" meaning that FSM can stay for longer periods in them. Other states are "transient" - the FSM have to switch away during next step. In practice this is implemented using E_COMPL signal which is issued in corresponding RTS handler or internal function (in case of SID First P1 $\rightarrow$ P2 transition).

The FSM states are grouped as follows:

- ST_VOICE: talkspurt, normal voice transmission
- ST_FACCH: transmission of higher-priority FACCH frame, interrupting current DTX state
- ST_ONSET*: handling of Onset event (interruption of silence period)
- ST_SID_F*: silence initiation
- ST_*_REC: sending of 2nd event for L1 during the same FN
- ST_*_INH: handling of Inhibition event (interruption of transmission of previous event)

The latter are specific to AMR HR where transmission of particular message by L1 takes longer so it can be aborted due to another incoming event. Note that for AMR FR only subset of this FSM is active (ST_SID_F2 and *INH states are never reached). This is implemented by signal emitting functions.

Note that the FSM states describe only the situation when we got to send something to L1, the transmission of Empty frames suppressing the actual radio transmission is done by other code.