Internet Engineering Task Force SIPPING WG
Internet Draft
Document: <draft-ietf-sipping-toip-00.txt>
A. van Wijk (editor)
Viataal
October 17 2004
Expires: April 15 2005
Informational

Framework of requirements for real-time text conversation using SIP.

Status of this Memo

Abstract

This document provides the framework of requirements for text conversation with real time character-by-character interactive flow over the IP network using the Session Initiation Protocol. The requirements for general real-time text-over-IP telephony, point-to point and conference calls, transcoding, relay services, user mobility, interworking between text-over-IP telephony and existing text-telephony, and some special features including instant messaging have been described.

Table of Contents

1. Introduction

Text-over-IP (ToIP) is becoming popular as a part of total conversation among a range of users although this medium of communications may be the most convenient to certain categories of people (e.g., deaf, hard of hearing and speech-impaired individuals). The Session Initiation Protocol (SIP) has become the protocol of choice for control of Multimedia IP telephony and Voice-over-IP (VoIP) communications. Naturally, it has become essential to define the requirements for how ToIP can be used with SIP to allow text conversations as an equivalent to voice. This document defines the framework of requirements for using ToIP, either by itself or as a part of total conversation using SIP for session control.

2. Scope

The primary scope of this document is to define the requirements for using ToIP with SIP, either stand-alone or as a part of a total conversation approach. In general, the scope of the requirements is:

a. Features in Real-Time ToIP
b. Real-time Multimedia Conversational Sessions using SIP
c. General Requirements for Real-Time ToIP using SIP
d. Interworking Requirements for ToIP

e. Text gateways to interconnect the different networks

The subsequent sections describe those requirements in detail.

3. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [2].

4. Definitions

Audio bridging - a function of a gateway or relay service that enables an audio path through the service between the users involved in the call.

Full duplex - user information is sent independently in both directions.

Half duplex - user information can only be sent in one direction at a time or, if an attempt to send information in both directions is made, errors can be introduced into the user information.

Interactive text - a term for real time transmission of text in a character-by-character fashion for use in conversational services.

TTY - name for text telephone, often used in USA, see textphone. Also called TDD, Telecommunication Device for the Deaf.

Textphone - text telephone. A terminal device that allow end-to-end real time text communication. A variety of textphone protocols exists world-wide, both in the PSTN and other networks. A textphone can often be combined with a voice telephone, or include voice communication functions for simultaneous or alternating use of text and voice in a call.

Text bridging - a function of a gateway or relay service that enables the flow of text through the service between the users involved in the call.

Text gateway - a multi functional gateway that is able to transcode between different forms of text transport methods. E.g. Between ToIP in IP networks and Baudot text telephony in the PSTN.

Text telephony - Analog textphone services

Text Relay Service - A third-party or intermediary that enables communications between deaf, hard of hearing and speech-impaired people, and voice telephone users by translating between voice and text in a call.

Transcoding Services - Services of a third-party user agent (human or automated) that transcodes one stream into another.

Total Conversation - A multimedia service offering real time conversation in video, text and voice according to interoperable standards. All media flow in real time. Further defined in ITU-T F.703 Multimedia conversational services description.

Video Relay Service - A service that enables communications between deaf and hard of hearing people with total conversation devices, and hearing persons with voice telephones by translating between sign language and spoken language in a call.

Acronyms:

2G	Second generation cellular (mobile)
2.5G	Enhanced second generation cellular (mobile)
3G	Third generation cellular (mobile)
CDMA	Code Division Multiple Access
CTM	Cellular Text Telephone Modem
GSM	Global System of Mobile Communication
ISDN	Integrated Services Digital Network
ITU-T	International Telecommunications Union-Telecommunications standardisation Sector
PSTN	Public Switched Telephone Network
SIP	Session Initiation Protocol
TDD	Telecommunication Device for the Deaf
TDMA	Time Division Multiple Access
ToIP	Text over Internet Protocol
UTF-8	Universal Transfer Format-8

5. Background and General Requirements

The main purpose of this document is to provide a set of requirements for real-time text conversation over the IP network using the Session Initiation Protocol (SIP) [3]. The overall requirement is that real-time text conversation can be part of a conversational service like any other media. Participants can negotiate all media including real-time text conversation[4, 5]. This is a highly desirable function for all IP telephony users, and essential for deaf, hard of hearing, or speech impaired people who have limited or no use of the audio path of the call.

It is important to understand that real-time text conversations are significantly different from other text based communications like email or instant messaging. Real-time text conversations deliver an equivalent mode to voice conversations by providing transmission of text character by character as it is entered, so that the conversation can be followed closely and immediate interaction take place, therefore providing the same mode of interaction as voice telephony does. Store-and-forward systems like email or messaging on mobile networks or non-streaming

systems like instant messaging are unable to provide that functionality.

One particular application where real-time text is absolutely essential, is the use of relay services between conversational modes, like between text and voice.

Direct text emergency service calls, where time and continuous connection are of the essence, is another essential application.

6. Features in Real-time Text-over-IP

While real-time Text-over-IP will be used for a wide variety of services, an important field of application will be to provide a text equivalent to voice conversation, in particular for deaf, hard of hearing and speech-impaired users.
As such, it is crucial that the conversational nature of this service is maintained. Text based communications exist in a variety of forms, some non-conversational (SMS, text paging, E-mail, newsgroups, message boards, etc.), others conversational (TTY/TDD, Textphone, etc).

Real-time Text-over-IP will sometimes be used in conjunction with a relay service [I] to allow text users to communicate with voice users. With relay services, it is crucial that text characters are sent as soon as possible after they are entered. While buffering MAY be done to improve efficiency, the delays SHOULD be kept as small as possible. In particular, buffering of whole lines of text MUST NOT be used.

In order to make Real-Time Text-over-IP the equivalent of what voice is to hearing people, it needs to offer equivalent features in terms of conversation as voice communications provides to hearing people. To achieve that, real-time Text-over-IP MUST:

a. Offer Real-Time presentation of the conversation. This means that text MUST be sent as soon as available, or with very small delays. The delay MUST not be longer than 300 milliseconds,

b. Provide simultaneous transmission in both directions,

c. Provide interoperability with text conversation features in other networks, e.g. PSTN, accepting functional limitations that this will lead to during interoperation.

d. Support a transmission rate of at least 30 characters/second.

e. Support suitable reliability of text transmission. A character error rate of 0.2% is regarded good, and 1% usable.

f. Be possible to merge with video and voice transmission.

g. The end-to-end delay in transmission MUST be less than 2000 milliseconds.

Many users will want to use multiple modes of communication during the conversation, either at the same time or by switching between modes e.g. between real-time Text-over-IP and voice. Native real-time Text-over-IP systems MUST support simultaneous use of modalities so that the text interface is always available.

When communicating via a gateway to other networks and protocols, the system MUST completely support the functionality for alternating or simultaneous modalities as offered by the gateway.

When voice is supported on the terminal, the terminal MUST provide volume control.

7. Real-Time Multimedia Conversational Sessions using SIP

The Session Initiation Protocol (SIP) [3] provides mechanisms for creating, modifying, and terminating sessions for real-time conversation with one or more participants using any combination of media: Text, Video and Audio. However, participants are allowed to negotiate on a set of compatible media types (e.g., Text, Video, Audio) with session descriptions used in SIP invitations.

The standardized T.140 real-time text conversation [4], in addition to audio and video communications, will be a valuable service to many. Real-time text can be expressed as a part of the session description in SIP and is a useful subset of the Total Conversation (which is Real-time text, Video and Audio simultaneously).

This specification describes the framework for using the T.140 text conversation in SIP as a part of the multimedia session establishment in real-time over a SIP network.

The session establishment using SIP defines procedures for how T.140 text conversation can be supported using the text/t140 RTP payload defined in RFC 2793 [5]. The performance characteristics of T.140 will be determined using RTCP.

The session will not only define procedures between the SIP devices having text conversation capability, but will also define how sessions in SIP can be established between the text conversation and audio/video/text capable devices transparently.

If there is any incompatibility between the terminals, e.g. T.140 only and audio-only terminals, the necessary transcoding services will need to be invoked. This important service feature offers a variety of rich capabilities in the transcoding server. For example, speech-to-text (STT), text-to-speech (TTS), text bridging after conversion from speech, audio bridging after conversion from text, and other services can also be provided by the transcoding

and/or translation server. The session description protocol (SDP) [6] used in SIP to describe the session also needs to be capable of expressing these attributes of the session (e.g., uniqueness in media mapping for conversion from one media to another for each communicating party).

Real-time text can also be presented in conjunction with video and audio. Making real-time text part of total conversation.

Visual and/or Tactile alerting for T.140 capable terminals should to be provided.

Users may set up text conversation sessions using SIP from any location. In addition, user privacy and security MUST be provided for text conversation sessions at least equal to that for voice.

The transcoding/translation services can be invoked in SIP using different session establishment models [7]: Third party call control [8] and Conference Bridge model [9].

Both point-to-point and multipoint communication need to be defined for the session establishment using T.140 text conversation. In addition, the interworking between T.140 text conversation and text telephony conversation [10] is needed.

The general requirements for real-time text conversation using SIP can be described as follows:

a. Session setup, modification and teardown procedures for point-to-point and multimedia calls

b. Registration procedures and address resolutions

c. Registration of user preferences

d. Negotiation procedures for device capabilities

e. Discovery and invocation of transcoding/translation services between the media in the call

f. Different session establishment models for transcoding/translation services invocation: Third party call control and Conference bridge model

g. Uniqueness in media mapping to be used in the session for conversion from one media to another by the transcoding/translation server for each communicating party

h. Media bridging services for T.140 real-time text, audio, and video for multipoint communications

i. Transparent session setup, modification, and teardown between text conversation capable and voice/video capable devices

j. Conversations to be carried out using T.140-over-RTP and RTCP will provide performance report for T.140

k. Altering capability using text conversation during the session establishment

l. T.140 real-time text presentation mixing with voice and video

m. T.140 real-time text conversation sessions using SIP, allowing users to move from one place to another

n. User privacy and security for sessions setup, modification, and teardown as well as for media transfer

o. Interoperability between T.140 conversations and analogue text telephones

p. Routing of emergency calls according to national or regional policy to the same level of a voice call.

8. General Requirements for Real-Time Text-over-IP using SIP

The communications environments for ToIP using SIP to set up the conversation in real-time may vary from a simple point-to-point call to multipoint calls in addition to the fact that ToIP can be used in combination with other media like audio and video. In order to establish the session in real-time, the communicating parties SHOULD be provided with experiences like those of normal telephony call setup. There may also be some need for pre-call setup e.g. storing registration information in the SIP registrar to provide information about how a user can be contacted. This will allow calls to be set up rapidly and with proper addressing.

Similarly, there are requirements that need to be satisfied during call set up when another media is preferred by a user. For instance, some users may prefer to use audio while others want to use text as their preferred choice of conversational mode. In this case, transcoding services will need to be invoked for text-to-speech (TTS) and speech-to-text (STT). The requirements for transcoding services need to be negotiated in real-time to set up the session.

The subsequent subsections describe those requirements in great detail.

8.1 Pre-Call Requirements

The desire of the users for using ToIP as a medium of communications can be expressed during registration time. Two situations need to be considered in the pre-call setup environment:

a. User Preferences: It MUST be possible for a user to indicate a preference for ToIP by registering that preference in a SIP server. If the user is called by other party, preferences can be invoked by the SIP server to accept or reject the call based on the rules defined by the user. If the rules require that a transcoding server is needed, the call can be re-directed or handled accordingly.

b. Server to support User Preferences: SIP servers MUST have the capability to act on users preferences for ToIP, based on the user preferences defined during the pre-call setup registration time.

8.2 Basic Point-to-Point Call Requirements

The point-to-point call will take place between two parties. The requirements are described in subsequent sub-sections. They assume that one or both of the communicating parties will indicate ToIP as a possible or preferred medium for conversation using SIP in the session setup.

8.2.1 General Requirements

The general requirements are that ToIP will be chosen from the available media as the preferred means of communication for the session. However, there may be a need to invoke some underlying capabilities in some cases, for example, a transcoding server may be invoked if one of the users want to use a communication medium other than ToIP.
The following features MAY need to be involved to facilitate the session establishment using ToIP as another medium:

a. Caller Preferences: SIP headers (e.g., Contact) can be used to show that ToIP is the medium of choice for communications.

b. Called Party Preferences: The called party being passive can formulate a clear rule indicating how a call should be handled either using ToIP as a preferred medium or not, and whether a designated SIP proxy needs to handle this call or it is handled in the SIP user agent (UA).

c. SIP Server support for User Preferences: SIP servers can also handle the incoming calls in accordance to preferences expressed for ToIP. The SIP Server can also enforce ToIP policy rules for communications (e.g., use of the transcoding server for ToIP).

8.2.2 Session Setup

Users will set up a session by identifying the remote party or the service they will want to connect to. However, conversations could be started using a mode other than real-time Text-over-IP. For instance, the conversation might be established using voice and the user could elect to switch to text, or add text, during the conversation. Systems supporting real-time Text-over-IP MUST allow users to select any of the supported conversation modes at any time, including mid-conversation.

Systems SHOULD allow the user to specify a preferred mode of communication, with the ability to fall back to alternatives that the user has indicated are acceptable.

If the user requests simultaneous use of text and voice, and this is not possible either because the system only supports alternate modalities or because of resource management on the network, the system MUST try to establish a text-only communication. and the user MUST be informed of this change throughout the process, either in text or in a combination of modalities that MUST include text.

Session setup, especially through gateways to other networks, MAY require the use of specially formatted addresses or other mechanisms for invoking gateways.
Such mechanisms MUST be supported by the terminal.

8.2.3 Addressing

The SIP [3] addressing schemes MUST be used for all entities. For example SIP URL and Tel URL will be used for caller, called party, user devices, and servers (e.g., SIP server, Transcoding server).

The right to include a transcoding service MUST NOT require user registration in any specific SIP registrar, but MAY require authorisation of the SIP registrar in the service.

8.2.4 Alerting

Systems supporting real-time Text-over-IP MUST have an alerting method (e.g., for incoming calls) that can be used by deaf and hard of hearing people or provide a range of alternative, but equivalent, alerting methods that are suitable for all users, regardless of their abilities and preferences.

It should be noted that general alerting systems exist, and one common interface for triggering the alerting action is a contact closure between two conductors.

Among the alerting options are alerting on the user equipment and specific alerting user agents registered to the same registrar as the main user agent.

If present, identification of the originating party (for example in the form of a URL or CLI) MUST be clearly presented to the user in a form suitable for the user BEFORE answering the request. When the invitation to initiate a conversation involving real-time Text-over-IP originates from a gateway, this MAY be signalled to the user.

8.2.5 Call Negotiations

The Session Description Protocol (SDP) used in SIP [3] provides the capabilities to indicate ToIP as a media in the call setup. RFC 2793 [5] provides the RTP payload type text/t140 for support of ToIP which can be indicated in the SDP as a part of SDP INVITE, OK and SIP/200/ACK for media negotiations. In addition, SIPÆs offer/answer model can also be used in conjunction with other capabilities including the use of a transcoding server for enhanced call negotiations [7,8,9].

8.2.6 Answering

Systems SHOULD provide a best-effort approach to answering invitations for session set-up and users should be kept informed at all times about the progress of session establishment. On all systems that both inform users of session status and support real-time Text-over-IP, this information MUST be available in text, and may be provided in other visual media.

8.2.6.1 Answering Machine

Systems for real-time Text-over-IP MAY support an auto-answer function, equivalent to answering machines on telephony networks. If an answering machine function is supported, it MUST support at least 160 characters for the greeting message. It MUST support incoming text message storage of a minimum of 16000 characters, although systems MAY support much larger storage.

When the answering machine is activated, user alerting MUST still take place. The user MUST be allowed to monitor the auto-answer progress and MUST be allowed to intervene during any stage of the answering machine and take control of the session.

8.2.7 Session progress and status presentation

During a conversation that includes real-time Text-over-IP, status and session progress information MUST be provided in text. That information MUST be equivalent to session progress information delivered in any other format, for example audio. Users MUST be able to manage the session and perform all session control functions based on the textual session progress information.

The user MUST be informed of any change in modalities.

Session progress information MUST use simple language as much as possible so that it can be understood by as many users as

possible.
The use of jargon or ambiguous terminology SHOULD be avoided at all times. It is RECOMMENDED to let text information be used together with icons symbolising the items to be reported.

There MUST be a clear indication, both visually as well as audibly whenever a session gets connected and disconnected. The user should never be in doubt as to what the status of the connection is, even if he/she is not able to use audio feedback or vision.

8.2.8 Actions During Calls

Certain actions need to be performed for the ToIP conversation during the call and these actions are describe briefly as follows:

a. Text transmission SHALL be done character by character as entered, or in small groups transmitted so that no character is delayed between entry and transmission by more than 300 milliseconds.
b. The text transmission SHALL allow a rate of at least 30 characters per second so that human typing speed as well as speech to text methods of generating conversation text can be supported.

c. After text connection is established, the mean end-to-end delay of characters SHALL be less than two seconds, measured between two ToIP users. This requirement is valid as long as the text input rate is lower or equal to the text reception and display rate.

d. The character corruption rate SHALL be less than 1% in conditions where users experience the quality of voice transmission to be low but useable. This is in accordance with ITU-T F.700 Annex A.3 quality level T1.

e. When interoperability functions are invoked, there may be a need for intermediate storage of characters before transmission to a device receiving slower than the typing speed of the sender. Such temporary storage SHALL be dimensioned to adjust for receiving at 30 characters per second and transmitting at 6 characters per second during at least 4 minutes [less than 3k characters].

f. If text is detected to be missing after transmission, there SHALL be an indication in the text marking the loss.

g. When used from a terminal designed for PSTN text telephony, or in interworking with such a terminal, ToIP shall enable alternating between text and voice in a similar manner as the PSTN text telephone handles this mode of operation. (This mode is often called VCO/HCO in USA).

h. The transmission of the text conversation SHALL be made according to an internationally suitable character set and control protocol for text conversation as specified in ITU-T T.140.

i. When display of the conversation on end user equipment is included in the design, display of the dialogue SHALL be made so that it is easy to read text belonging to each party in the conversation.

8.2.8.1 Text and other Media Handling Between ToIP Devices

The ToIP devices do not need transcoding from speech to text and can communicate directly using text/t140. The following requirements are valid for media handling during calls:

a. When used between terminals designed for ToIP, it SHALL be possible to send and receive text simultaneously with the other media (text, audio and/or video) supported by the same terminals.

b. When used between terminals designed for ToIP, it SHALL be possible to send and receive text simultaneously.

c. It should be possible to know during the call that ToIP is available, even if it is not invoked at call setup (only voice and/or video is used). To disable this, the user must disable the use of ToIP.

8.2.8.2 General Actions

a. It SHALL be possible to establish a session with text capabilities enabled at the beginning of a Call. Note: a call is in this document defined as one or more sessions).

b. It SHALL be possible to place a call without text capabilities, and to add text capabilities later in the call.

c. It SHALL be possible to transfer text at at least 30 characters per second

d. It SHALL be possible to talk and listen simultaneously with typing and reading.

8.2.8.3 Call Action with Native ToIP Devices

a. It SHOULD be possible to answer a call with text capabilities enabled.

b. It SHOULD be possible to use video simultaneously with the other media in the call.

c. It SHOULD be possible to answer a call in voice or video without text enabled, and add text later in the call.

d. It MUST be possible to disconnect the call.

e. It SHOULD be possible to control IVR (Interactive Voice Response) services from a numeric keypad.

f. It SHOULD be possible to control ITR (Interactive Text Response) services from the alphanumeric keyboard.

g. It SHOULD be possible to invoke multi-party calls.

h. It SHALL be possible to transfer the call.

i. It MUST be possible to use text characters (numbers) instead of DTMF tones (numbers) in interactions where the person is using a keyboard to interact with a service and the service asks for a number.

8.2.8.4 Audio/Visual/Tactile Indicators

It SHOULD be possible to observe visual or tactile indicators about:
- Call progress
- Availability of text, voice and video channels.
- Incoming call.
- Incoming text.
- Typed and transmitted text.
- Any loss in incoming text.

8.2.9 Additional session control

Systems that support additional session control features, for example call waiting, forwarding, hold etc on voice calls, MUST offer equivalent functionality for real-time Text-over-IP functions. In addition, all these features MUST be controllable by text users at any time, in an equivalent way as for other users.

8.2.10 File storage

Systems that support real-time Text-over-IP MAY save the text conversation to a file. This SHOULD be done using a standard file format.

8.3 Conference Call Requirements

The conference call requirements deal with multipoint conferencing calls where there will be at least one or more ToIP capable devices along with other end user devices where the total number end user devices will be at least three.

It SHOULD be possible to use the text medium in conference calls, in a similar way as video is handled and displayed. Text in conferences can be used both for letting individual participants use the text medium, and for central support of the conference with real time text interpretation of speech.

8.4 Transport

ToIP uses RTP as the default transport protocol for transmission of real-time text medium text/t140 as specified in RFC 2793 [5]. Signaling and other media will use the transport protocol

specified in SIP [3] and/or their revised versions as specified in standards.

The redundancy method of RFC 2198 SHOULD be used for making text transmission reliable with transmission of three generations.

Text capability SHOULD be announced in SDP by a declaration in line with this example:

```
m=text 11000 RTP/AVP 98 100
a=rtpmap:98 t140/1000
a=rtpmap:100 red/1000
a=fmtp:100 98/98/98
```

Characters SHOULD BE buffered for transmission and transmitted every 300 ms.

By having this single coding and transmission scheme for real time text defined, in the SIP call control environment, the opportunity for interoperability is optimized.

However, if good reasons exist, other transport mechanisms MAY be offered and used for the T.140 coded text, provided that proper negotiation is introduced, and RFC 2793 transport MUST be used as the defaut fallback solution.

8.5 Character Set

a. Real-Time Text-over-IP protocols MUST use UTF-8 encoding as specified in ITU-T T.140 [12].

b. Real-time Text-over-IP SHOULD handle characers with editing effect such as new line, erasure and alerting during session as specified in ITU-T T.140.

8.6 Transcoding

Transcoding of text may need to take place in gateways between ToIP and other forms of text conversation. ToIP makes use of ISO 10646 character set.
Most PSTN textphones use a 7-bit character set, or a character set that is converted to a 7-bit character set by the V.18 modem.

When transcoding between these character sets and T.140 in gateways, special consideration MUST be paid to the national variants of the 7 bit codes, with national characters mapping into different codes in the ISO 10 646 code space. The national variant to be used SHOULD be possible to select by the user per call, or be configured as a national default for the gateway.

The missing text indicator in T.140, specified in T.140 amendment 1, cannot be represented in the 7 bit character codes. Therefore

these characters SHOULD be translated to be represented by the ' (apostrophe) character in legacy text telephone systems where this character exists. For legacy systems where the character ' does not exist, the character . (full stop) SHOULD be used instead.

8.7 Relay Services

The relay service acts as an intermediary between 2 or more callers.
The basic relay service allows a translation of speech to text and text to speech, which enables hearing and speech impaired callers to communicate with hearing callers. Even though this document focuses on ToIP, we do not exclude video relay services for e.g., speech to sign language and vice versa and other possible relay services. It will be possible to use ToIP simultaneously with other relay services if desired.

It is very important for the users that a relay session is invoked as transparently as possible. It SHOULD happen automatically when the call is being set-up or by a simple user action. A transcoding framework document using SIP [7] describes invoking relay services, where the relay acts as a conference bridge or uses the third party control mechanism.

Adding or removing a relay service MUST be possible without disrupting the current call.

When setting up a call, the relay service MUST be able to determine the type of service requested (e.g. speech to text or text to speech), to indicate if the caller wants voice carry over, the language of the text including the sign language being used.

The user MUST be provided with a method to indicate which service is desired.

Relay services MUST be reachable all the time, even if the users are visiting networks from different operators.

It SHOULD be possible to route the call to a preferred relay service even if the user makes the call from another region or network than usually used.

It MUST be possible to identify ToIP sessions as emergency sessions.

If it is decided that a relay service supports emergency calls, the relay service operator MUST be able to process such a session correctly and quickly with the following functionality:

a. The relay service operatorÆs network MUST give priority to this incoming call.

b. The relay service operator MUST forward this session if they are unable to process it to an alternative emergency relay operator.

c. The relay service MUST label the transcoded stream as an emergency call (in case of text to speech and/or vice versa).

d. The relay service MUST provide all session information to the emergency centre (e.g., location information of the caller if available).

8.8 Emergency services

a. It MUST be possible to support emergency service calls with text only or simultaneously with voice.

b. All session information that accompanies a voice session to the emergency centre, MUST also be provided to the emergency center if it is a ToIP session.(e.g, phone number and location information of the user placing the emergency call).

c. A text over IP stream MUST be labelled as an emergency stream to ensure that the emergency service center is able to receive this call.

8.9 User Mobility

ToIP terminals SHOULD use the same mechanisms as other terminals to resolve mobility issues. It is RECOMMENDED to use a SIP-adress for the users, resolved by a SIP REGISTRAR, to enable basic user mobility. Further mechanisms are defined for the 3G IP multimedia systems.

8.10 Confidentiality and Security

User confidentiality and privacy need to be met as described in SIP [3]. For example, nothing should reveal the fact that the user of ToIP is a person with a disability unless the user prefers to make this information public. If a transcoding server is being used, this SHOULD be transparent. Encryption SHOULD be used on end-to-end or hop-by-hop basis as described in SIP [3] and SRTP [19]

Authentication needs to be provided for users in addition to the message integrity and access control.

Protection against Denial-of-service (DoS) attacks needs to be provided considering the case that the ToIP users might need transcoding servers.

8.11 Call Scenarios

ToIP is a way of establishing the real-time conversation. Call flow for ToIP MUST be similar to session establishment with audio and video. For example, ToIP services MAY be invoked in the following situations (among others):

- Noisy environment (e.g., in a machine room of a factory where listening is difficult)Busy with another call and want to participate in two calls at the same time.

- Text and/or speech recording services (e.g., text documentation/audio recording for legal/clarity/flexibility purposes)

- Overcoming of language barriers through speech translation and/or transcoding services

- Not hearing well or not at all (e.g., hearing loss due to aging, hard of hearing, deaf)

NOTE: In many of the above scenarios, text may accompany speech in a subtitling like fashion. This would occur for individuals who are hard of hearing and also for mixed calls with a hearing and deaf person listening to the call.

All call flows either for the point-to-point or for the multipoint situation need to consider that ToIP services may be invoked for many different reasons by users as explained. When the transcoding/translation services are needed, call flows will be shown for both session establishment models: Third-party call control model and Conferencing bridge model.

8.11.1 Call Scenarios

There are 2 different terminal types possible:

1. The terminal itself has the intelligence to initiate a relay service for incoming and outgoing calls (based on address book, user preferences programmed on the terminal etc. This terminal can be used in a conference bridge call as well as a third party control call.

2. Dumb terminals, so that the relay service server actually initiates the correct call handling (the dumb terminal can only REFER the call to the relay center, which then sets up the call using the conference bridge flow.).

The following call scenarios are shown:

- Communications between two ToIP/Multimedia capable, end user devices using the same language.

- Communications between ToIP capable, end user devices using translation services to provide language translation.

- Communications between ToIP/Multimedia capable and Audio (non-ToIP) capable end user devices.

- Communications between ToIP/Multimedia and/or Audio (non-ToIP)/Multimedia end user devices maintaining privacy.

8.11.2 Point-to-Point Call Scenarios

The point-to-point call scenarios will contain at least one or both ToIP/Multimedia devices in setting up the session. The detail call scenarios will include:

- ToIP/Multimedia devices that use the same language.

- ToIP/Multimedia devices invoke translation services for using different languages.
  * Third-party call control model.
  * Conference bridge service model.

- ToIP/Multimedia devices invoke translation services for using different languages maintaining privacy.
  * Third-party call control model.
  * Conference bridge service model.

- ToIP/Multimedia device and Audio (non-ToIP)/Multimedia device invoking transcoding server.
  * Call initiated by Audio (non-ToIP)/Multimedia user
    - Third-party call control model.
    - Conference bridge service model.
  * Call initiated by ToIP user.
    - Third-party call control model.
    - Conference bridge service model.

- ToIP/Multimedia device and Audio (non-ToIP)/Multimedia device invoking transcoding server maintaining privacy.
  * Call initiated by Audio (non-ToIP)/Multimedia user
    - Third-party call control model.
    - Conference bridge service model.
  * Call initiated by ToIP user.
    - Third-party call control model.
    - Conference bridge service model.

8.11.3 Conference Call Scenarios

The conference call scenarios only contain the multipoint communications, and only the centralized bridge model is considered. The following multipoint conference call scenarios will contain at least one more ToIP/Multimedia devices:

- ToIP/Multimedia devices that use the same language.

- ToIP/Multimedia devices invoke translation services for using different languages.

- ToIP/Multimedia devices invoke translation services for using different languages maintaining privacy.

- ToIP/Multimedia device and Audio (non-ToIP)/Multimedia device invoking transcoding server.
   * Call initiated by Audio (non-ToIP)/Multimedia user.
   * Call initiated by ToIP/Multimedia user.

- ToIP/Multimedia device and Audio (non-ToIP)/Multimedia device invoking transcoding server maintaining privacy.
   * Call initiated by Audio (non-ToIP)/Multimedia user.
   * Call initiated by ToIP/Multimedia user.

9. Interworking Requirements for Text-over-IP

A number of systems for real time text conversation already exist as well as a number of message oriented text communication systems. Interoperability is of interest between ToIP and some of these systems. This section describes requirements on this interoperability.

9.1 Real-Time Text-over-IP Interworking Gateway Services

Interactive texting facilities exist already in various forms and on various networks. On the PSTN, it is commonly referred to as text telephony. The simultaneous or alternating use of voice and text is used by a large number of users who can send voice, but must receive text or who can hear but must send text due to a speech disability.

9.2 Text-over-IP and PSTN/ISDN Text-Telephony

On PSTN networks, transmission of interactive text takes place using a variety of codings and modulations, including ITU-T V.21 [II], Baudot, DTMF, V.23 [III] and others. Many difficulties have arisen as a result of this variety in text telephony protocols and the ITU-T V.18 [10] standard was developed to address some of these issues.

ITU-T-V.18 [10] offers a native text telephony method plus it defines interworking with current protocols. In the interworking mode, it will recognise one of the older protocols and fall back to that transmission method when required.

In order to allow systems and services based on Real-time Text-over-IP to communicate with PSTN text telephones, text gateways are the recommended approach. These gateways MUST use the ITU-T V.18 [10] standard at the PSTN side.

Buffering MUST be used to support different transmission rates. At least 1K buffer MUST be provided. A buffer of at least 2K characters is recommended. In addition, the gateway MUST provide a minimum throughput of at least 30 characters/second or the highest speed supported by the PSTN text telephony protocol side, whichever is the lowest.

PSTN-Real-time Text-over-IP gateways MUST allow alternating use of text and voice.

PSTN and ISDN to real-time Text-over-IP gateways that receive CLI information from the originating party MUST pass this information to the receiving party as soon as possible.

Priority MUST be given to calls labeled as emergency calls.

9.3 Text-over-IP and Cellular Wireless circuit switched Text-Telephony

Cellular wireless (or Mobile) circuit switched connections provide a digital real-time transport service for voice or data.
The access technologies include GSM, CDMA, TDMA, iDen and various 3G technologies.

Alternative means of transferring the Text telephony data have been developed when TTY services over cellular was mandated by the FCC in the USA. They are a) "No-gain" codec solution, b) the Cellular Text Telephony Modem (CTM) solution and c) "Baudot mode" solution.

The GSM and 3G standards from 3GPP make use of the CTM modem in the voice channel for text telephony.
However, implementations also exist that use the data channel to provide such functionality. Interworking with these solutions SHOULD be done using text gateways that set up the data channel connection at the GSM side and provide real-time Text-over-IP at the other side.

9.3.1 "No-gain"

The "No-gain" text telephone transporting technology uses specially modified EFR [15] and EVR [16] speech vocoders in both mobile terminals used provide a text telephony call. It provides full duplex operation and supports alternating voice and text.("VCO/HCO"). It is dedicated to the CDMA and TDMA mobile technologies and the US Baudot type of text telephones.

9.3.2 Cellular Text Telephone Modem (CTM)

CTM [17] is a technology independent modem technology that provides the transport of text telephone characters at up to 10 characters/sec using modem signals that are at or below 1 kHz and uses a highly redundant encoding technique to overcome the fading and cell changing losses. On any interface that uses analog transmission, half-duplex operation must be supported as the "send" and "receive" modem frequencies are identical. The use of

CTM may have to be modified slightly to support half-duplex operation.

9.3.3 "Baudot mode"

This term is often used by cellular terminal suppliers for a GSM cellular phone mode that allows TTYs to operate into a cellular phone and to communicate with a fixed line TTY.

9.3.4 Data channel mode

Many mobile terminals allow the use of the data channel to transfer data in real-time. Data rates of 9600 bit/s are usually supported on the mobile connection.Gateways or the interworking function provides interoperability with PSTN textphones.

9.3.5 Common Text Gateway Functions

Text Gateways MUST cover the differences that result from different text protocols. The protocols to be supported will depend on the service requirements of the Gateway.

Different data rates of different protocols MAY require text buffering.

Interoperation of half-duplex and full-duplex protocols MAY require text buffering and some intelligence to determine when to change direction when operating in half-duplex.

Identification may be required of half-duplex operation either at the "user" level (ie. users must inform each other) or at the "protocol" level (where an indication must be sent back to the Gateway).

A Text Gateway MUST be able to route text calls to emergency service providers when any of the recognised emergency numbers that support text communications for the country or region are called eg. "911" in USA and "112" in Europe.

A text gateway MUST act transparently on the IP side. It acts then as a virtual end-point terminal.

9.4 Text-over-IP and Cellular Wireless Text-over-IP

Text-over-IP MAY be supported over the cellular wireless packet switched service. It interfaces to the Internet. For 3GPP 3G services, the support is described to use ToIP in 3G TS 26.235 [20].

A Text gateway with cellular wireless packet switched services MUST be able to route text calls into emergency service providers when any of the recognized emergency numbers that support text communication for the country are called.

9.5 Instant Messaging Support

Instant Messaging is used by many people to communicate using text via the Internet. Instant Messaging transfers blocks of text rather than streaming as is used for real-time Text-over-IP. As such, it is not a replacement for real-time Text-over-IP and in particular does not meet the needs for real time conversations of deaf, hard of hearing and speech-impaired users. It is unsuitable for communications through a relay service [I]. The streaming character of real-time Text-over-IP provides a better user experience and, when given the choice, users often prefer real-time Text-over-IP.

However, since some users might only have Instant Messaging available, text gateways might be developed that allow interworking between Instant Messaging systems and real-time Text-over-IP solutions.

Because Instant Messaging is based on blocks of text, rather than on a continuous stream of characters, such gateways need to transform between these two formats. Text gateways for interworking between Instant Messaging and real-time Text-over-IP MUST concatenate individual characters originating at the real-time Text-over-IP side into blocks of text and:

a. When the length of the concatenated message becomes longer than 50 characters, the buffered text MUST be transmitted to the Instant Messaging side as soon as any non-alphanumerical character is received from the real-time Text-over-IP side.

b. When a new line is received from the real-time Text-over-IP side, the buffered characters up to that point, including the carriage return and/or line feed characters, MUST be transmitted to the Instant Messaging side.

c. When the real-time Text-over-IP side has been idle for at least 5 seconds, all buffered text up to that point MUST be transmitted to the Instant Messaging side.

It is recommended that during the session, both users are constantly updated on the progress of the text input.
For example, many Instant Messaging protocols signal that a user is typing to the other party in the conversation. Text gateways between Instant Messaging and real-time Text-over-IP MUST provide this signaling to the Instant Messaging side when characters start being received, or at the beginning of the conversation.
Also at the real-time text-over-IP side, an indicator of writing the Instant Message MUST be present. For example, the real-time text user will see . . . waiting for replying IM. . . And per 5 seconds a . (dot) can be shown.
Those solutions will reduce the difficulties between a streaming versus blcoked text.

Even though that the text gateway can connect Instant Messaging and real-time Text-over-IP. The best solution is to take advantage of the fact that the user interfaces and the user communities for instant messaging and real-time text-over-IP telephony are extremely similar.

After all, the character input, the character display, Internet connectivity, SIP stack, etc are the same for Instant Messaging and real-time Text-over-IP.

Devices that implement Instant Messaging SHOULD implement real-time text-over-IP telephony, using standard SIP and text/t140 mechanisms.

9.6 IP Telephony with Traditional RJ-11 Interfaces

Analogue adapters using SIP based IP communication and RJ-11 connectors for connecting traditional PSTN devices SHOULD enable connection of legacy PSTN text telephones [18]. These adapters SHOULD contain V.18 modem functionality, voice handling functionality, and conversion functions to/from SIP based ToIP with T.140 transported according to RFC 2793, in a similar way as it provides interoperability for voice calls. If a call is set up and text/t140 capability is not declared by the endpoint (by the end-point terminal or the text gateway in the network at the end-point), a method for invoking a transcoding server shall be used. If no such server is available, the signals from the textphone MAY be transmitted in the voice channel as audio with high quality of service.
NOTE: It is preferred that such analogue adaptors do use RFC2793 on board and thus act as a text gateway. Sending textphone signals over the voice channel is undesirable due posible filtering and

compression and packet loss between the end-points. Which can result in dropping characters in the textphone conversation or even not allowing the textphones to connect with each other.

9.7 Interworking Call Flows

The call scenarios in chapter 8.11 deal with end to end ToIP. These call flows do not change on the IP side of the network when one end-point is actually a text gateway. The text gateway actually acts like a ToIP/Multimedia device. Separate call flows will show the interworking between the ToIP/Multimedia devices [4] over the IP network and the text telephony devices [10] over the PSTN/ISDN network using the IP-PSTN/ISDN interworking functional (IWF) entity. It is assumed that the IWF will provide ToIP and text telephony interworking in addition to other capabilities. Thus acting as a Text gateway.

The point-to-point call flows will contain at least one ToIP/Multimedia and one text telephony/multimedia (or POTS) device for the following cases:

- ToIP/Multimedia device and text telephony/multimedia device that use the same/different language.
- ToIP/Multimedia device and PSTN/ISDN-based POTS/Multimedia device.

For multipoint conferencing calls, it is assumed that only the centralized conferencing will be considered, and the media bridge is supposed to be located somewhere in the SIP network. However, it is considered that the ToIP and text telephony interworking function will be located in the IWF.

The multipoint conference call flows will contain at least one ToIP/Multimedia device, at least one text telephony/multimedia device, and other devices where total number of devices will be three or more for the following cases:

- ToIP/Multimedia and text telephony/multimedia devices that use the same/different language.
- ToIP/Multimedia devices, telephony/multimedia devices, and/or PSTN/ISDN-based POTS/Multimedia devices.

9.8 Multi-functional gateways

The scenarios described in this document deal with single pairs of interworking protocols or services. However, in practice many of these interworking systems will be implemented as gateways that combine different functions. As such, a text gateway could be build to have modems to interwork with the PSTN and support both Instant Messaging as well as real-time ToIP. Such interworking functions are called Combination gateways.

Combination gateways MUST provide interworking between all of their supported text based functions. For example, a text gateway that has modems to interwork with the PSTN and that support both Instant Messaging and real-time ToIP MUST support the following interworking functions:

- PSTN text telephony to real-time ToIP.
- PSTN text telephony to Instant Messaging.
- Instant Messaging to real-time ToIP.

9.9 Gateway Discovery

To get a smooth invocation of the text gateways, where those gateways are transparant on the IP side, it requires a method how and when to invoke the text gateway. As described previously in this draft. The text gateways must act as the end-terminal. The capabilities of the text gateway will in that call be determined by the call capabilities of the terminal that is using the gateway. For example, a PSTN textphone is only able to receive

voice and streaming text. Thus the text gateway will only allow ToIP and audio.

The PSTN devices or other non IP multimedia devices that require the text gateways to connect to the IP must be able to locate the text gateway, and ensure that the correct call capabilities of the non IP multimedia device is used by the text gateway.

The following possible solutions for using the text gateway are:

- PSTN Textphone users using a prefix number before dialing out.
- In band text dialogue, where the gateway asks the user for the destination address.
- separate text subscriptions, linked to the phone number or terminal identifier/ IP address.
- text capability indicators.
- text preference indicator.
- listen for text activity in all calls.
- call transfer request by the called user.
- placing a call via the web, and use one of the methods described here
- text gateways with its own telephone number and/or SIP address. (this requires user interaction with the text gateway to place a

call).
- ENUM address analysis and number plan
- number or address analysis leads to the gateway for all PSTN calls.
- etc

9.10 Text Gateway in the call Scenarios

9.10.1 IP terminal calling an analogue textphone (PSTN)

The ToIP stream will be converted into an analogue text telephone protocol (using the voice channel) and vice versa by the text gateway.

The PSTN knows that it may be a textphone call thanks to the SDP description (for example: m=text 11000 RTP/AVP 98 a=rtpmap:98 t140/1000 for T.140 text on port 11000). It can then activate text gateway functions on the PSTN side listening for a text answer.

The PSTN will also know that all those incoming calls are only for analogue textphones. Thus the speed of the text stream is adjusted to the selected analogue textphone protocol.
If there is no analogue textphone on the called number, the call setup will be terminated by the text gateway.

The text gateway can be implemented in two ways: The PSTN has its own text gateway (the IWF), or it redirects the media stream to the nearest IP-PSTN gateway with text transcoding abilities.

Text gateway detection: In the SIP messages.

9.10.2 IP terminal calling a mobile text telephone (CTM)

The ToIP stream will be converted into CTM and vice versa by the text gateway located in the network of the cellular/mobile operator. It is similar to the PSTN.

Text gateway detection: In the SIP messages.

9.10.3 IP terminal calling a mobile telephone (GPRS based)

A text gateway located in the mobile network converts the incoming T.140/RTP stream into for example T.140 over TCP (T.140/TCP) or tunnels the T.140 stream over HTTP (T.140/HTTP). Or any other temporarily non standard solution necessary to connect the text gateway with the text telephone client on the mobile phone.

This is necessary, since RTP over GPRS is not possible in many

mobile phones.
Note, those server-client solutions are ONLY acceptable for the GPRS and non RTP stack phones. It is encouraged to use T.140/RTP as soon as possible for all mobile phones.
Allowing UDP transport over the GPRS link will enable RFC2793 text over GPRS.

Text gateway detection: In the SIP messages.

9.10.4 IP terminal calling a mobile telephone(UMTS)

No text gateway is required here since this will be end to end IP.

9.10.5 Analogue textphone (PSTN) user calling an IP terminal using prefix

The PSTN is unable to distinguish between an analogue voice call and an analogue textphone, both use the voice channel. The text gateway needs to transcode the analogue textphone protocol into T.140/RTP.

One way for a PSTN to separate an incoming voice call into text telephony or normal voice is by using a prefix number for all incoming text telephone calls to the PSTN. For example , the text telephone user (e.g Boudot) places a call and enters a prefix e.g. 600 and then continues with the original number. The PSTN will recognize all incoming 600 calls as an analogue textphone call and redirects the call to a text gateway (unless it is a number connecting the same PSTN).

It is undesirable to allow a PSTN to transport all the analogue textphone tones/signals through a VoIP stream! (In band text dialogue).

Text gateway detection: Prefix number for incoming textphone calls.

9.10.6 Mobile text telephone (CTM) user calling an IP terminal

The voice channel of the cellular network is used. The MSC is able to separate between the text call and voice only, it is just a matter of redirecting the voice channel to the text gateway.

Text gateway detection: CTM signal detection.

9.10.7 Mobile telephone user (GPRS) calling an IP terminal

The text telephone client on the mobile telephone connects the text gateway located in the network. The text gateway transcodes the text stream into ToIP.

Text gateway detection: pre-programmed in the mobile textphone client.

9.10.8 Mobile telephone (UMTS) user calling an IP terminal

No text gateway is required here since this will be end to end IP.

9.10.9 Voice over DSL user using an analogue text telephone.

Voice over DSL is a widespread service. When connecting analogue text telephones to this service there is a risk that they just use the voice channel that result in corrupted text transmission. The VoDSL gateway located in the network of the (A)DSL operator itself should connect with a text gateway as soon it turns into VoIP.

Text gateway detection: prefix number similar to the PSTN.

9.10.10 VoIP user via a building telephone switch (at an apartment building) owning an analogue text telephone.

This is the case where only VoIP is possible and no other IP traffic between the telephone switch and the apartments.
The only solution would be a forced analogue text telephone protocol over the Voice channel, in band text dialogue . If that must happen. Then the telephone switch MUST convert the analogue text telephone protocol into ToIP and vice versa before the telephone switch connects the IP network.
Note: The in band text dialogue is undesirable. This scenario SHOULD be avoided at any cost.

Text gateway detection: prefix number or in band text signalling.

9.10.11 VoIP user via a gateway/box connected to his/her own Broadband connection owning an analogue text telephone.

The gateway box should natively transcode analogue text telephony into ToIP and vice versa when an analogue text phone is plugged in the RJ-11 socket [18].

Text gateway detection: RJ-11 socket preconfigured by the box via jumpers or software, or listen for textphone tones and perform V.18 text telephone detection.

10. Terminal Features

Implementers of products that support interactive Text-over-IP SHOULD NOT assume that all users of text are able to use mainstream input and output devices. People with arthritis or other dexterity problems might not be able to use very small keyboards. Visually impaired people might not be able to use standard sized characters on a display. Colour-blind people might suffer from badly chosen colour-schemes. People with motor disabilities might require specialised input devices.

Implementers SHOULD make their products as open as possible with regard to this wide range of abilities and preferences and they MUST use standard interfaces wherever they provide such interfaces.

10.1 Text input

Systems that support real-time interactive Text-over-IP SHOULD support suitable input mechanisms, either built-in or connectable through the use of a standard interface: PS/2, USB, Bluetooth, or virtual keyboard. In particular Braille users should be able to connect Braille keyboards to the terminal. Terminals MAY support a web interface for input and output of text.

It is recommended that systems that fixed terminals that support real-time interactive Text-over-IP allow the user to enter the standard alphanumerical characters directly, rather than through a cycle of key presses or other indirect means. This could be done using full-sized keyboards, smaller sized keyboards or fastap keyboards for example. It is highly recommended to provide a standard interface to allow attachment of an external input device, especially for terminals that have only limited input systems built-in.

Systems should provide means to add voice-to-text translation as text input.

All IP phones with a display of 12 or more characters MUST support at least text input through the regular phone keypad (and display of any incoming text) in order to provide basic emergency text communication from any IP phone.

Input devices that have automatic key repeat MUST allow the user to specify the key-repeat rate.

10.2 Text presentation

Systems that support real-time interactive Text-over-IP SHOULD support suitable displays, either built-in or connectable through the use of a standard interface: S-VGA, USB, Bluetooth or IP. Braille readers should be connectable to the terminal using a standard interface.

Terminals MAY support a web interface for input and output of text.

A variety of handsets and terminals might be developed for a number of equally varied scenarios.

In the case of fixed terminals or software applications on Personal Computers, implementers MUST:

a. Use either separate screen areas for displaying sent and received text OR clearly indicate the difference between sent and received text. Systems MAY allow the user to chose either on of these presentation methodologies.

b. Provide at least 5 lines of 35 monospaced characters each for each direction (sent and received text) OR at least 10 lines of 35 characters when sent and received text are presented together.

In the case of Mobile terminals, implementers MUST:

c. Use either separate screen areas for displaying sent and received text OR clearly indicate the difference between sent and received text. Systems MAY allow the user to chose either on of these presentation methodologies.

d. Provide at least 3 lines of 20 monospaced characters each for each direction (sent and received text) OR at least 6 lines of 20 characters when sent and received text are presented together.

On both types of terminals, scrolling back through both sent and received text MUST be supported, even after the conversation has ended. Lines SHOULD be wrapped at word boundaries .

There MUST be an easy-to-use function to clear the screen at any time during the session, and if the implementation has chosen to present sent and received text separately, clearing the screen SHOULD be possible as a separate function for sent and received text.

The function of the new line and erasure controls as explained in section 9.5. MUST be supported by the presentation in the consistent way described by T.140. Presentation layers MUST support the full UTF-8 character set.

When real-time Text-over-IP is used in conjunction with other modalities, like voice, the presentation MUST clearly indicate this to the user in an area outside the display region for send and received text.

Identification information for other parties in the conversation, like URLÆs, user-friendly names from an address book, or CLI in the case of conversations with text telephones, SHOULD be displayed throughout the entire conversation in a region outside the sent and received text area.

10.3 Call control

Call (Session) Control procedures MUST use the SIP protocol. Text sessions MUST be identified in accordance with requirements described earlier.

Text services SHOULD be part of a Total Conversation environment in which voice, text and video sessions can be added, modified or deleted individually.

To enable interworking with Textphones in telephone and cellular (mobile) networks, terminals MUST be able to access Gateways automatically when a PSTN or cellular (mobile) E.164-based telephone number is used as the called address.

Users MUST be able to establish text sessions to emergency service providers using the widely recognised emergency numbers in use in the country or region of operation of the terminal eg. æ911Æ in USA and ³112³in Europe.

The ability to transfer Location information SHALL be provided if the information is available from the terminal.

10.4 Device control

ToIP devices shall support multiple means of setting up and performing calls as well as controlling the device itself. The built-in controls and presentation systems shall take accessibility aspects into account as far as possible. The device shall include external interfaces that makes it possible to attach user interface devices for people with needs beyond what the built-in user interface can support. It is preferrable if such external interfaces are wireless.

10.5 Alerting

The form of Alerting indication(s) provided to the user should be selectable to suit particular users. Alerting indications MAY include Sound, Tactile (eg. vibrational), Visual (on-screen symbols; separate flashing light), Motion (eg. movement of something).

The ability to send an Alerting signal to an external interface SHOULD be provided. This will allow Alerting devices that are specific to users requirements to be attached.

As many as possible of the following alternatives for alerting SHOULD be provided:

* Internal flash.
* Two-pole connector for external alerting systems triggered by contact between the two poles when a ring signal is generated (if necessary with 1.5-9 V battery power for alerting systems requiring electrical currents to activate).
* Bluetooth serial profile with AT command interface, sending the "RING" message, intended for a Bluetooth alerting receiver with flash, vibration or sound action.
* SIP connected alerting device, that get its stimuli by being registered on the same sip address as the terminal.

10.6 External interfaces

Terminals for ToIP SHOULD provide external interfaces for the following functions:

* Text input.
* Text display.
* Terminal control.
* Session control.

10.7 Power

As terminals could remain active for very long periods of time, the electrical power requirements of all the terminals SHOULD be as low as possible.

If the terminal is to be used for calling Emergency services or where the mains power supply is unreliable, back-up power systems SHOULD be provided for the terminal and all equipment used to provide the ToIP service. This can be implemented in many different ways eg. via the line powering option on some Ethernet interfaces, or by using a "no break" power supply (a battery back-up system with inverters that can recreate a limited amount of mains power).

11. Security Considerations

There are no additional security requirements other than described earlier.

12. Outstanding issues

A number of outstanding issues yet need to be resolved. This is possible in this draft, or in a separate draft.

- Call flows diagrams based on the scenarios discussed in this draft.
- Service labelling of media streams to be able to determine which kind of service the text stream contains. For example, is it english, spanish text? Is it an emergency text stream? Etc.

13. Authors Addresses

The following people provided substantial technical and writing contributions to this document, listed alphabetically:

Barry Dingle
ACIF, 32 Walker Street
North Sydney, NSW 2060 Australia
Tel +61 (0)2 9959 9111
Fax +61 (0)2 9954 6136
TTY +61 (0)2 9923 1911
Mob +61 (0)41 911 7578
Email barry.dingle@bigfoot.com.au

Guido Gybels
RNID, 19-23 Featherstone Street
London EC1Y 8SL, UK
Tel +44(0)20 7294 3713
Txt +44(0)20 7296 8019
Fax +44(0)20 7296 8069
EMail: guido.gybels@rnid.org.uk

Gunnar Hellstrom
Omnitor AB
Renathvagen 2
SE 121 37 Johanneshov
Sweden
Phone: +46 708 204 288 / +46 8 556 002 03
Fax: +46 8 556 002 06

Email: gunnar.hellstrom@omnitor.se

Radhika R. Roy
AT&T
Room C1-2B03
200 Laurel Avenue S.
Middletown, NJ 07748
USA
Phone: +1 732 420 1580
Fax: +1 732 368 1302
Email: rrroy@att.com

Henry Sinnreich

MCI
400 International Parkway
Richardson, Texas 75081
Email: henry.sinnreich@mci.com

Gregg C Vanderheiden
University of Wisconsin-Madison
Trace R & D Center
1550 Engineering Dr (Rm 2107)
Madison, Wi 53706
USA
gv@trace.wisc.edu
Phone +1 608 262-6966
FAX +1 608 262-8848

Arnoud A. T. van Wijk
Viataal (Dutch Institute for the Deaf)
Research & Development
Afdeling RDS
Theerestraat 42
5271 GD Sint-Michielsgestel
The Netherlands.
Email: a.vwijk@viataal.nl

14. Acknowledgements

The authors wish to thank Snowshore for providing the ToIP mailing list, which allows many discussions necessary for this draft.

15. Full Copyright Statement

16. References

16.1 Normative

1. Bradner, S., "The Internet Standards Process -- Revision 3", BCP 9, RFC 2026, October 1996.

2. Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997

3. J. Rosenberg, H. Schulzrinne, G. Camarillo, A. R. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, "SIP: Session Initiation Protocol, RFC 3621, IETF, June 2002.

4. ITU-T Recommendation T.140, "Protocol for Multimedia Application Text Conversation (February 1998) and Addendum 1 (February 2000).

5. G. Hellstrom, "RTP Payload for Text Conversation, RFC 2793, May 2000.

6. G. Camarillo, H. Schulzrinne, and E. Burger, "The Source and Sink Attributes for the Session Description Protocol," IETF, August 2003 û Work in Progress.

7. G.Camarillo, "Framework for Transcoding with the Session Initiation Protocol" IETF august 2003 - Work in progress.

8. G. Camarillo, H. Schulzrinne, E. Burger, and A. van Wijk, "Transcoding Service Invocation in SIP using Third Party Call Control," IETF, September 2004 - Work in Progress.

9. G. Camarillo, "The SIP Conference Bridge Transcoding Model," IETF, August 2003 - Work in Progress.

10. ITU-T Recommendation V.18,"Operational and Interworking Requirements for DCEs operating in Text Telephone Mode," November 2000.

11. "XHTML 1.0: The Extensible HyperText Markup Language: A Reformulation of HTML 4 in XML 1.0", W3C Recommendation. Available at http://www.w3.org/TR/xhtml1.

12. Yergeau, F., "UTF-8, a transformation format of ISO 10646", RFC 2279, January 1998.

13. TIA/EIA/825 "A Frequency Shift Keyed Modem for Use on the Public Switched Telephone Network." (The specification for 45.45 and 50 bit/s TTY modems.)

14. Bell-103 300 bit/s modem.

15. TIA/EIA/IS-823-A "TTY/TDD Extension to TIA/EIA-136-410 Enhanced Full Rate Speech Codec (must used in conjunction with TIA/EIA/IS-840)"

16. TIA/EIA/IS-127-2 "Enhanced Variable Rate Codec, Speech Service Option 3 for Wideband Spread Spectrum Digital Systems. Addendum 2."

17. 3GPP TS26.226 "Cellular Text Telephone Modem Description" (CTM).

18. I. Butcher, S. Lass, D. Petrie, H. Sinnreich, and C. Stredicke, "SIP Telephony Device Requirements, Configuration and Data," IETF, February 2004- Work in Progress.

19 Baugher, McGrew, Carrara, Naslund, Norrman, "The Secure Real-Time Transport Protocol (SRTP)", RFC 3711, March 2004.

20. IP Multimedia default codecs. 3GPP TS 26.235

16.2 Informative

I. A relay service allows the users to transcode between different modalities or languages. In the context of this document, relay services will often refer to text relays that transcode text into voice and vice-versa. See for example http://www.typetalk.org.

II. International Telecommunication Union (ITU), "300 bits per second duplex modem standardized for use in the general switched telephone network". ITU-T Recommendation V.21, November 1988.

III. International Telecommunication Union (ITU), "600/1200-baud modem standardized for use in the general switched telephone network. ITU-T Recommendation V.23, November 1988.

IV. Third Generation Partnership Project (3GPP), "Technical Specification Group Services and System Aspects; Cellular Text Telephone Modem; General Description (Release 5)". 3GPP TS 26.226 V5.0.0,