

MMUSIC Working Group
Internet Draft
Document: <[draft-dcsgroup-mmusic-resource-00.txt](#)>
Category: Informational

W. Marshall
K. Ramakrishnan
AT&T

E. Miller
G. Russell
CableLabs

B. Beser
M. Mannette
K. Steinbrenner
3Com

D. Oran
Cisco

J. Pickens
Com21

P. Lalwaney
J. Fellows
General Instrument

D. Evans
Lucent Cable

K. Kelly
NetSpeak

F. Andreassen
Telcordia

June, 1999

Integration of Resource Management and Call Signaling for IP Telephony

Status of this Memo

This document is an Internet-Draft and is NOT offered in accordance with [Section 10 of RFC2026](#) [1], and the author does not provide the IETF with any rights other than to publish as an Internet-Draft.

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>.

The distribution of this memo is unlimited. It is filed as <[draft-dcsgroup-mmusic-resource-00.txt](#)>, and expires December 31, 1999. Please send comments to the authors.

1. Abstract

This document describes a two-stage call-setup mechanism, with the resource management protocol interleaved between the two stages. The objective of such a mechanism is to enable deployment of robust IP Telephony services. The first phase ensures that resources are made available before the phone rings and the participants of the call are "invited" to participate. The second phase involves the alerting of the user after the network resources are reserved. This draft presents the call signaling requirements that lead to this design, and the specific SIP extension of an INVITE that does not ring the far-end phone.

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](#) [2].

3. Introduction

For an Internet Telephony service to be successfully used by a large number of subscribers, it must offer few surprises to those accustomed to the behavior of existing telephony services. One expectation is that of connection quality, implying resources must be set aside for each call.

A key contribution of the DCS architecture, as described in [3], is a recognition of the need for coordination between call signaling, which controls access to telephony specific services, and resource

management, which controls access to network-layer resources. This coordination is designed to meet the user expectations and human factors associated with telephony.

While customers may expect, during times of heavy load, to receive a "fast busy" or an announcement saying "all circuits are busy now," the general expectation is that once the destination phone rings that the connection can be made. Blocking a call after ringing the

DCS Group	Category Informational - Expiration 12/31/99	2
	Integration of Resource Mgmt and Signaling	June 1999

destination is considered a "call defect" and is a very undesirable exception condition.

We consider the case where a provider may choose to block a call when adequate resources for the call are not available. It may be argued that best-effort connections may be an alternative in such a case. However, public policy demands that the phone system provide adequate quality at least in certain cases: e.g., for emergency communications during times of disasters. Call blocking enables a provider to meet such requirements. This draft and the overall DCS architecture assumes that the provider blocks calls when resources are unavailable.

It is often the case that calls into a disaster area are blocked, to ensure resources are available for calls out of the disaster area. Such policy-level controls also need to be available for the service provider.

A further expectation of customers of a carrier-grade telephony system is that the first few syllables of a conversation not be clipped. While residential users often introduce a long delay between pickup and use of the voice path (e.g., time taken to lift receiver and placing it to the ear before speaking), automatic answering devices and operators with headsets have no such delay and demand fast cut-through of the voice path.

Coordination between call signaling and resource management is also needed to prevent fraud and theft of service. The coordination between call-signaling and QoS setup protocols ensures that users are authenticated and authorized before receiving access to the enhanced QoS associated with the telephony service.

3.1 Requirements

The basic motivation in this work is to meet and possibly exceed the user expectations and human factors associated with telephony.

In this section, we briefly describe the application requirements that led to the set of DCS signaling design principles. In its basic implementation, DCS supports a residential telephone service comparable to the local telephone services offered today. Some of the requirements for resource management, in concert with call signaling, are as follows:

The system must minimize call defects. These are situations where either the call never completes, or an error occurs after the destination is alerted. Requirements on call defects are typically far more stringent than call blocking. Note that we expect the provider and the endpoints to attempt to use lower bandwidth CODECs as the first line of defense against limited network capacity, and to avoid blocking calls.

DCS Group	Category Informational - Expiration 12/31/99	3
	Integration of Resource Mgmt and Signaling	June 1999

The system must minimize the post-dial-delay, which is the time between the user dialing the last digit and receiving positive confirmation from the network. This delay must be short enough that users do not perceive a difference with post-dial delay in the circuit switched network or conclude that the network connectivity no longer exists.

The system must minimize the post-pickup-delay, which is the time between the user picking up a ringing phone and the voice path being ready for use. This must be short enough that the "hello" is not clipped (both from the called party and the caller). In practical terms, this delay cannot be more than tens of milliseconds beyond the one-way propagation delay.

Call signaling needs to provide enough information to the resource management protocol so as to enable resources to be allocated in the network. This typically requires most if not all of the components of a packet classifier (source IP, destination IP, source port, destination port, protocol) to be available for resource allocation.

3.2 Overview

For acceptable interactive voice communication it is important to achieve end-to-end QoS. The end-to-end QoS assurance implies achieving low packet delay and packet loss. End-to-end packet delay must be small enough that it does not interfere with normal voice conversations. The ITU recommends no greater than 300 ms roundtrip

delay for telephony service. Packet loss must be small enough to not perceptibly impede voice quality or the performance of fax and voice band modems.

If it is found that the network does not have end-to-end QoS resources there are two alternatives: either (1) allow call signaling to proceed with high probability of excessive delay and packet loss which could impair any interactive real-time communication between the participants, or (2) to block the call prior to the called party being alerted. When calls are blocked because of a lack of resources in a particular segment of the network, it is highly desirable that such blocking occur before the called party picks up.

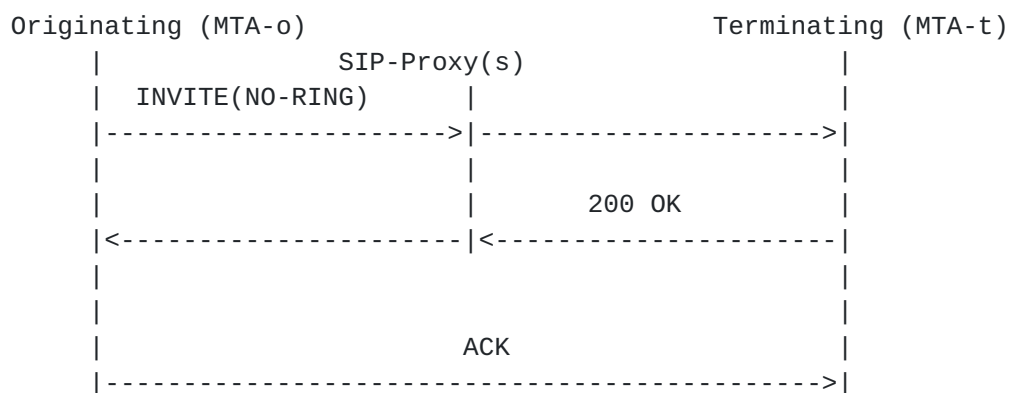
We do expect the endpoints to attempt to use lower bandwidth CODECs, thereby avoiding blocking calls, as the first line of defense against limited network capacity.

The call-signaling and resource reservation must be achieved in such a way that the post-dial-delay must be minimized without increasing the probability of call defects. This means that the number of round-trip messages must be kept at an absolute minimum and messages must be sent directly end system-to-end system if possible.

DCS Group	Category Informational - Expiration 12/31/99	4
	Integration of Resource Mgmt and Signaling	June 1999

Since the post-pickup-delay has a much more stringent requirement than the post-dial-delay, post-pickup-delay must not depend on any signals that are not sent directly end-to-end. Otherwise it may be difficult to avoid quality impairments such as clipping of the initial user "hello".

3.3 Basic Call Flow



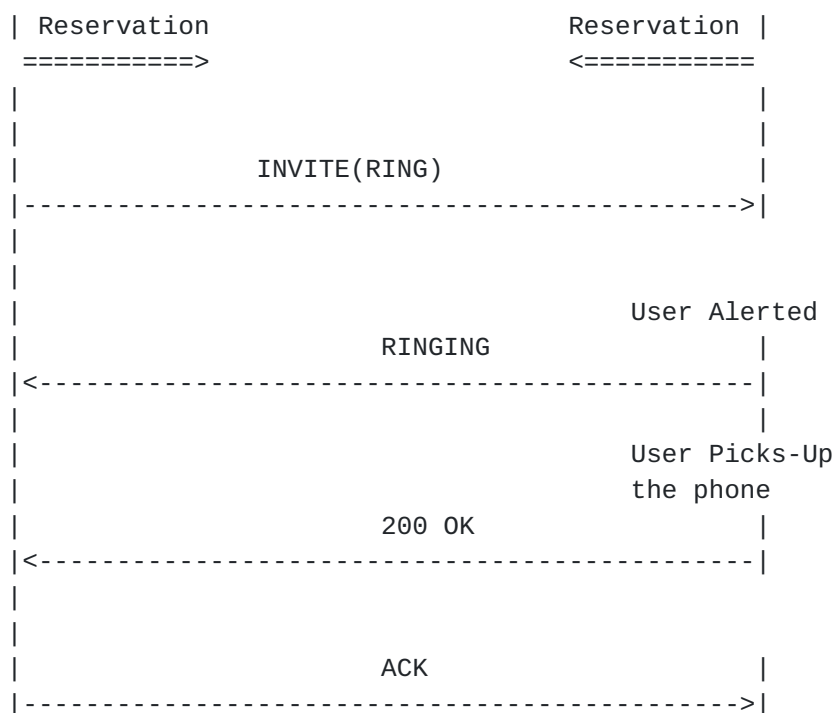


Figure 1

Figure 1 presents a high-level overview of a basic end-point to end-point(MTA-to-MTA) call flow.

When a user goes off-hook and dials a telephone number, the originating MTA (MTA-o) collects the dialed digits and sends the initial call INVITE message to a SIP-Proxy.

Upon reception the initial INVITE message at the destination MTA (MTA-t), the MTA invokes call feature handling, such as call-

forwarding but does not alert the user. This action is indicated by the existence of the NO-RING statement.

Assuming that the call is not forwarded, MTA-t communicates the coding style and bandwidth requirements for the media streams. The 200 OK response to the initial INVITE is forwarded back through the SIP-proxies.

The INVITE request and 200 OK response contain a SIP Contact header to indicate the IP address of the remote MTA to be used for subsequent end-to-end SIP signaling exchanges. MTA-o acknowledges the 200 OK directly to MTA-t.

Once the initial INVITE/200 OK exchange has completed, each MTA has sufficient information regarding the other end-point, bandwidth and other characteristics of the media exchange. Now the endpoints may reserve the resources that will be needed for the media streams.

Once MTA-o has successfully made its reservation, it sends a second INVITE message directly to MTA-t without the NO-RING statement, stating that the MTA-t should alert the user. If MTA-t successfully reserved the resources needed for the call, it responds with a 180 Ringing to indicate that the user is alerted, and that the calling user should be given a ring-back call progress tone. When the called party answers, by going off-hook, MTA-t sends a 200 OK final response directly to MTA-o, which MTA-o acknowledges.

Either party can terminate the call. An MTA that detects an "on-hook" (request to terminate the call) releases the QoS resources held for the connection, and sends a SIP BYE message to the remote MTA, which is acknowledged.

4. Changes to SIP to Support Two-Stage Call Setup

The profile/extension defined in this section allows network resources to be reserved prior to alerting the user and simultaneously reduces the user perceived delays.

The following syntax specification uses the augmented Backus-Naur Form (BNF) as described in [RFC-2234](#) [4].

4.1 SIP Header Extension: NO-RING

The NO-RING request header allows the caller to indicate that an INVITE should not alert the user.

NO-RING = "NoRing:"

A SIP server, on receiving an INVITE with "NO-RING" MUST execute the feature logic associated with the "line" (the logical end-point) and seize the line if it is not busy and the number has not been

forwarded, but the server SHOULD NOT alert the user. Once network resources have been reserved, the originating client MUST issue a second INVITE without the NO-RING header. This second INVITE instructs the destination server to alert the user.

4.2 SIP Profile

This section defines a SIP [[RFC 2543](#)] profile for integrating resource management.

4.2.1 Phase One

Phase one consists of an INVITE message which does not alert the user. The INVITE message can be sent through proxies or (in the case that the desired end point is known, and no resource reservation needs to be authorized by the network, and enhanced QoS is not desired for communication) to the destination itself. Because of the possibility of call forwarding, the originator does not know the final end-point and its characteristics (e.g. CODECs supported, etc.); therefore the resource reservation MUST NOT be attempted yet.

Additionally, to support CODEC selection an SDP session description MUST be included in:

@ The INVITE message MUST contain the NO_RING header.

@ The 200 OK response to the INVITE MUST contain the NO_RING header, and a CONTACT header that indicates the final recipient of the call.

@ The ACK associated with INVITE containing the NO-RING header MUST be send end-to-end.

4.2.2

Phase Two

The message sent during phase two depends on the result of the QoS reservation carried out after phase one. If the reservation is unsuccessful the originating MTA sends a BYE message to terminate the session.

If the reservation is successful, the originating MTA sends a second INVITE message directly to the destination MTA, to the address as indicated by the CONTACT header. This direct end-to-end INVITE message is necessary to minimize post-pickup delay. The 200 OK final response message generated after the called user picks up will also be sent directly end-to-end to minimize post-pickup delay as well as avoid clipping the initial "hello". Otherwise, there is the potential of excessive delay if the 200 OK goes through proxies after the phone is picked up, resulting in a delay to cut-through the voice path.

The second INVITE and 200 OK messages MUST NOT contain the NO-RING statement since during this phase the user should be alerted. The INVITE message MUST NOT carry different SDP descriptions than that communicated in the first phase.

Upon reception of the INVITE message, the terminating MTA responds based upon its QoS reservation. If the reservation is unsuccessful then it MUST return an error indication. If the reservation is successful, the terminating MTA MUST return 180 ringing message and alert the user, or return a 200 OK message directly when the called user picks up.

5. Advantages of the Proposed Approach

The use of two-phase call signaling makes it possible for SIP to meet user expectations that come from the telephony services.

The reservation of resources before the user is alerted makes sure that the network resources are assured before the destination end-point is informed about the call.

The number of messages and the total delay introduced by these messages is kept to a minimum without sacrificing compatibility with classic SIP.

Because the response messages to second phase goes directly end-to-end between the endpoints without traversing the DCS-proxies, the post-dial-delay is minimized. This reduces the chance that the "hello" greeting of the caller (in response to the callee's greeting) gets clipped.

6. Security Considerations

There are no security implication recognized by the extensions proposed in this draft.

7. References

- 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 DCS Group, "Architectural Considerations for Providing Carrier Class Telephony Services Utilizing SIP-based Distributed Call Control Mechanisms", [draft-dcsgroup-mmusic-arch-00.txt](#), June 1999.

DCS Group	Category Informational - Expiration 12/31/99	8
	Integration of Resource Mgmt and Signaling	June 1999

- 4 Crocker, D. and Overell, P.(Editors), "Augmented BNF for Syntax Specifications: ABNF", [RFC 2234](#), Internet Mail Consortium and Demon Internet Ltd., November 1997

8.

Acknowledgments

The Distributed Call Signaling work in the PacketCable project is the work of a large number of people, representing many different companies. The authors would like to recognize and thank the following for their assistance: John Wheeler, Motorola; David Boardman, Daniel Paul, Arris Interactive; Bill Blum, Jon Fellows, Jay Strater, Jeff Ollis, Clive Holborow, General Instruments; Doug Newlin, Guido Schuster, Ikhlaq Sidhu, 3Com; Jiri Matousek, Bay Networks; Farzi Khazai, Nortel; John Chapman, Bill Guckle, Cisco; and Chuck Kalmanek, Doug Nortz, John Lawser, James Cheng, and Partho Mishra, AT&T.

9. Author's Addresses

Bill Marshall
AT&T
Florham Park, NJ 07932
Email: wtm@research.att.com

K. K. Ramakrishnan
AT&T
Florham Park, NJ 07932
Email: kkrama@research.att.com

Ed Miller
CableLabs

Louisville, CO 80027
Email: E.Miller@Cablelabs.com

Glenn Russell
CableLabs
Louisville, CO 80027
Email: G.Russell@Cablelabs.com

Burcak Beser
3Com
Rolling Meadows, IL 60008
Email: Burcak_Beser@3com.com

Mike Mannette
3Com

DCS Group Category Informational - Expiration 12/31/99 9

Integration of Resource Mgmt and Signaling June 1999

Rolling Meadows, IL 60008
Email: Michael_Mannette@3com.com

Kurt Steinbrenner
3Com
Rolling Meadows, IL 60008
Email: Kurt_Steinbrenner@3com.com

Dave Oran
Cisco
Acton, MA 01720
Email: oran@cisco.com

John Pickens
Com21
San Jose, CA
Email: jpickens@com21.com

Poornima Lalwaney
General Instrument
San Diego, CA 92121
Email: plalwaney@gi.com

Jon Fellows
General Instrument
San Diego, CA 92121
Email: jfellows@gi.com

Doc Evans
Lucent Cable Communications
Westminster, CO 80020

Email: n7dr@lucent.com

Keith Kelly
NetSpeak
Boca Raton, FL 33587
Email: keith@netspeak.com

Flemming Andreasen
Telcordia
Piscataway, NJ
Email: fandreas@telcordia.com

DCS Group	Category Informational - Expiration 12/31/99	10
	Integration of Resource Mgmt and Signaling	June 1999

Full Copyright Statement

"Copyright (C) The Internet Society (date). All Rights Reserved.
This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English. The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns. This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS 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."

Expiration Date This memo is filed as <[draft-dcsgroup-mmmusic-resource-00.txt](#)>, and expires December 31, 1999.