SIPPING Workgroup
Internet-Draft
Category: Informational
Expires: March 6 2006

A. van Wijk (editor)
Viataal

September 7 2005

Framework of requirements for real-time text conversation using SIP

draft-ietf-sipping-toip-03.txt

Abstract

This document provides the framework of requirements for real-time character-by-character interactive text conversation over the IP network using the Session Initiation Protocol and the Real-Time Transport Protocol. It discusses requirements for real-time Text-over-IP as well as interworking between Text-over-IP and existing text telephony on the PSTN and other networks.

Table of Contents

1. Introduction

For many years, text has been in use as a medium for conversational, interactive dialogue between users in a similar way to how voice telephony is used. Such interactive text is different from messaging and semi-interactive solutions like Instant Messaging in that it offers an equivalent conversational experience to users who cannot, or do not wish to, use voice. It therefore meets a different set of requirements from other text-based solutions already available on IP networks.

Traditionally, deaf, hard of hearing and speech-impaired people are amongst the most prolific users of conversational, interactive text but, because of its interactivity, it is becoming popular amongst mainstream users as well.

This document describes how existing IETF protocols can be used to implement a Text-over-IP solution (ToIP). This ToIP framework is specifically designed to be compatible with Voice-over-IP (VoIP) environments, as well as meeting the userÆs requirements, including those of deaf, hard of hearing and speech-impaired users as described in RFC3351 [19].

The Session Initiation Protocol (SIP) is the protocol of choice for control of Multimedia communications and Voice-over-IP (VoIP) in particular. It offers all the necessary control and signaling required for the ToIP framework.

The Real-Time Transport Protocol (RTP) is the protocol of choice for real-time data transmission, and its use for interactive text payloads is described in RFC4103 [5].

This document defines a framework for ToIP to be used either by itself or as part of integrated, multi-media services, including Total Conversation.

2. Scope

This document defines a framework for the implementation of real-time ToIP, either stand-alone or as a part of multimedia services, including Total Conversation. It defines the:

a. Requirements of Real-time, interactive text;
b. Requirements for ToIP interworking;
c. Description of ToIP using SIP and RTP;
d. Description of ToIP interworking with other text services.

3. Terminology

In this document, the key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" are to be interpreted as described in BCP 14, RFC 2119 [2] and indicate requirement levels for compliant implementations.

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.

Cellular - Telephone systems based on radio transmission to become wireless. Also called Wireless or Mobile systems.

Full duplex - media is sent independently in both directions.

Half duplex - media 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 presented media.

Interactive text - a term for real time transmission of text in a character-by-character fashion for use in conversational services, often as a text equivalent to voice based conversational services.

Textphone û also "text telephone". A terminal device that allows end-to-end real-time, interactive text communication using analog transmission. A variety of PSTN textphone protocols exists world-wide. 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 service that enables the flow of text through the service between the users involved in the call.

Text gateway - a function that transcodes between different forms of text transport methods, e.g., between ToIP in IP networks and Baudot or ITU-T V.21 text telephony in the PSTN.

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.

Text telephony û analog textphone service.

Total Conversation - a multimedia service offering real time conversation in video, text and voice according to interoperable standards. All media flow in real time. (See ITU-T F.703 "Multimedia conversational services".)

Transcoding Services - services of a third-party user agent that transcodes one stream into another. Transcoding can be done by human operators, in an automated manner or a combination of both methods. Text Relay Services are examples of a transcoding service between text and audio.

TTY û alternative designation for a text telephone or textphone, often used in USA. Also called TDD, Telecommunication Device for the Deaf.

Video Relay Service - A service that enables communications between deaf and hard of hearing people, 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
CLI	Calling Line Identification
CTM	Cellular Text Telephone Modem
ENUM	E.164 number storage in DNS (see RFC3761)
GSM	Global System of Mobile Communication
ISDN	Integrated Services Digital Network
ITU-T	International Telecommunications Union-Telecommunications Standardisation Sector
NAT	Network Address Translation
PSTN	Public Switched Telephone Network
RTP	Real Time Transport Protocol
SDP	Session Description Protocol
SIP	Session Initiation Protocol
SRTP	Secure Real Time Transport Protocol
TDD	Telecommunication Device for the Deaf
TDMA	Time Division Multiple Access
TTY	Analog textphone (Teletypewriter)
ToIP	Text over Internet Protocol
UTF-8	Universal Transfer Format-8
VCO/HCO	Voice Carry Over/Hearing Carry Over
VoIP	Voice over Internet Protocol

5. Framework Description

This framework defines the requirements of a text-based conversational service that is the text equivalent of voice based telephony. Real-time text conversation can be combined with other conversational services like video or voice.

ToIP also offers an IP equivalent of analog text telephony services as used by deaf, hard of hearing and speech-impaired individuals.

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 takes place. This provides the same mode of interaction as voice telephony does for hearing people.

Store-and-forward systems like email or messaging on mobile networks or non-streaming systems like instant messaging are unable to provide that functionality. In particular, they do not allow for smooth communication through a Text Relay Service.

This framework uses existing standards that are already commonly used for voice based conversational services on IP networks. It uses the Session Initiation Protocol (SIP) to set up, control and tear down the connections between users whilst the media is transported using the Real-Time Transport Protocol (RTP) as described in RFC4103 [5].

This framework is designed to meet the requirements of RFC3351 [19]. As such, it offers a standardized way for offering text-based, conversational services that can be used as an equivalent to voice telephony by deaf, hard of hearing and speech-impaired individuals.

SIP allows participants to negotiate all media including real-time text conversation [4,5]. This is a highly desirable function for all IP telephony users but essential for deaf, hard of hearing, or speech impaired people who have limited or no use of the audio path of the call.

5.1. General requirements for ToIP

In order to make ToIP the text equivalent of voice services, it needs to offer equivalent features in terms of conversationality as voice telephony provides. To achieve that, ToIP needs to:

a. Offer real-time presentation of the conversation;
b. Provide simultaneous transmission in both directions;
c. Support both point-to-point and multipoint communication;

d. Allow other media, like audio and video, to be used in conjunction with ToIP;
e. Ensure that the text service is always available.

Real-time text is a useful subset of Total Conversation defined in ITU-T F.703 [23]. Users could use multiple modes of communication during the conversation, either at the same time or by switching between modes, e.g., between text and audio.

Users may invoke ToIP services for many different reasons:

- Because they are in a noisy environment, e.g., in a machine room of a factory where listening is difficult.
- Because they are busy with another call and want to participate in two calls at the same time.
- For implementing text and/or speech recording services (e.g., text documentation/ audio recording for legal/clarity/flexibility purposes).
- To overcome language barriers through speech translation and/or transcoding services.

- Because of hearing loss, deafness or tinnitus as a result of the aging process or for any other reason, thus creating a need to replace or complement voice with text in conversational sessions.

NOTE: In many of the above examples, text may accompany speech. The text could be displayed side by side, in a manner similar to subtitling in broadcasting environments, or in any other suitable manner. This could occur for users who are hard of hearing and also for mixed media calls with both hearing and deaf people participating in the call.

User Agents providing ToIP functionality need to provide suitable alerting indications, specifically offering visual and/or tactile alerting for deaf and hard of hearing users.

The ability of SIP to set up conversation sessions from any location, as well as its privacy and security provisions, MUST be maintained by ToIP services.

Where ToIP is used in conjunction with other media, exposure of SIP functions through the User Interface needs to be done in an equivalent manner for all supported media. In other words, where certain SIP call control functions are available for the audio media part of the session, these functions MUST also be supported for the text media part of the same session. For example, call transfer must act on all media in the session.

T.140 real-time text conversation [4], in addition to audio and video communications, is a valuable service for many users, including those on non-IP networks. T.140 also provides for real-time editing of the text.

5.1.1 General ToIP Summary

The general requirements for ToIP are:

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. Support of text media transport using T.140 over RTP as described in RFC 4103 [5]

f. Signaling of status information, call progress and the like in a suitable manner, bearing in mind that the user may have a hearing impairment

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

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

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

j. Routing of emergency calls according to national or regional policy with the same level of functionality as a voice call.

5.2. General Requirements for ToIP Interworking

This section describes the general ToIP interworking requirements and gives some background information to many of the issues.

There is a range of existing text services. There is also a range of network technologies that could support text services (see examples below). ToIP needs to provide interoperability with text conversation features in other networks, for instance the PSTN, and with some text messaging services.

Text gateways are used for converting between different media types. They could be used between networks or within networks where different transport technologies are used.

When communicating via a gateway to other networks and protocols, the ToIP service SHOULD support the functionality for alternating or simultaneous use of modalities as offered by the destination network.

Address information, both called and calling, SHOULD be transferred, and possibly converted, when interworking between different networks.

ToIP will often be used to access a relay service [I], allowing 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 minimal. In particular, buffering of whole lines of text will not meet character delay requirements.

If the User Agents of different participants indicate that there is an incompatibility between their capabilities to support certain media types, e.g. one terminal only offering T.140 over IP as described in RFC4103 [5] and the other one only supporting audio, the user might want to invoke a transcoding service.

Examples of possible scenarios for including a relay service in the conversation are: speech-to-text (STT), text-to-speech (TTS), text bridging after conversion from speech, audio bridging after conversion from text, etc.

The general requirements for ToIP Interworking are:

a. Interoperability between T.140 conversations [4] and analog text telephones

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

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

d. 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

e. Media bridging services for T.140 real-time text, as described in RFC4103 [5], audio and video for multipoint communications

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

g. Buffering of text when interworking with media that transport text at different rates.

5.2.1 PSTN Interworking

Analog text telephony is cumbersome because of incompatible national implementations where interworking was never considered.

A large number of these implementations have been documented in ITU-T V.18 [10], which also defines the modem detection sequences for the different text protocols. The modem type identification may in rare cases take considerable time depending on user actions.

To resolve analog textphone incompatibilities, text telephone gateways are needed to transcode incoming analog signals into T.140 and vice versa. The modem capability exchange time can be reduced by the text telephone gateways initially assuming the analog text telephone protocol used in the region where the gateway is located. For example, in the USA, Baudot [III] might be tried as the initial protocol. If negotiation for Baudot fails, the full V.18 modem capability exchange will take place. In the UK, ITU-T V.21 [II] might be the first choice.

5.2.2 Cellular 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 ToIP at the other side.

5.2.2.1 Cellular "No-gain"

The "No-gain" text telephone transporting technology uses specially modified EFR [13] and EVR [14] speech vocoders in mobile terminals used to provide a text telephony call. It provides full duplex operation and supports alternating voice and text ("VCO/HCO"). It is dedicated to CDMA and TDMA mobile technologies and the US Baudot (i.e. 45 bit/s) type of text telephones.

5.2.2.2 Cellular Text Telephone Modem (CTM)

CTM [15] is a technology independent modem technology that provides the transport of text telephone characters at up to 10 characters/sec using modem signals that can be carried by many voice codecs and uses a highly redundant encoding technique to overcome the fading and cell changing losses.

5.2.2.3 Cellular "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.

5.2.3 Cellular 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 network. Gateways provide interoperability with PSTN textphones.

5.2.4 Cellular Wireless ToIP

ToIP could be supported over cellular wireless packet switched services that interface to the Internet. For 3GPP 3G services, the support is described to use ToIP in 3G TS 26.235 [18]. Low data rates and additional delays can affect performance.

5.2.5 Instant Messaging Support

Many people use Instant Messaging to communicate via the Internet using text. Instant Messaging transfers blocks of text rather than streaming as is used by ToIP. As such, it is not a replacement for ToIP and in particular does not meet the needs for real time conversations including those of deaf, hard of hearing and speech-impaired users as defined in RFC 3351 [19]. It is unsuitable for communications through a relay service [I]. The streaming nature of ToIP provides a more direct conversational user experience and, when given the choice, users may prefer ToIP.

Text gateways could be developed to allow interworking between Instant Messaging systems and ToIP solutions.

6. Detailed requirements for ToIP

A ToIP user may wish to call another ToIP user, or join a conference session involving several users or initiate or join a multimedia session, such as a Total Conversation session.

There may be some need for pre-session setup e.g. storing of registration information in the SIP registrar, to provide information about how a user can be contacted. This will allow sessions to be set up rapidly and with proper routing and addressing.

Similarly, there are requirements that need to be satisfied during session set up when other media are preferred by a user. For instance, some users may indicate their preferred modality to be audio while others may indicate text. In this case, transcoding services might be needed 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 some of these requirements in detail.

6.1. Pre-Session Requirements

The need to use text as a medium of communications can be expressed by users during registration time. Two situations need to be considered in the pre-session setup environment:

a. User Preferences: It MUST be possible for a user to indicate a preference for text by registering that preference with a SIP server that is part of the ToIP service.

b. Server to support User Preferences: SIP servers that support ToIP services MUST have the capability to act on calling user preferences for text in order to accept or reject the session-, based on the called userÆs preferences defined as part of the pre-session setup registration. For example, if the user is called by another party, and it is determined that a transcoding server is needed, the session MUST be re-directed or otherwise handled accordingly.

6.2 Basic Point-to-Point Session Requirements

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

6.2.1 Session control

ToIP services MUST use the Session Initiation Protocol (SIP) [3] for setting up, controlling and terminating sessions for real-time text conversation with one or more participants and possibly including other media like video or audio. The session description protocol (SDP) [6] used in SIP to describe the session is used to express the attributes of the session and to negotiate a set of compatible media types.

6.2.2 Text transport

A ToIP service MUST always support at least one Text media type.

ToIP services MUST support the Real-Time Transport Protocol (RTP) [24] according to the specification of RFC4103 [5] for the transport of text between participants.

RFC4103 describes the transmission of T.140 [4] on IP networks.

6.2.3 Session Setup

Users will set up a session by identifying the remote party or the service they want to connect to. However, conversations could be started using a mode other than text. For instance, the conversation might be established using audio and the user could subsequently elect to switch to text, or add text as an additional modality, during the conversation. Systems supporting ToIP 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 audio, and this is not possible either because the system only supports alternate modalities or because of constraints in the network, the system MUST try to establish communication with best effort. If the user has expressed a preference for text, establishment of a connection including text MUST have priority over other outcomes of the session setup.

The following features MAY be implemented to facilitate the session establishment using ToIP:

a. Caller Preferences: SIP headers (e.g., Contact)[24] 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 session should be handled either using text as a preferred medium or not, and whether a designated SIP proxy needs to handle this session or it will be handled in the SIP user agent.

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

6.2.4 Addressing

The SIP [3] addressing schemes MUST be used for all entities in a ToIP session. For example, SIP URLÆs or Tel URLÆs are 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 to invoke the service.

6.2.5 Alerting

User Agents supporting ToIP MUST have an alerting method (e.g., for incoming sessions) that can be used by deaf and hard of hearing people or provide a range of alternative, but equivalent, alerting methods that can be selected by all users, regardless of their abilities.

It should be noted that external 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 by the User AgentÆs User Interface and specific alerting user agents registered to the same registrar as the main user agent.

6.2.6 Session information

If present, identification of the originating party (for example in the form of a URL or a CLI) MUST be clearly presented to the user in a form suitable for the user BEFORE the session invitation is answered. When a session invitation involving ToIP originates from a gateway, this MAY be signaled to the user.

The user MUST be informed of any change in modalities.

6.2.7 Session progress information

During a conversation that includes ToIP, status and session progress information MUST be provided in a textual form so users can perform all session control functions. That information MUST be equivalent to session progress information delivered in any other format, for example audio.

Session progress information SHOULD use simple language so that as many users as possible can understand it. The use of jargon or ambiguous terminology SHOULD be avoided. It is RECOMMENDED that text information be used together with icons to symbolise the session progress information.

There MUST be a clear indication, in a modality useful to the user, whenever a session is connected or disconnected. A user SHOULD never be in doubt about the status of the session, even if the user is unable to make use of the audio or visual indication. For example, tactile indications could be used by deafblind individuals.

In summary, it SHOULD be possible to observe indicators about:
- Incoming session
- Availability of text, voice and video channels
- Session progress
- Incoming text
- Any loss in incoming text

- Typed and transmitted text.

For users who cannot use the audible alerter for incoming sessions, it is RECOMMENDED to include a tactile as well as a visual indicator.

6.2.8 Session Negotiations

The Session Description Protocol (SDP) used in SIP [3] provides the capabilities to indicate text as a medium in the session setup. RFC 4103 [5] uses the RTP payload type "text/t140" for support of ToIP which can be indicated in the SDP as a part of the SIP INVITE, OK and SIP/200/ACK media negotiations. In addition, SIPÆs offer/answer model [20] can also be used in conjunction with other capabilities including the use of a transcoding server for enhanced session negotiations [7,8,9].

6.2.9 Answering

Systems SHOULD provide a best-effort approach to answering invitations for session set-up and users SHOULD be informed when the session is accepted by the other party. On all systems that both inform users of session status and support ToIP, this information MUST be available in textual form and MAY also be provided in other media.

6.2.9.1 Answering Machine

Systems for ToIP 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 4096 characters, although systems MAY support much larger storage. It is RECOMMENDED that systems support storage of at least 20 incoming messages of up to 16000 characters per message.

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

6.2.10 Actions During a Session

Certain actions need to be performed during ToIP conversation:

a. Text transmission from a terminal SHALL be performed character by character as entered, or in small groups of characters, so that no character is delayed from entry to 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. To enable the use of international character sets, the transmission format for text conversation SHALL be UTF-8 [12], in accordance with ITU-T T.140.

d. If text is detected to be missing after transmission, there SHOULD be a "text loss" indication in the text as specified in T.140 Addendum 1 [4].

e. When the display of text conversation is included in the design of the end user equipment, the display of the dialogue SHOULD be made so that it is easy to differentiate the text belonging to each party in the conversation.

6.2.10.1 Text Transport

ToIP uses RTP as the default transport protocol for the transmission of real-time text via the medium "text/t140" as specified in RFC 4103 [5].

The redundancy method of RFC 4103 [5] SHOULD be used to significantly increase the reliability of the text transmission. A redundancy level using 2 generations gives very reliable results and is therefore RECOMMENDED.

Text capability MUST be announced in SDP by a declaration similar to 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
```

By having this single coding and transmission scheme for real time text defined in the SIP session 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, but RFC 4103 [5] transport MUST be used as both the default and the fallback transport.

6.2.10.2 Handling Text and other Media.

A call is one or more related sessions. The following requirements apply to media handling during a call:

a. When used between User Agents designed for ToIP, it SHALL be possible to send and receive text simultaneously.

b. When used between User Agents that support 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.

c. It SHOULD be possible to know during a call that ToIP is available, even if it is not invoked at call setup (e.g. when only voice and/or video is used initially). To disable this, the user MUST disable the use of ToIP. This is possible during registration at the SIP registrar.

6.2.11 Additional session control

Systems that support additional session control features, for example call waiting, forwarding, hold etc on voice sessions, MUST offer this functionality for text sessions.

6.2.12 File storage

Systems that support ToIP MAY save the text conversation to a file. This SHOULD be done using a standard file format. For example: a UTF8 text file in XML format [11] including timestamps, party names (or addresses) and the text conversation.

6.3 Conference Session Requirements

The conference session requirements deal with multipoint conferencing sessions where there will be one or more ToIP capable devices and/or other end user devices where the total number of end user devices will be at least three.

It SHOULD be possible to use the text medium in conference sessions in a similar way to how audio is handled and video is displayed. Text in conferences can be used both for letting individual participants use the text medium (for example, for sidebar discussions in text while listening to the main conference audio), as well as for central support of the conference with real time text interpretation of speech.

6.4 Real-time Editing and User Alerting

ToIP SHOULD handle characters such as new line, erasure and alerting during a session as specified in ITU-T T.140.

6.5 Emergency services

It MUST be possible to place an emergency call using ToIP and it MUST be possible to use a relay service in such call. The emergency service provided to users utilising the text medium MUST be equivalent to the emergency service provided to users utilising speech or other media.

6.6 User Mobility

ToIP User Agents SHOULD use the same mechanisms as other SIP User Agents to resolve mobility issues. It is RECOMMENDED that users use a SIP-address, resolved by a SIP registrar, to enable basic user mobility. Further mechanisms are defined for all session types for 3G IP multimedia systems.

6.7 Firewalls and NATs

ToIP uses the same signaling and transport protocols as VoIP hence, the same firewall and NAT solutions and network functionality that apply to VoIP MUST also apply to ToIP.

7. Interworking Requirements for ToIP

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 the interoperability requirements, especially for PSTN text telephony, to ensure full backward interoperability with ToIP.

7.1 ToIP 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.

Simultaneous or alternating use of voice and text is used by a large number of users who can send voice but must receive text (due to a hearing impairment), or who can hear but must send text (due to a speech impairment).

Session setup through gateways to other networks MAY require the use of specially formatted addresses or other mechanisms for invoking those gateways.

Different data rates of different protocols MAY require text buffering.

Transcoding of text to and from other coding formats MAY need to take place in gateways between ToIP and other forms of text conversation, for example to connect to a PSTN text telephone.

7.2 ToIP and PSTN/ISDN Text-Telephony Interworking

On PSTN networks, transmission of interactive text takes place using a variety of codings and modulations, including ITU-T V.21 [II], Baudot [III], DTMF, V.23 [IV] 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.

V.18 MUST be supported on the PSTN side of a PSTN-ToIP gateway.

PSTN-ToIP gateways MUST allow alternating use of text and voice if the PSTN textphone involved at the PSTN side of the session supports this. (This mode is often called VCO/HCO).

Calling party identification information, such as CLI, MUST be passed by gateways and converted to an approapriate form if required.

7.3 ToIP and Cellular Wireless ToIP

ToIP MAY be supported over the cellular wireless packet switched service. It interfaces to the Internet.

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

7.4 Instant Messaging Support

Text gateways MAY be developed to allow interworking between Instant Messaging systems and ToIP solutions. Because Instant Messaging is based on blocks of text, rather than on a continuous stream of characters, gateways MUST transcode between the two formats. Text gateways for interworking between Instant Messaging and ToIP MUST concatenate individual characters originating at the ToIP side into blocks of text and:

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

b. When a new line indicator is received from the ToIP side, the buffered characters up to that point, including the carriage return and/or line feed characters, SHOULD be transmitted to the Instant Messaging side.

c. When the ToIP side has been idle for at least 5 seconds, all buffered text up to that point SHOULD 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.

Many Instant Messaging protocols signal that a user is typing to the other party in the conversation. Text gateways between such Instant Messaging protocols and ToIP MUST provide this signaling to the Instant Messaging side when characters start being received, or at the beginning of the conversation.

At the ToIP side, an indicator of writing the Instant Message MUST be present where the Instant Messaging protocol provides one. For example, the real-time text user MAY see ". . . waiting for replying IM. . . " and when 5 seconds have passed another . (dot) can be shown.

Those solutions will reduce the difficulties between streaming and blocked text services.

Even though the text gateway can connect Instant Messaging and ToIP, the best solution is to take advantage of the fact that the user interfaces and the user communities for instant messaging and ToIP telephony are very similar. After all, the character input, the character display, Internet connectivity and SIP stack are the same for Instant Messaging (SIMPLE) and ToIP.

Devices that implement Instant Messaging SHOULD implement ToIP as described in this document so that a more complete text communication service can be provided.

7.5 Common Text Gateway Functions

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

7.5.1 Protocol support

Text gateways MUST use the ITU-T V.18 [10] standard at the PSTN side. A text gateway MUST act as a SIP User Agent on the IP side and support RFC4103 text transport.

7.5.2 Relay buffer storage

When text gateway functions are invoked, there will be a need for intermediate storage of characters before transmission to a device receiving text slower than the transmitting 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 for up to 4 minutes (i.e. less than 3k characters).

Interoperation of half-duplex and full-duplex protocols MAY require text buffering. Some intelligence will be needed to determine when to change direction when operating in half-duplex mode. 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).

7.5.3 Emergency calls through gateways

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 e.g. "911" in USA and "112" in Europe. Routing text calls to emergency services MAY require the use of a transcoding service.

7.5.4 Text Gateway Invocation

ToIP interworking requires a method to invoke a text gateway. As described previously in this draft, these text gateways MUST act as User Agents at the IP side. The capabilities of the text gateway during the call will be determined by the call capabilities of the terminal that is using the gateway. For example, a PSTN textphone is generally only able to receive voice and streaming text, so the text gateway will only allow ToIP and audio.

Examples of possible scenarios for invocation of the text gateway are:

a. PSTN textphone users dial a prefix number before dialing out.
b. Separate text subscriptions, linked to the phone number or terminal identifier/ IP address.
c. Text capability indicators.
d. Text preference indicator.
e. Listen for V.18 modem modulation text activity in all PSTN calls and routing of the call to an appropriate gateway.
f. Call transfer request by the called user.
g. Placing a call via the web, and using one of the methods described here
h. Text gateways with its own telephone number and/or SIP address. (This requires user interaction with the text gateway to place a call).
i. ENUM address analysis and number plan
j. Number or address analysis leads to a gateway for all PSTN calls.

7.6 Home Gateways or Analog Terminal Adapters

Analog terminal adapters (ATAs) using SIP based IP communication and RJ-11 connectors for connecting traditional PSTN devices SHOULD enable connection of legacy PSTN text telephones [16].

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 4103 [5], in a similar way as it provides interoperability for voice sessions.

If a session is set up and text/t140 capability is not declared by the destination 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 analog terminal adaptors do use RFC 4103 [5] on board and thus act as a text gateway. Sending textphone signals over the voice channel is undesirable due to possible filtering and compression and packet loss between the end-points. This can result in character loss in the textphone conversation or even not allowing the textphones to connect to each other.

7.7 Multi-functional Combination gateways

In practice many interworking gateways will be implemented as gateways that combine different functions. As such, a text gateway could be built to have modems to interwork with the PSTN and support both Instant Messaging as well as 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.

7.8 Transcoding

Gateways between the ToIP network and other networks MAY need to transcode text streams. ToIP makes use of the 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 character sets and T.140 in gateways, special consideration MUST be given to the national variants of the 7 bit codes, with national characters mapping into different codes in the ISO 10646 code space. The national variant to be used could be selectable by the user on a per call basis, or be configured as a national default for the gateway.

The indicator of missing text in T.140, specified in T.140 amendment 1, cannot be represented in the 7 bit character codes. Therefore the indicator of missing text SHOULD be transcoded to the ' (apostrophe) character in legacy text telephone systems,

where this character exists. For legacy systems where the character ' does not exist, the . (full stop) character SHOULD be used instead.

7.9 Relay Services

The relay service acts as an intermediary between two or more callers using different media or different media encoding schemes.

7.9.1 Basic function of the relay service

The basic text 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 want to remind readers that other relay services exist, like video relay services transcoding speech to sign language and vice versa where the signing is communicated using video.

7.9.2 Invocation of relay services

It is RECOMMENDED that ToIP implementations make the invocation and use of relay services as easy as possible. It MAY happen automatically when the session is being set up based on any valid indication or negotiation of supported or preferred media types. 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. ToIP implementations SHOULD support this transcoding framework.

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

When setting up a session, 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, the sign language being used (in the video stream), etc.

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

A number of requirements, motivations and implementation guidelines for relay service invocation can be found in RFC 3351 [19].

8. Security Considerations

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 [17].

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.

9. Authors Addresses

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

Willem P. Dijkstra
TNO Informatie- en Communicatietechnologie
Postbus 15000
9700 CD Groningen
The Netherlands
Tel: +31 50 585 77 24
Fax: +31 50 585 77 57
Email: willem.dijkstra@tno.nl

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

Guido Gybels
Department of New Technologies
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
SAIC
3465-B Box Hill Corporate Center Drive
Abingdon, MD 21009
Tel: 443 402 9041
Email: Radhika.R.Roy@saic.com

Henry Sinnreich
pulver.com
115 Broadhollow Rd
Suite 225
Melville, NY 11747
USA
Tel: +1.631.961.8950

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

Arnoud A. T. van Wijk
Viataal
Centre for R & D on sensory and communication disabilities.
Theerestraat 42
5271 GD Sint-Michielsgestel
The Netherlands.

Email: a.vwijk@viataal.nl

10. References

10.1 Normative references

1. S. Bradner, "Intellectual Property Rights in IETF Technology
", BCP 79, RFC 3979, IETF, March 2005.

2. S. Bradner, "Key words for use in RFCs to Indicate Requirement
Levels", BCP 14, RFC 2119, IETF, 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 4103,
IETF, June 2005.

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

7. G.Camarillo, "Framework for Transcoding with the Session
Initiation Protocol" IETF June 2005 - Work in progress.

8. G. Camarillo, H. Schulzrinne, E. Burger, and A. van Wijk,
"Transcoding Services Invocation in the Session Initiation
Protocol (SIP) Using Third Party Call Control (3pcc)" RFC 4117,
IETF, June 2005.

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, IETF, January 1998.

13. 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)"

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

15. 3GPP TS26.226 "Cellular Text Telephone Modem Description" (CTM).

16. H. Sinnreich, S. Lass, and C. Stredicke, "SIP Telephony Device Requirements and Configuration," IETF, June 2005 - Work in Progress.

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

18. "IP Multimedia default codecs". 3GPP TS 26.235

19. Charlton, Gasson, Gybels, Spanner, van Wijk, "User Requirements for the Session Initiation Protocol (SIP) in Support

of Deaf, Hard of Hearing and Speech-impaired Individuals", RFC 3351, IETF, August 2002.

20. J. Rosenberg, H. Schulzrinne, "An Offer/Answer Model with the Session Description Protocol (SDP)", RFC 3624, IETF, June 2002.

21. ITU-T Recommendation F.700,"Framework Recommendation for Multimedia Services", November 2000.

22. H. Schulzrinne, S.Casner, R. Frederick, V. Jacobsone, "RTP: A Transport Protocol for Real-Time Applications", RFC 3550, IETF, July 2003.

23. ITU-T Recommendation F.703,"Multimedia Conversational Services", November 2000.

24. J. Rosenberg, H. Schulzrinne, P. Kyzivat, "Indicating User Agent Capabilities in the Session Initiation Protocol (SIP)", RFC 3840, IETF, August 2004

10.2 Informative references

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. 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.)

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

Intellectual Property Statement

Disclaimer of Validity

PARTICULAR PURPOSE.

Copyright Statement

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.