

Network Working Group  
Request for Comments: 4239  
Category: Standards Track

S. McRae  
IBM  
G. Parsons  
Nortel Networks  
November 2005

## Internet Voice Messaging (IVM)

### Status of This Memo

This document specifies an Internet standards track protocol for the Internet community, and requests discussion and suggestions for improvements. Please refer to the current edition of the "Internet Official Protocol Standards" (STD 1) for the standardization state and status of this protocol. Distribution of this memo is unlimited.

### Copyright Notice

Copyright (C) The Internet Society (2005).

### Abstract

This document describes the carriage of voicemail messages over Internet mail as part of a unified messaging infrastructure.

The Internet Voice Messaging (IVM) concept described in this document is not a successor format to VPIM v2 (Voice Profile for Internet Mail Version 2), but rather an alternative specification for a different application.

### 1. Introduction

For some forms of communication, people prefer to communicate using their voices rather than typing. By permitting voicemail to be implemented in an interoperable way on top of Internet Mail, voice messaging and electronic mail no longer need to remain in separate, isolated worlds, and users will be able to choose the most appropriate form of communication. This will also enable new types of devices, without keyboards, to be used to participate in electronic messaging when mobile, in a hostile environment, or in spite of disabilities.

There exist unified messaging systems that will transmit voicemail messages over the Internet using SMTP/MIME, but these systems suffer from a lack of interoperability because various aspects of such a message have not hitherto been standardized. In addition, voicemail systems can now conform to the Voice Profile for Internet Messaging

(VPIM v2 as defined in RFC 2421 [VPIMV2] and revised in RFC 3801, Draft Standard [VPIMV2R2]) when forwarding messages to remote voicemail systems. VPIM v2 was designed to allow two voicemail systems to exchange messages, not to allow a voicemail system to interoperate with a desktop e-mail client. It is often not reasonable to expect a VPIM v2 message to be usable by an e-mail recipient. The result is messages that cannot be processed by the recipient (e.g., because of the encoding used), or look ugly to the user.

This document therefore proposes a standard mechanism for representing a voicemail message within SMTP/MIME, and a standard encoding for the audio content, which unified messaging systems and mail clients MUST implement to ensure interoperability. By using a standard SMTP/MIME representation and a widely implemented audio encoding, this will also permit most users of e-mail clients not specifically implementing the standard to still access the voicemail messages. In addition, this document describes features an e-mail client SHOULD implement to allow recipients to display voicemail messages in a more friendly, context-sensitive way to the user, and intelligently provide some of the additional functionality typically found in voicemail systems (such as responding with a voice message instead of e-mail). Finally, how a client MAY provide a level of interoperability with VPIM v2 is explained.

It is desirable that unified messaging mail clients also be able to fully interoperate with voicemail servers. This is possible today, providing the client implements VPIM v2 [VPIMV2R2], in addition to this specification, and uses it to construct messages to be sent to a voicemail server.

The definition in this document is based on the IVM Requirements document [GOALS]. It references separate work on critical content [CRITICAL] and message context [HINT]. Addressing and directory issues are discussed in related documents [ADDRESS], [VPIMENUM], [SCHEMA].

Further information on VPIM and related activities can be found at <http://www.vpim.org> or <http://www.ema.org/vpim>.

## 2. Conventions Used in This Document

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

### 3. Message Format

Voice messages may be created explicitly by a user (e.g., recording a voicemail message in their mail client) or implicitly by a unified messaging system (when it records a telephone message).

All messages **MUST** conform with the Internet Mail format, as updated by the DRUMS working group [DRUMSIMF], and the MIME format [MIME1].

When creating a voice message from a client supporting IVM, the message header **MUST** indicate a message context of "voice-message" (see [HINT]). However, to support interoperability with clients not explicitly supporting IVM, a recipient **MUST NOT** require its presence in order to correctly process voice messages.

The receiving agent **MUST** be able to support multipart messages [MIME5]. If the receiving user agent identifies the message as a voice message (from the message context), it **SHOULD** present it to the user as a voice message rather than as an electronic mail message with a voice attachment (see [BEHAVIOUR]).

Any content type is permitted in a message, but the top level content type on a new, forwarded or reply voice message **SHOULD** be multipart/mixed. If the recipient is known to be VPIM v2 compliant, then multipart/voice-message **MAY** be used instead (in which case, all the provisions of [VPIMV2R2] **MUST** be implemented in constructing the message).

If the message was created as a voice message, and so is not useful if the audio content is omitted, then the appropriate audio body part **MUST** be indicated as critical content, via a Criticality parameter of CRITICAL on the Content-Disposition (see [CRITICAL]). Additional important body parts (such as the original audio message if a voicemail is being forwarded) **MAY** also be indicated via a Criticality of CRITICAL. Contents that are not essential to communicating the meaning of the message (e.g., an associated vCard for the originator) **MAY** be indicated via a Criticality of IGNORE.

When forwarding IVM messages, clients **MUST** preserve the content type of all audio body parts in order to ensure that the new recipient is able to play the forwarded messages.

The top level content type, on origination of a delivery notification message, **MUST** be a multipart/report. This will allow automatic processing of the delivery notification, for example, so that text-to-speech processing can render a non-delivery notification in the appropriate language for the recipient.

#### 4. Transport

The message MUST be transmitted in accordance with the Simple Mail Transport Protocol, as updated by the DRUMS working group [DRUMSMTP].

Delivery Status Notifications MAY be requested [DSN] if delivery of the message is important to the originator and a mechanism exists to return status indications to them (which may not be possible for voicemail originators).

#### 5. Addressing

Any valid Internet Mail address may be used for a voice message.

It is desirable to be able to use an onramp/offramp for delivery of a voicemail message to a user, which will result in specific addressing requirements, based on service selectors defined in [SELECTOR]. Further discussion of addressing requirements for voice messages can be found in the VPIM Addressing document [ADDRESS].

It is desirable to permit the use of a directory service to map between the E.164 phone number of the recipient and an SMTP mailbox address. A discussion on how this may be achieved using the ENUM infrastructure is in [VPIMENUM]. A definition of the VPIM LDAP schema that a system would use is found in [SCHEMA].

If a message is created and stored as a result of call answering, the caller's name and number MAY be stored in the message headers in its original format per [CLID].

#### 6. Notifications

Delivery Status Notifications MUST be supported. All non-delivery of messages MUST result in an NDN, if requested [DSN, DSN2, DSN3, DSN4]. If the receiving system supports content criticality and is unable to process all of the critical media types within a voice message (indicated by the content criticality), then it MUST non-deliver the entire message per [CRITICAL].

Message Disposition Notifications SHOULD be supported (but according to MDN rules, the user MUST be given the option of deciding whether MDNs are returned) per [MDN].

If the recipient is unable to display all of the indicated critical content components indicated, then it SHOULD give the user the option of returning an appropriate MDN (see [CRITICAL]).

## 7. Voice Contents

Voice messages may be contained at any location within a message and **MUST** always be contained in an audio/basic content-type, unless the originator is aware that the recipient can handle other content. Specifically, Audio/32kadpcm **MAY** be used when the recipient is known to support VPIM v2 [VPIMV2R2].

The VOICE parameter on Content-Disposition from VPIM v2 [VPIMV2R2] **SHOULD** be used to identify any spoken names or spoken subjects (as distinct from voice message contents). As well, the Content-Duration header [DUR] **SHOULD** be used to indicate the audio length.

The originator's spoken name **MAY** be included with messages as separate audio contents, if known, and **SHOULD** be indicated by the Content-Disposition VOICE parameter as defined in VPIM v2 [VPIMV2R2]. If there is a single recipient for the message, the spoken name **MAY** also be included (per VPIM v2). A spoken subject **MAY** also be provided (per VPIM v2).

A sending implementation **MAY** determine the recipient capabilities before sending a message and choose a codec accordingly (e.g., using some form of content negotiation). In the absence of such recipient knowledge, sending implementations **MUST** use raw G.711 mu-law, which is indicated with a MIME content type of "audio/basic" (and **SHOULD** use a filename parameter that ends ".au") [G711], [MIME2]. A sending implementation **MAY** support interoperability with VPIM v2 [VPIMV2R2], in which case, it **MUST** be able to record G.726 (indicated as audio/32kadpcm) [G726], [ADPCM].

Recipients **MUST** be able to play a raw G.711 mu-law message, and **MAY** be able to play G.726 (indicated as audio/32kadpcm) to provide interoperability with VPIM v2. A receiving implementation **MAY** also be able to play messages encoded with other codecs (either natively or via transcoding).

These requirements may be summarized as follows:

Codec -----	No VPIM v2 Support		With VPIM v2 Support	
	Record -----	Playback -----	Record -----	Playback -----
G.711 mu-law	MUST	MUST	MUST	MUST
G.726	*	MAY	MUST	MUST
Other	*	MAY	*	MAY

\* = MUST NOT, but MAY only if recipient capabilities known

## 8. Fax Contents

Fax contents SHOULD be carried according to RFC 2532 [IFAX].

## 9. Interoperability with VPIM v2

Interoperability between VPIM v2 systems and IVM systems can take a number of different forms. While a thorough investigation of how full interoperability might be provided between IVM and VPIM v2 systems is beyond the scope of this document; three key alternatives are discussed below.

### 9.1. Handling VPIM v2 Messages in an IVM Client

If an IVM-conformant client is able to process a content type of multipart/voice-message (by treating it as multipart/mixed) and play a G.726 encoded audio message within it (indicated by a content type of audio/32kadpcm), then a VPIM v2 message that gets routed to that desktop will be at least usable by the recipient.

This delivers a level of partial interoperability that would ease the life of end users. However, care should be taken to ensure that any attempt to reply to such a message does not result in an invalid VPIM v2 message being sent to a VPIM v2 system. Note that replying to an e-mail user who has forwarded a VPIM v2 message to you is, however, acceptable.

A conformant IVM implementation MUST NOT send a non-VPIM v2 message to something it knows to be a VPIM v2 system, unless it also knows that the destination system can handle such a message (even though VPIM v2 systems are encouraged to handle non-VPIM v2 messages in a graceful manner). In general, it must be assumed that if a system sends you a conformant VPIM v2 message, then it is a VPIM v2 system, and so you may only reply with a VPIM v2 compliant message (unless you know by some other means that the system supports IVM).

In addition, it should be noted that an IVM client may not fully conform to VPIM v2, even if it supports playing a G.726 message (e.g., it may not respect the handling of the Sensitivity field required by VPIM v2). This is one reason why VPIM v2 systems may choose not to route messages to any system they do not know to be VPIM v2 compliant.

### 9.2. Dual Mode Systems and Clients

A VPIM v2 system could be extended to also be able to support IVM compliant messages, and an IVM conformant client could be extended to implement VPIM v2 in full when corresponding with a VPIM v2 compliant

system. This is simply a matter of implementing both specifications and selecting the appropriate one, depending on the received message content or the recipient's capabilities. This delivers full interoperability for the relevant systems, providing the capabilities of the target users can be determined.

Note that the mechanism for determining if a given recipient is using a VPIM v2 system or client is outside of the scope of this specification. Various mechanisms for capabilities discovery exist that could be applied to this problem, but no standard solution has yet been defined.

### 9.3. Gateways

It would be possible to build a gateway linking a set of VPIM v2 users with a set of IVM users. This gateway would implement the semantics of the two worlds, and translate between them according to defined policies.

For example, VPIM v2 messages with a Sensitivity of Private might be rejected instead of forwarded to an IVM recipient, because it might not implement the semantics of a Private message, while an IVM message containing content not supported in VPIM v2 (e.g., a PNG image), with a Criticality of CRITICAL, would be rejected in the gateway.

Such a gateway MUST fully implement this specification and the VPIM v2 specification [VPIMV2R2], unless it knows somehow that the specific originators/recipients support capabilities beyond those required by these standards.

## 10. Security Considerations

This document presents an optional gateway between IVM and VPIM systems. Gateways are inherently lossy systems and not all information can be accurately translated. This applies to both the transcoding of the voice and the translation of features. Two examples of feature translation are given in 9.3, but the risk remains that different gateways will handle the translation differently since it is undefined in this document. This may lead to unexpected behavior through gateways.

In addition, gateways present an additional point of attack for those interested in compromising a messaging system. If a gateway is compromised, "monkey in the middle" attacks, conducted from the compromised gateway, may be difficult to detect or appear to be authorized transformations.

Aside from the gateway issue, it is anticipated that there are no new additional security issues beyond those identified in VPIM v2 [VPIMV2R2] and in the other RFCs referenced by this document -- especially SMTP [DRUMSMTP], Internet Message Format [DRUMSIMF], MIME [MIME2], Critical Content [CRITICAL], and Message Context [HINT].

## 11. References

### 11.1. Normative References

- [ADDRESS] Parsons, G., "Voice Profile for Internet Mail (VPIM) Addressing", RFC 3804, June 2004.
- [ADPCM] Vaudreuil, G. and G. Parsons, "Toll Quality Voice - 32 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM) MIME Sub-type Registration", RFC 3802, June 2004.
- [BEHAVIOUR] Parsons, G. and J. Maruszak, "Voice Messaging Client Behaviour", RFC 4024, July 2005.
- [CLID] Parsons, G. and J. Maruszak, "Calling Line Identification for Voice Mail Messages", RFC 3939, December 2004.
- [CRITICAL] Burger, E., "Critical Content Multi-purpose Internet Mail Extensions (MIME) Parameter", RFC 3459, January 2003.
- [DSN] Moore, K., "Simple Mail Transfer Protocol (SMTP) Service Extension for Delivery Status Notifications (DSNs)", RFC 3461, January 2003.
- [DSN2] Vaudreuil, G., "The Multipart/Report Content Type for the Reporting of Mail System Administrative Messages", RFC 3462, January 2003.
- [DSN3] Vaudreuil, G., "Enhanced Mail System Status Codes", RFC 3463, January 2003.
- [DSN4] Moore, K. and G. Vaudreuil, "An Extensible Message Format for Delivery Status Notifications", RFC 3464, January 2003.
- [DRUMSMTP] Klensin, J., "Simple Mail Transfer Protocol", RFC 2821, April 2001.



- [DRUMSIMF] Resnick, P., "Internet Message Format", RFC 2822, April 2001.
- [DUR] Vaudreuil, G. and G. Parsons, "Content Duration MIME Header Definition", RFC 3803, June 2004.
- [HINT] Burger, E., Candell, E., Eliot, C., and G. Klyne, "Message Context for Internet Mail", RFC 3458, January 2003.
- [IFAX] Masinter, L. and D. Wing, " Extended Facsimile Using Internet Mail", RFC 2532, March 1999.
- [KEYWORDS] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [MDN] Hansen, T. and G. Vaudreuil, "Message Disposition Notification", RFC 3798, May 2004.
- [MIME1] Freed, N. and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies", RFC 2045, November 1996.
- [MIME2] Freed, N. and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types", RFC 2046, November 1996.
- [MIME5] Freed, N. and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part Five: Conformance Criteria and Examples", RFC 2049, November 1996.
- [SELECTOR] Allocchio, C., "Minimal GSTN address format in Internet Mail", RFC 3191, October 2001.
- [SCHEMA] Vaudreuil, G., "Voice Messaging Directory Service", RFC 4237, October 2005.
- [VPIMENUM] Vaudreuil, G., "Voice Message Routing Service", RFC 4238, October 2005.
- [VPIMV2] Vaudreuil, G. and G. Parsons, "Voice Profile for Internet Mail - version 2", RFC 2421, September 1998.
- [VPIMV2R2] Vaudreuil, G. and G. Parsons, "Voice Profile for Internet Mail - version 2 (VPIMv2)", RFC 3801, June 2004.

## 11.2. Informative References

- [GOALS] Candell, E., "High-Level Requirements for Internet Voice Mail", RFC 3773, June 2004.
- [G726] CCITT Recommendation G.726 (1990), General Aspects of Digital Transmission Systems, Terminal Equipment - 40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM).
- [G711] ITU-T Recommendation G.711 (1993), General Aspects of Digital Transmission Systems, Terminal Equipment - Pulse Code Modulation (PCM) of Voice Frequencies.

## Authors' Addresses

Stuart J. McRae  
IBM  
Lotus Park, The Causeway<  
Staines, TW18 3AG  
United Kingdom

Phone: +44 1784 445 112  
Fax: +44 1784 499 112  
EMail: stuart.mcrae@uk.ibm.com

Glenn W. Parsons  
Nortel Networks  
3500 Carling Avenue  
Ottawa, ON K2H 8E9  
Canada

Phone: +1-613-763-7582  
Fax: +1-613-967-5060  
EMail: gparsons@nortel.com

## Full 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.

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

## Intellectual Property

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).

## Acknowledgement

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

