

Internet Engineering Task Force

SIPPING WG

Internet Draft

Document: <[draft-ietf-sipping-toip-01.txt](#)>

A. van Wijk (editor)

July 18 2005

Viataal

Expires: January 17 2006

Informational

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

#### Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with [Section 6 of BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/1id-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

This Internet-Draft will expire on January 17, 2006.

#### Copyright Notice

Copyright (C) The Internet Society (2005).

#### 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 Transport Protocol for Real-Time Applications. It discusses requirements for real-time Text-over-IP telephony as well as interworking between Text-over-IP telephony and existing text telephony on the PSTN and other networks.

Table of Contents

1. Introduction	3
2. Scope	3
3. Terminology	3
4. Definitions	4
5. Framework Description	5
5.1. Background	5
5.2. Requirements for ToIP	6
5.3. Use of SIP and RTP	6
5.4. Requirements for ToIP Interworking	9
6. Detailed requirements for Text-over-IP	9
6.1. Pre-Call Requirements	10
6.2 Basic Point-to-Point Call Requirements	10
6.2.1 Session Setup	10
6.2.2 Addressing	11
6.2.3 Alerting and session progress presentation	11
6.2.4 Call Negotiations	12
6.2.5 Answering	12
6.2.6 Actions During Calls	13
6.2.7 Additional session control	14
6.2.8 File storage	15
6.3 Conference Call Requirements for ToIP User Agents	15
6.4 Transport via RTP	15
6.5 Character Set	16
6.6 Transcoding	16
6.7 Relay Services	16
6.8 Emergency services	17
6.9 User Mobility	17
6.10 Confidentiality and Security	17
7. Interworking Requirements for ToIP	17
7.1 ToIP Interworking Gateway Services	17
7.2 ToIP and PSTN/ISDN Text-Telephony	18
7.3 ToIP and Cellular Wireless circuit switched Text-Telephony	18
7.3.1 "No-gain"	19
7.3.2 Cellular Text Telephone Modem (CTM)	19
7.3.3 "Baudot mode"	19
7.3.4 Data channel mode	19
7.3.5 Common Text Gateway Functions	19
7.4 ToIP and Cellular Wireless ToIP	20
7.5 Instant Messaging Support	20
7.6 IP Telephony with Traditional RJ-11 Interfaces	21
7.7 Multi-functional gateways	22
7.8 ToIP interoperability with PSTN text telephones.	22
7.9 Gateway Discovery	22

8. Afterword	23
9. Security Considerations	23
10. Authors Addresses	24
11. References	25
11.1 Normative	25
11.2 Informative	27

## **1. Introduction**

For many years, text has been in use as a medium for conversational, interactive dialogue between users in a similar way as 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 that cannot, or do not wish to, use voice. It therefore meets a different set of requirements than other text-based solutions already available on IP networks.

Traditionally, deaf, hard of hearing and speech-impaired people are amongst the most proliferate users of conversational, interactive text, but because of its interactivity, it is becoming popular amongst mainstream user groups 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 environments, as well as meeting the user's requirements, including those of deaf, hard of hearing and speech-impaired users as described in [RFC3351](#) [21].

The Session Initiation Protocol (SIP) is the protocol of choice for control of Multimedia IP telephony and Voice-over-IP (VoIP) communications. 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 services, including Total Conversation.

## **2. Scope**

The primary scope of this document is to define a framework for the implementation of ToIP, either stand-alone or as a part of wider services, including Total Conversation. In general, the scope is:

- a. Description of ToIP using SIP and RTP;

- b. Requirements of Real-time, interactive text;
- c. Requirements for ToIP interworking.

The subsequent sections describe those requirements in detail.

### **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.

[A. van Wijk](#)  
[draft-ietf-sipping-ToIP-01.txt](#)

[Page 3 of 28]  
July 18 2005

### **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 - 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.

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

Textphone - also text telephone. A terminal device that allows end-to-end real-time, interactive 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 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 ð 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 that transcodes one stream into another. Transcoding can be done by human operators, in automated manner or a combination of both methods. Text Relay Services are examples of a transcoding service between text and audio.

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.

[A. van Wijk](#)  
[draft-ietf-sipping-ToIP-01.txt](#)

[Page 4 of 28]  
July 18 2005

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
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. Framework Description](#)

### [5.1. Background](#)

The main purpose of this document is to provide a framework description for the implementation of real-time, interactive text based conversational services over IP networks, known as Text-

over-IP (ToIP).

This framework uses existing standards that are already commonly used for voice based conversational services on IP networks. In particular, the ToIP framework uses the Session Initiation Protocol (SIP) [3] to set up, control and tear down the connections between users.

Media is transported using the Real-Time Transport Protocol (RTP) in the manner described in [RFC4103](#).

This framework allows for implementation of services that meet the requirement of providing a text-based conversational service, equivalent to voice based telephony. In particular, ToIP offers an IP equivalent of text telephony services as used by deaf, hard of hearing and speech-impaired individuals.

In addition, real-time text conversations can be combined with other conversational services using different media like video or voice.

By using SIP, ToIP 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.

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, thus providing 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.

## **5.2. Requirements for ToIP**

In order to make ToIP the equivalent of what voice is to hearing people, it needs to offer equivalent features in terms of conversationality as voice telephony provides to hearing people. To achieve that, ToIP MUST:

- a. Offer real-time presentation of the conversation;
- b. Provide simultaneous transmission in both directions;
- c. Provide interoperability with text conversation features in other networks, for instance the PSTN, accepting functional limitations that will occur during interoperation.
- d. Not prevent other media, like audio and video, to be used in conjunction with ToIP.

Users might want to use multiple modes of communication during the conversation, either at the same time or by switching between modes, e.g., between text and audio for example. Native ToIP services **MUST** ensure that the text interface is always available.

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

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 as small as possible. In particular, buffering of whole lines of text **MUST NOT** be used.

### **5.3. Use of SIP and RTP**

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.

Thus, participants are allowed to negotiate on a set of compatible media types with session descriptions used in SIP invitations. A ToIP service **MUST** always support at least one Text media type.

ToIP services **MUST** use the Real-Time Transport Protocol (RTP) according to the specification of [RFC4103](#) for the transport of text between participants, which implements T.140 on IP networks.

The standardized T.140 real-time text conversation [4], in addition to audio and video communications, will be a valuable service to many, including on non-IP networks. Real-time text can be expressed as a part of the session description in SIP and is a useful subset of Total Conversation.

The ToIP specification describes a 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.

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](#) and the other one only supporting audio, the user might want to invoke a transcoding services.

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 session description protocol (SDP) [6] used in SIP to describe the session is used to express 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 other media like video and audio, as for example in Total Conversation services.

User Agents providing ToIP functionality SHOULD provide suitable alerting, specifically offering visual and/or tactile alerting so that deaf and hard of hearing users can use them.

The SIP abilities to set up text conversation sessions from any location, as well as privacy and security provisions SHOULD be implemented in ToIP services.

Where ToIP is used in conjunction with other media, exposure of SIP functions through the User Interface MUST be available in equivalent fashion 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.

Any ToIP implementation MUST also allow invocation and use of relevant transcoding services where these are available. This can be achieved through application of SIP techniques for 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, ToIP services SHOULD support interworking with text telephony [10].

The general framework for ToIP 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 as described in [RFC4103](#), audio, and video for multipoint communications
- i. Transparent session setup, modification, and teardown between text conversation capable and voice/video capable devices
- j. Support of text media transport using T.140 over RTP as laid out in [RFC 4103](#) [4]
- k. Signaling of status information, call progress and the like in a suitable manner, bearing in mind the user may have a hearing impairment
- 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.

#### **[5.4. Requirements for ToIP 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, which also defines modem detection sequences for the different text terminals. The full modem capability exchange between two wildly different terminals can take more than one minute to complete if both terminals have a common text modulation.

To resolve international analog textphone incompatibilities, text telephone gateways MUST transcode incoming analog signals into T.140 and vice versa. The modem capability exchange time is then also reduced, since V.18 allows the sequence of protocol discovery to be customized. Hence, the text telephone gateways will assume the analog text telephone protocol used in the region the gateway is located. For example, in the USA, Baudot might be tried as the initial protocol. If negotiation for Baudot fails, the full modem capability exchange will then take place. In contrast, in the UK, ITU-T V.21 might be the first choice.

## **6. Detailed requirements for Text-over-IP**

ToIP services MUST use SIP for call control and signaling.

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

There may 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 routing and addressing.

Similarly, there are requirements that need to be satisfied during call set up when other media are preferred by a user. For instance, some users may prefer to use audio while others want to use text as their preferred modality. 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-Call Requirements**

The need to use ToIP as a medium of communications can be expressed by users 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 with a SIP server that is part of the ToIP service.
- b. Server to support User Preferences: SIP servers that are part of ToIP services MUST have the capability to act on users preferences for ToIP to accept or reject the call, based on the user preferences defined during the pre-call setup registration time. For example, if the user is called by another party, and it is determined that a transcoding server is needed, the call MUST be re-directed or otherwise handled accordingly.

## **6.2 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.

### **6.2.1 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 ToIP. 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 resource management on the network, the system MUST try to establish a text-only communication. 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.

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

- 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).

### **6.2.2 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.

### **6.2.3 Alerting and session progress presentation**

User Agents supporting ToIP 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 by the User Agent's User Interface 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 ToIP originates from a gateway, this MAY be signaled to the user.

During a conversation that includes ToIP, 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 SHOULD use simple language as much as possible so that as many users as possible can understand it. 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 or 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.

In summary, 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.

#### **6.2.4 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 4103](#) [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, SIP's 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].

#### **6.2.5 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 ToIP, this information MUST be available in text, and MAY be provided in other visual media.

#### **6.2.5.1 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

**A. van Wijk**  
[draft-ietf-sipping-ToIP-01.txt](#)

**[Page 12 of 28]**  
July 18 2005

MAY support much larger storage. It is RECOMMENDED that systems support storage of at least 20 incoming messages of up to 16000 characters.

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 MUST be allowed to intervene during any stage of the answering machine and take control of the session.

#### **6.2.6 Actions During Calls**

Certain actions need to be performed for the ToIP conversation during the call and these actions are described 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. To enable the use of international character sets the transmission format for text conversation SHALL be UTF-8, in accordance with ITU-T T.140.

g. If text is detected to be missing after transmission, there SHALL be an indication in the text marking the loss. For 7 bit terminals this loss MAY be marked as an apostrophe: ¢.

g. When used from a terminal designed for PSTN text telephony, or in interworking with such a terminal, ToIP shall enable

[A. van Wijk](#)  
[draft-ietf-sipping-ToIP-01.txt](#)

[Page 13 of 28]  
July 18 2005

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 the USA and the UK).

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.

#### **6.2.6.1 Text and other Media Handling Between ToIP User Agents**

The following requirements are valid for media handling during calls:

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 the call that ToIP is available, even if it is not invoked at call setup (only voice and/or video is used for example). To disable this, the user must disable the use of ToIP. This is possible during registration at the REGISTRAR.

#### **6.2.6.2 Call Action with Native ToIP User Agents**

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

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

c. It MUST 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 invoke multi-party calls.

f. It MUST be possible to transfer the call.

#### **6.2.7 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 text calls.

#### **6.2.8 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: UTF8 text file in XML format including record timestamp, party and the text conversation.

### **6.3 Conference Call Requirements for ToIP User Agents**

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 the audio is handled and the 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 Transport via RTP**

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

The redundancy method of [RFC 4103](#) [5] SHOULD be used for making text transmission reliable.



Text capability MUST 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 4103](#) [5] transport MUST be used as both the default as well as the fallback transport.

## [6.5](#) Character Set

- a. ToIP services MUST use UTF-8 encoding as specified in ITU-T T.140 [12].
- b. ToIP SHOULD handle characters with editing effect such as new line, erasure and alerting during session as specified in ITU-T T.140.

## [6.6](#) Transcoding

Transcoding of text may need to take place in gateways between ToIP and other forms of text conversation. For example to connect to a PSTN text telephone.

## [6.7](#) Relay Services

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

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 there exist other relay services like, for example, speech to sign language and vice versa using video.

It is RECOMMENDED that ToIP implementations make the invocation and use of relay services as easy as possible. It MAY happen automatically when the call 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 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, the sign language being used (in the video stream), etc.

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.

## **[6.8](#) Emergency services**

Access to emergency services using ToIP SHOULD provide an equivalent service to the one offered by other supported media, like audio.

## **[6.9](#) User Mobility**

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

## **[6.10](#) 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.

## **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 requirements on this interoperability, especially for the 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 or who can hear but must send text due to a speech disability.

### **7.2 ToIP 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 ToIP 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-ToIP gateways MUST allow alternating use of text and voice.

PSTN and ISDN to ToIP 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.

### **7.3 ToIP 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 ToIP at the other side.

#### **7.3.1 "No-gain"**

The "No-gain" text telephone transporting technology uses specially modified EFR [15] and EVR [16] speech vocoders in both mobile terminals used to 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.

#### **7.3.2 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.

### **7.3.3 "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.

### **7.3.4 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 network. Gateways or the interworking function provides interoperability with PSTN textphones.

### **7.3.5 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. Routing text calls to emergency services MAY require the use of a transcoding service.

A text gateway MUST act as a SIP User Agent on the IP side.

#### **7.4 ToIP and Cellular Wireless ToIP**

ToIP 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.

#### **7.5 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 of deaf, hard of hearing and speech-impaired users as defined in [RFC 3351](#) [21]. It is unsuitable for communications through a relay service [I]. The streaming character of ToIP provides a better user experience and, when given the choice, users often prefer ToIP.

However, since some users might only have Instant Messaging available, 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, such gateways need to transform between these 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-alphanumeric character is received from the ToIP side.
- b. When a new line 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 per 5 seconds that pass a . (dot) can be shown.

Those solutions will reduce the difficulties between a streaming versus blocked text.

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 extremely 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.

## **7.6 IP Telephony with Traditional RJ-11 Interfaces**

Analogue adapters using SIP based IP communication and RJ-11 connectors for connecting traditional PSTN devices (ATA box) 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 4103](#) [5], 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 [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 dropping characters in the textphone conversation or even not allowing the textphones to connect with

each other.

### **[7.7](#) Multi-functional gateways**

In practice many interworking gateways 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 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](#) ToIP interoperability with PSTN text telephones.**

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 10 646 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 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 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](#) Gateway Discovery**

ToIP 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 only able to receive voice and streaming text, so the text gateway will only allow ToIP and audio.

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

**A. van Wijk**

[draft-ietf-sipping-ToIP-01.txt](#)

**[Page 22 of 28]**

July 18 2005

- PSTN textphone users dial a prefix number before dialing out.
- Separate text subscriptions, linked to the phone number or terminal identifier/ IP address.
- Text capability indicators.
- Text preference indicator.
- Listen for V.18 modem modulation text activity in all calls.
- Call transfer request by the called user.
- Placing a call via the web, and using 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.

## **8. Afterword**

The authors want to make it clear that ToIP is a way of allowing real-time, interactive text conversation between all users and is thus not only for the hearing and speech impaired users.

The users may invoke the ToIP services for many different reasons. For example:

- 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.
- Hearing loss, tinnitus or deafness due to the aging process or any other reason.

NOTE: In many of the above examples, text may accompany speech and could be displayed in a manner similar to subtitling in broadcasting environments or any other suitable manner. This could occur for individuals who are hard of hearing and also for

mixed calls with a hearing and deaf person listening to the call.

## **9. Security Considerations**

There are no additional security requirements other than described earlier.

**A. van Wijk**  
[draft-ietf-sipping-ToIP-01.txt](#)

**[Page 23 of 28]**  
July 18 2005

## **10. 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  
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  
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](mailto:gunnar.hellstrom@omnitor.se)

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

**A. van Wijk**  
[draft-ietf-sipping-ToIP-01.txt](#)

**[Page 24 of 28]**  
July 18 2005

Gregg C Vanderheiden  
University of Wisconsin-Madison  
Trace R & D Center  
1550 Engineering Dr (Rm 2107)  
Madison, Wi 53706  
USA  
[gv@trace.wisc.edu](mailto: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](mailto:a.vwijk@viataal.nl)

## **11. References**

### **11.1 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 4103](#), 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](#), June 2005.

**A. van Wijk**  
[draft-ietf-sipping-ToIP-01.txt](#)

**[Page 25 of 28]**  
July 18 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](#), 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](#), IETF, March 2004.

20. IP Multimedia default codecs. 3GPP TS 26.235

21. 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.

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

**A. van Wijk**  
[draft-ietf-sipping-ToIP-01.txt](#)

**[Page 26 of 28]**  
July 18 2005

## **11.2 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.typtalk.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.

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in [BCP 78](#) and [BCP 79](#).

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at [ietf-ipr@ietf.org](mailto:ietf-ipr@ietf.org).

#### Disclaimer of Validity

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR

[A. van Wijk](#)  
[draft-ietf-sipping-ToIP-01.txt](#)

**[Page 27 of 28]**  
July 18 2005

ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

#### Copyright Statement

Copyright (C) The Internet Society (2005). This document is subject to the rights, licenses and restrictions contained in [BCP 78](#), and except as set forth therein, the authors retain all their rights.

#### Acknowledgment

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

