# ETSI TR 102 793 V1.2.1 (2010-11)

*Technical Report*

**Speech and multimedia Transmission Quality (STQ); Measurements of Call Establishment Performance in IP Networks**

Reference

RTR/STQ-00180

Keywords

IP, packet mode, protocol, QoS, quality

***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

# Contents

# 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 (http://webapp.etsi.org/IPR/home.asp).

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 Technical Committee Speech and multimedia Transmission Quality (STQ).

# Introduction

The existing definitions of Call Establishment Delay are related only to the context of telephony on traditional PSTN networks.

In VoIP context the Call Establishment may differ for several reasons. Firstly, the tones played by terminal do not necessarily correspond to the state of the network. For example, the user terminal may play the ring back tone even though it has not received any signalling message from the network informing that the called party has been alerted. In addition it can happen that some network resources are not available immediately after pick-up on caller part because the media channel needs to be established separately.

In these conditions it is necessary to define new terms characterizing the Call Establishment as well as the related measurement methodology.

# 1 Scope

The present document provides a state of art concerning definitions of call establishment delay existing in standardization. As call establishment in PSTN context and in IP context differ, the present document defines a new term of call establishment delay which better reflects the course of call establishment in VoIP networks. The call establishment delay can be reflected by two parameters: Call Setup Delay and Media Establishment Delay.

# 2 References

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the reference document (including any amendments) applies.

Referenced documents which are not found to be publicly available in the expected location might be found at http://docbox.etsi.org/Reference.

NOTE: While any hyperlinks included in this clause were valid at the time of publication ETSI cannot guarantee their long term validity.

## 2.1 Normative references

The following referenced documents are necessary for the application of the present document.

Not applicable.

## 2.2 Informative references

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

[i.1] ETSI EG 202 765-2: "Speech Processing, Transmission and Quality Aspects (STQ); QoS and network performance metrics and measurement methods; Part 2: Transmission Quality Indicator combining Voice Quality Metrics".

[i.2] ITU-T Recommendation E.800: "Definitions of terms related to quality of service".

[i.3] ITU-T Recommendation E.600: "Terms and Definitions of Traffic Engineering".

[i.4] ITU-T Recommendation H.323: " Packet-based multimedia communications systems".

[i.5] IETF RFC 3261: "SIP: Session Initiation Protocol".

[i.6] IETF RFC 5245: "Interactive Connectivity Establishment (ICE): A Protocol for Network Address Translator (NAT) Traversal for Offer/Answer Protocols".

[i.7] IETF RFC 5389: "Session Traversal Utilities for NAT (STUN)".

[i.8] ETSI TS 123 228: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; IP Multimedia Subsystem (IMS); Stage 2 (3GPP TS 23.228 version 9.3.0 Release 9)".

[i.9] ETSI TS 124 229: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Internet Protocol (IP) multimedia call control protocol based on Session Initiation Protocol (SIP) and Session Description Protocol (SDP); Stage 3 (3GPP TS 24.229 version 9.3.1 Release 9)".

[i.10] TIA FS 1037C: "Telecommunications: Glossary of Telecommunication Terms".

[i.11] IETF RFC 3389: "Real-time Transport Protocol (RTP) Payload for Comfort Noise (CN)".

[i.12] IETF RFC 3551: "RTP Profile for Audio and Video Conferences with Minimal Control".

# 3 Abbreviations

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

| | |
|---|---|
| ETSI | European Telecommunications Standards Institute |
| IP | Internet Protocol |
| ISDN | Integrated Services Digital Network |
| ICE | Interactive Connectivity Establishement |
| IMS | IP Multimedia Subsystem |
| ITU-T | International Telecommunication Union - Telecommunication standardization sector |
| MTSI | Multimedia Telephony Service for IMS |
| NAT | Network Adress Translation |
| PSTN | Public Switched Telephone Network |
| RTP | Realtime Transport Protocol |
| SBC | Session Boarder Controller |
| SDP | Session Description Protocol |
| SIP | Session Initiation Protocol |
| STUN | Session Traversal Utilities for NAT |
| TIA | Telecommunications & Information Administration |
| UDP | User Datagram Protocol |
| UE | User Equipment |
| VoIP | Voice over Internet Protocol |
| WLAN | Wireless Local Area Network |

# 4 State of Art

Several definitions of call establishment delay are present in standards issued by different standardisation bodies: ETSI, ITU-T and TIA.

According to document EG 202 765-2 [i.1] the Post Dialling Delay *"is the time interval between the end of dialling by the caller and the reception back by him of the appropriate ringing tone or recorded announcement"*. This definition regards call establishment delay from end-user perspective.

ITU-T defines this parameter in two recommendations: ITU-T Recommendations E.600 [i.3] and E.800 [i.2].

The definition in ITU-T Recommendation E.600 [i.3] adds some significant adjustments to the definition by EG 202 765-2 [i.1]. In fact according to ITU-T Recommendation E.600 [i.3] the post-dialling delay is *"the time interval between the end of dialling by the user and the reception by him of the appropriate tone or recorded announcement, or the abandon of the call without tone"*. The fact of taking into consideration the abandon of the call without tone as an instant of call establishment termination can create some problems for supervision of service and for gathering of statistics concerning service functioning. For real calls this means taking into account the calls which have been terminated prematurely and for test calls (carried out by robots which terminate a call after timeout) it means taking into account calls which in fact ware not successful.

ITU-T Recommendation E.600 [i.3] defines in addition the post-selection delay used in ISDN networks which is defined as follows:

*"a) post-selection delay (overlap sending) is defined as the time interval from the instant the first bit of the INFORMATION message containing the last selection digit is passed by the calling terminal to the access signalling system until the last bit of the first message indicating call disposition is received by the calling terminal (ALERTING message in case of successful call).*

*b) post-selection delay (en-bloc sending) is defined as the time interval from the instant the first bit of the initial SETUP message containing all the selection digits is passed by the calling terminal to the access signalling system until the last bit of the first message indicating call disposition is received by the calling terminal (ALERTING message in case of successful call)."*

It is worthy of noticing that ITU-T Recommendation E.600 [i.3] defines the event of call establishment in terms of protocol messages. It is the only definition which regards the call establishment delay from both perspectives: user and network.

In addition, one can also find a definition of call set-up time by ITU-T Recommendation E.800 [i.2]: *"The period starting when the address information required for setting up a call is received by the network (recognized on the calling user's access line) and finishing when the called party busy tone, or ringing tone or answer signal is received by the calling party (i.e., recognized on the calling user's access line). Local, national and service calls should be included, but calls to Other Licensed Operators should not, as a given operator cannot control the QoS delivered by another network."* This definition regards the call establishment delay from network perspective.

The last definition of call delay defined by TIA is based on document FS 1037C [i.10] issued by US Federal Standards which defines the following: *"The time between the instant a system receives a call attempt and the instant of initiation of ringing at the call receiver end instrument"*. Yet again although named differently, this definition describes the call establishment delay from network perspective.

It can be noticed that the essence of all definitions is the same although they use different terms to describe it. There are some minor differences between these definitions, which generally stem from the fact that they consider the call establishment either from network or user point of view. In our opinion both approaches are valid. Call establishment delay considered from user perspective is generally called post dialling delay; whereas from network perspective different names are used: post-selection delay, call set-up time or call delay.

# 5 Need for a new indicator

The aforesaid discussion reflects well the call establishment delay definitions in PSTN and thus can be directly employed for measurements of this indicator. In PSTN networks the invitation to dialling signifies that the necessary resources to pass a call are available. In addition the ring back tone signal indicates that the call is established and that a circuit is reserved. The call can begin at any time which depends only on called party, as the network has prepared everything for the call to start.

The situation in IP networks is different. Not only the signal of invitation to dialling does not signify that the resources are available, also the ring back tone does not indicate that the network has prepared everything for the call to start. In fact, due to longer transmission delay in IP networks the call establishment for VoIP may take longer than in PSTN. Some devices are therefore inclined to diminish the waiting by groundlessly introducing ring back tone signal before the network has actually informed that the called party has been alerted. Therefore the fact of hearing the ring back tone signal in handset does not necessarily indicate that the telephone rings at the called party.

In addition, no matter which signalling protocol is used for call establishment, in VoIP telephony a media channel needs to be established separately. This may take less or more time. In some situations the interlocutors are obliged to wait for several milliseconds after the called party has answered the phone before they can actually hear themselves speaking, as the media channel is not yet ready. It seems natural to claim that the call has been established when the network has provided the means for both interlocutors to converse and hear one another. As this delay is usually perceived by users it should therefore be taken into consideration when defining call establishment indicators. Media establishment delay can be equal to zero if the media flow has been established before answering the phone. Notice that in situations where media flow establishment begins as soon as the called user is being alerted, this indicator may be user dependant. In fact, in such cases the more time it will take for the called party to answer the phone, the smaller this indicator may be. However this should not be seen as a drawback as the purpose of this indicator is mainly to evaluate user perception.

Consequently, it seems important to define a new term of call establishment delay which will better reflect the course of call establishment in VoIP networks. The call establishment delay can be reflected by two parameters: Call Setup Delay and Media Establishment Delay.

# 6 Measurements of call establishment performance

The measurements of call establishment performance in a passive way are done at one point while analysing the IP flows. The different phases of call establishment and associated signalling messages for SIP and H.323 have been presented on figure 1.

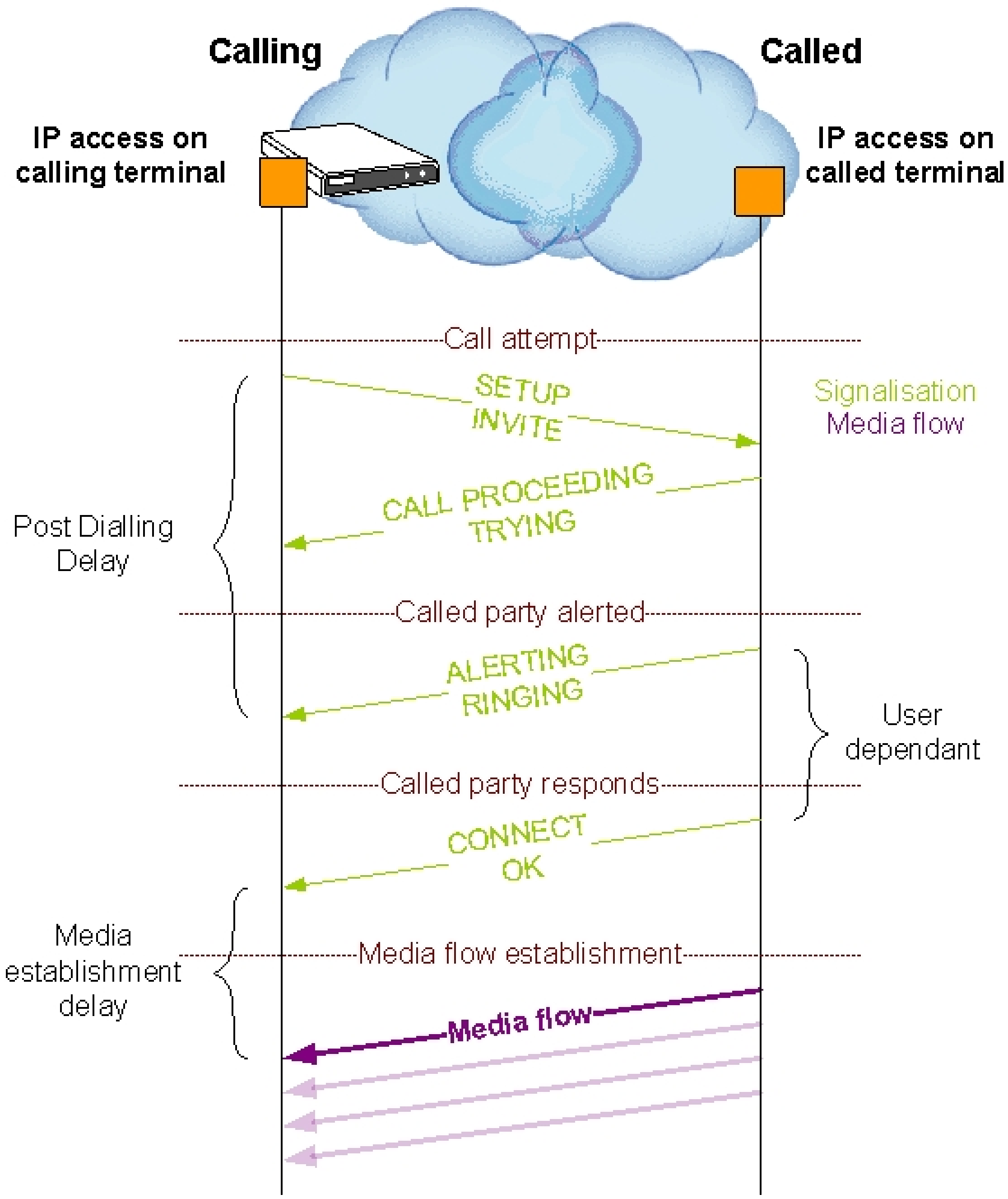



**Figure 1: Call establishment phases in VoIP context**

## 6.1 Call Setup Delay [s]

### 6.1.1 Abstract Definition

Time between sending of complete address information and receipt of call set-up notification.

### 6.1.2 Abstract Equation

$$\boxed{\text{Call Setup Delay [s]} = \left(t_{\text{connect established}} - t_{\text{user presses send button on handset}}\right)[\text{s}]}$$

NOTE: This parameter is not calculated unless the telephony call setup attempt is successful. Note that a call with "user busy" is also considered as a successful call setup.

### 6.1.3 Trigger Points

**Table 1**

<table>
<tr><th>Event from abstract equation</th><th>Trigger point from user's point of view</th><th>Technical description/protocol part</th></tr>
<tr><td rowspan="2">$t_{\text{user presses send button on handset}}$: Time of call attempt</td><td rowspan="2">Start: Push send button (e.g. #) or end of timeout. (See note 3).</td><td>Start (H.323): The forming of first CONNECTION REQUEST message 'SETUP'.</td></tr>
<tr><td>Start (SIP): The forming of first CONNECTION REQUEST message 'INVITE'.</td></tr>
<tr><td rowspan="2">$t_{\text{connection established}}$: Time when connection is established (successful call attempt)</td><td rowspan="2">Stop: Alerting tone is heard by the A-party coming from the B-party.</td><td>Stop (H.323): The message informing of ring back tone 'ALERTING' or that the B-party user is occupied 'RELEASE COMPLETE' with release code 17 is received from the B-party.</td></tr>
<tr><td>Stop (SIP): The message informing of ring back tone '180 RINGING' or that the B-party user is occupied '486 Busy Here' or '600 Busy Everywhere' is received from the B-party.</td></tr>
<tr><td colspan="3">NOTE 1: With automatic tools there is not a significant difference between considering the alerting or the connect message, as the answer machine should always answer immediately.<br>NOTE 2: For the trigger points of the technical description/protocol part, see figure 1.<br>NOTE 3: For technical reasons if the push send button is not pressed by user telecom systems wait a timeout before processing the dialled number. This timeout is excluded from the measurement as it can be cut short by the user.</td></tr>
</table>

## 6.2 Media Establishment Delay [ms]

### 6.2.1 Abstract Definition

Time between the moment the B-party answers the phone and the moment when media flows are established in both directions.

### 6.2.2 Abstract Equation

$$\boxed{\text{Media Establisment delay}\,[\text{ms}] = \max\left(\left(t_{\text{media flows established}} - t_{\text{B-party handset off-hook}}\right);0\right)[\text{ms}]}$$

NOTE 1: This parameter is not calculated unless the telephony call setup attempt is successful and the call was established (e.g. no busy tone).

NOTE 2: In some systems the media flow is established before the B-party handset is off-hook. In such situations the Media Establishment delay should be equal to zero (not negative). See annex B for additional information about such situations.

### 6.2.3 Trigger Points

**Table 2**

| Event from abstract equation | Trigger point from user's point of view | Technical description/protocol part |
|---|---|---|
| $t_{\text{B-party handset off hook}}$: Time of taking the handset off-hook at the B-party | Start: B-party answers the phone and end of ring back tone at the A-party. | Start (H.323): The receiving of CONNECTION message 'CONNECT'. |
| | | Start (SIP): The receiving of CONNECTION message '200 OK'. |
| $T_{\text{media flows established}}$: Time when media flows are established | Stop: A-party can hear B-party and inversely. | Stop: The reception of first media packet. (See note 5). |

NOTE 1: With automatic tools there is not a significant difference between considering the alerting or the connect message, as the answer machine should always answer immediately. The user dependant time is therefore negligible.
NOTE 2: This parameter cannot be evaluated at analog interface as the connection message is not received.
NOTE 3: Even if there is a difference between the time the called user lifts the handset and the time a CONNECTION message (e.g. CONNECT for ITU-T Recommendation H.323 [i.4] or 200 OK for RFC 3261 [i.5] SIP) arrives at the caller side, this difference is small enough to allow the measurement of media establishment delay at the caller side in order to avoid expensive time synchronisation between the caller and the called side.
NOTE 4: For the trigger points of the technical description/protocol part, see figure 1.
NOTE 5: The first packet sent is not necessarily the first packet received. See considerations regarding NAT functions in clause 6.2.5.

### 6.2.4 Informer section

For analog terminals there exists a delay between the instant of taking the handset off-hook and the moment when the terminal is ready to transmit and receive speech. This delay is not embraced by Media Establishment Delay.

### 6.2.5 Special Considerations

Special caution needs to be taken if the network implements the NAT functions (e.g. in Session Border Controller or in Home Gateway). In fact NAT will drop all packets (usually around 6 packets to 10 packets) arriving from the B-party unless A-party also sent some media packets (see figure 2). If NAT-discovery mechanisms as described in annex B are used, this drop can be reduced.

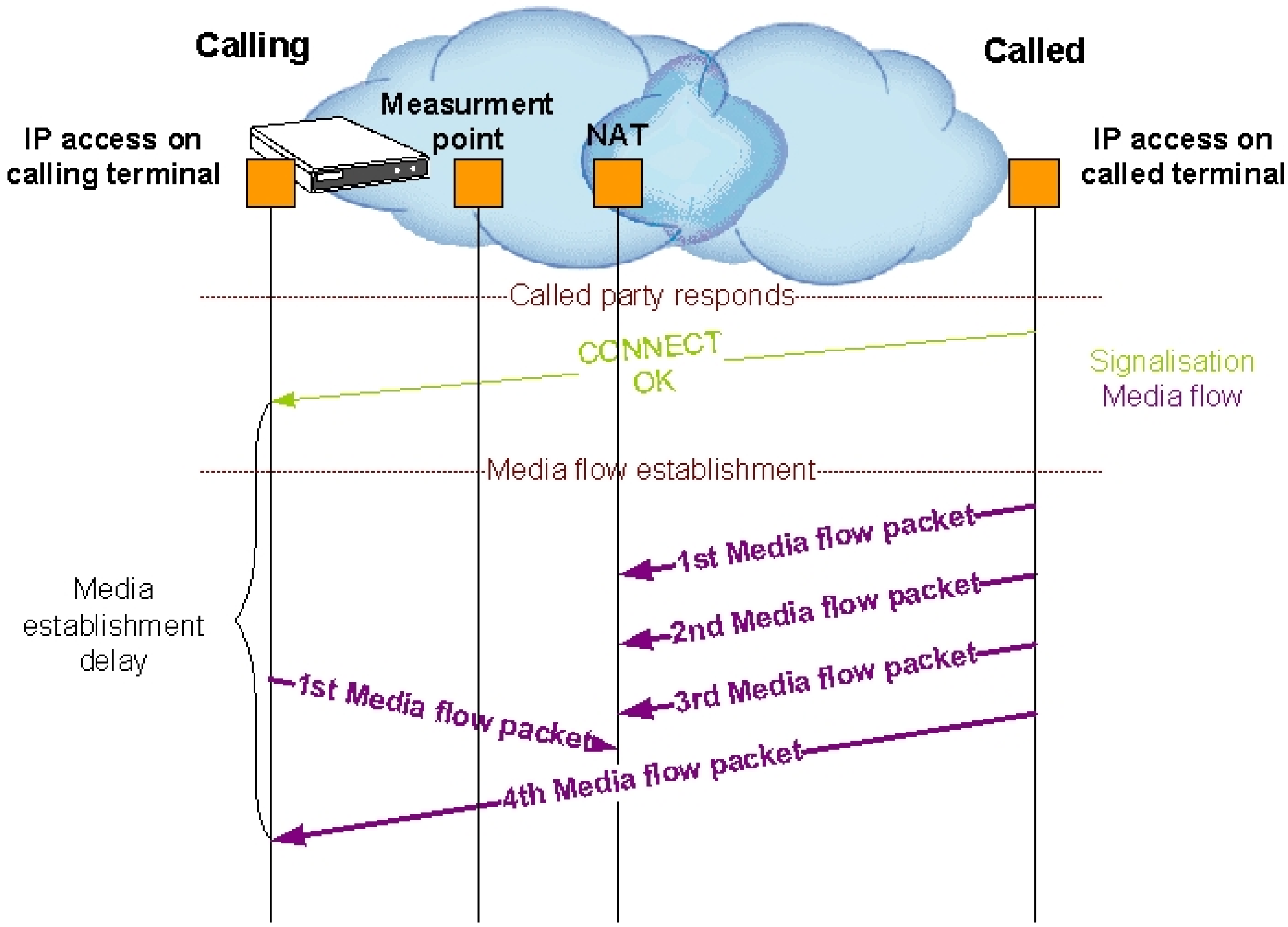


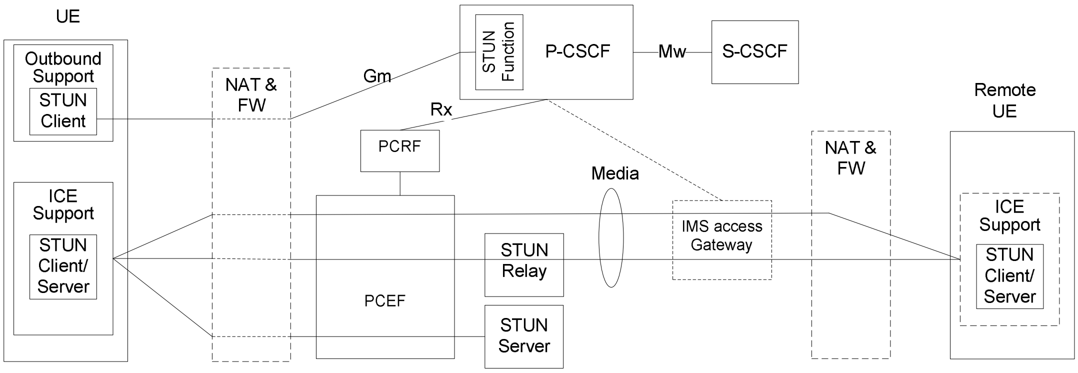

**Figure 2: Dropping of media flow packets by NAT functionality**

Measurement of media establishment delay can therefore take place in terminal or in network as long as the measurement takes place outside the NAT functions.

# Annex A:
# Void

# Annex B: Alternative NAT Handling in IMS

During the call setup of an MTSI call, both the control signalling and the media data need to pass from the A party to the B part and vice versa. As shown in figure B.1, there might be NAT and/or Firewall devices located between the clients and the IMS network.




NOTE: This figure corresponds to figure G.2a of TS 123 228 [i.8].

**Figure B.1: Reference model for ICE and Outbound Methodology**

If cellular access is used (e.g. GPRS, EDGE, HSPA or similar) these NATs should normally be managed by the operator, and thus open up the NAT pinholes needed for control signalling and media transport. Such managed NATs are often called Session Border Controllers (SBCs), which besides opening up the necessary ports also makes the relevant rewrites/translations of the address parts of the protocol headers. In the IMS architecture the SBCs are the P-CSCF in the signalling plane, and the IMS Access Gateway in the media plane.

However, it is also possible that non-managed NATs are present outside the IMS network. This is typically true if the client uses the MTSI service over WLAN or similar hot-spot access types, where the NAT might be hosted at the internet café or at the customer's home (there typically also for wired access). It can also be valid for cellular accesses if the operator does not coordinate the NATs in the IP network and the IMS network.

For non-managed NAT cases it can be the responsibility of the clients to detect the presence of the NAT, and to handle the call setup accordingly (if not done by a SBC). This is typically done using the ICE [i.6] and STUN [i.7] protocols, where NATs are detected and opened by communicating with a STUN server during the session setup. Other methods might also be used, and TS 123 228 [i.8] describes in clause G.1:

> *It shall be possible for an operator to use one or more of NAT traversal methods in its IMS domain. The selection of the method for a particular case shall depend on the UE's capabilities, the capabilities of the network and policies of the operator.*

In all cases this means that the call setup when non-managed NATs are present will take longer time, and it also means that the NATs at both ends might open the NAT pinholes at slightly different times during the setup phase. Note that signalling and media are normally handled separately by the NATs (due to the different ports used), so media pinholes are not opened during the SIP negotiation phase until the client actually starts to send something on the media port.

Thus some RTP messages might be thrown away by the NATs in the beginning of the call until the media paths for both NATs are open. To keep this time short, TS 124 229 [i.9] specifies in clause F.5:

> *To allow the IMS access gateway to perform address latching, for a given UDP-based media stream, the UE shall use the same port number for sending and receiving packets.*
>
> *To allow early media flows, the UE shall send keepalive messages for each UDP-based media stream as soon as an SDP offer or answer is received in order to allow the IMS access gateway to perform address latching before the call is established.*

*To keep NAT bindings and firewall pinholes open for the UDP-based media streams, and enable the IMS access gateway to perform address latching, the UE shall send keepalive messages for each media stream as defined in subclause K.5.2.1.*

Clause K.5.2.1 in TS 124 229 [i.9] further describes the keepalive messages for cases when ICE/STUN is used or for cases when it not supported:

*NAT bindings also need to be kept alive for media. draft-ietf-mmusic-ice [99] provides requirements for STUN based keepalive mechanisms. UEs that do not implement the ICE procedures as defined in draft-ietf-mmusic-ice [99] should implement the keepalive procedures defined in draft-ietf-mmusic-ice [99].*

*In the case where keepalives are required and the other end does not support ICE (such that STUN cannot be used for a keepalive) or the UE can not discover STUN or TURN servers to gather candidates, the UE shall send an empty (no payload) RTP packet with a payload type of 20 as a keepalive as long as the other end has not negotiated the use of this value.*

*If this value has already been negotiated, then some other unused static payload type from table 5 of RFC 3551 [55A] shall be used. When sending an empty RTP packet, the UE shall continue using the sequence number (SSRC) and timestamp as the negotiated RTP stream.*

Draft-ietf-mmusic-ice [i.6] further describes the keepalive messages in section 10:

*The keepalive SHOULD be sent using a format that is supported by its peer. ICE endpoints allow for STUN-based keepalives for UDP streams, and as such, STUN keepalives MUST be used when an agent is a full ICE implementation and is communicating with a peer that supports ICE (lite or full).*

*If the peer does not support ICE, the choice of a packet format for keepalives is a matter of local implementation. A format that allows packets to easily be sent in the absence of actual media content is RECOMMENDED. Examples of formats that readily meet this goal are RTP No-Op [NO-OP-RTP], and in cases where both sides support it, RTP comfort noise [RFC 3389].*

*If the peer doesn't support any formats that are particularly well suited for keepalives, an agent SHOULD send RTP packets with an incorrect version number, or some other form of error that would cause them to be discarded by the peer.*

This means that for a 3GPP-compliant MTSI client, the first packets (keepalive packets) sent on the media port during the session setup can be either STUN packets, empty RTP payload packets, or RTP packets with some unused payload type. A non-3GPP-compliant client can also use comfort noise RTP packets, or even intentionally corrupt RTP packets.

On the other hand, for all the above cases, as soon as any UDP packet is received by the client on the media port, it knows that the NAT is open for incoming media packets, and thus the call setup phase, including the media path setup, can be seen as complete. So the reception of the first UDP packet can be used as an end trigger for the media setup phase.

Another more restrictive option could be to use the first received RTP media packet with a valid payload type (thus excluding keepalive and RTCP packets) but this assumes that the B-party starts sending such media packets directly after call establishment.

# History

| Document history | | |
|---|---|---|
| V1.1.1 | March 2010 | Publication |
| V1.2.1 | November 2010 | Publication |
| | | |
| | | |
| | | |