

## Media Server Control Protocol Requirements

### Status of This Memo

This memo provides information for the Internet community. It does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

### Abstract

This document addresses the communication between an application server and media server. The current work in IETF working groups shows these logical entities, but it does not address the physical decomposition and the protocol between the entities.

This document presents the requirements for a Media Server Control Protocol (MCP) that enables an application server to use a media server. It will address the aspects of announcements, Interactive Voice Response, and conferencing media services.

### Table of Contents

1. Introduction . . . . .	2
2. Terminology . . . . .	2
3. Requirements . . . . .	3
3.1. Media Control Requirements . . . . .	3
3.2. Media mixing Requirements . . . . .	5
3.3. IVR Requirements . . . . .	6
3.4. Operational Requirements . . . . .	6
4. Security Considerations . . . . .	6
5. Acknowledgments . . . . .	7
6. Informative References . . . . .	7

## 1. Introduction

The IETF conferencing framework in RFC 4353 [CARCH] presents an architecture that is built of several functional entities. RFC 4353 [CARCH] does not specify the protocols between the functional entities since it is considered out of scope.

Based on RFC 4353 [CARCH], this document defines the requirements for a protocol that will enable one functional entity, known as an Application Server (AS), that includes the conference/media policy server, the notification server, and the focus, all defined in RFC 4353 [CARCH], to interact with one or more functional entities, called Media Server (MS), that serves as mixer or media server.

The media server can also be used for announcements and Interactive Voice Response (IVR) functions.

Application servers host one or more instances of a communication application. Media servers provide real time media processing functions. An example of the decomposition of a media server and an application server is described in the media control framework document [MEDIACTRL-FW].

This document presents the requirements for a Media Server Control Protocol (MCP) that enables an application server to control a media server. It will address the aspects of announcements, IVR, and conferencing media services.

The requirements are for the protocol and do not address the AS or MS functionality discussed in the media control framework.

Since the media server is a centralized component, the charter of the working group states that this work will not investigate distributed media processing algorithms or control protocols.

## 2. Terminology

The media server work uses, when appropriate, and expands on the terminology introduced in the conferencing framework [CARCH] and Centralized Conferencing (XCON) framework [XCON-FRMWRK]. The following additional terms are defined:

Application Server (AS) - A functional entity that hosts one or more instances of a communication application. The application server may include the conference policy server, the focus, and the conference notification server, as defined in [CARCH]. Also, it may include communication applications that use IVR or announcement services.

Media Server (MS) - The media server includes the mixer as defined in [CARCH]. The media server plays announcements, it processes media streams for functions like Dual Tone Multi-Frequency (DTMF) detection and transcoding. The media server may also record media streams for supporting IVR functions like announcing participants

Media Resource Broker (MRB) - A logical entity that is responsible for both the collection of appropriate published Media Server (MS) information and supplying of appropriate MS information to consuming entities. The MRB is an optional entity and will be discussed in a separate document.

Notification - A notification is used when there is a need to report event-related information from the MS to the AS.

Request - A request is sent from the controlling entity, such as an application server, to another resource, such as a media server, asking that a particular type of operation be executed.

Response - A response is used to signal information, such as an acknowledgement or error code in reply to a previously issued request.

### 3. Requirements

#### 3.1. Media Control Requirements

The following are the media control requirements:

REQ-MCP-01 - The MS Control Protocol shall enable one or more application servers to request media services from one or more media servers.

REQ-MCP-02 - The MS Control Protocol shall use a reliable transport protocol.

REQ-MCP-03 - The applications supported by the protocol shall include conferencing and Interactive Voice Response media services.

Note: Though the protocol enables these services, the functionality is invoked through other mechanisms.

REQ-MCP-04 - Media types supported in the context of the applications shall include audio, tones, text, and video. Tones media include in-band audio or RFC 4733 payload.

- REQ-MCP-05 - The MS control protocol should allow, but must not require, a media resource broker (MRB) or intermediate proxy to exist with the application server and media server.
- REQ-MCP-06 - On the MS control channel, there shall be requests to the MS, responses from the MS, and notifications to the AS.
- REQ-MCP-07 - SIP/SDP (Session Initiation Protocol / Session Description Protocol) shall be used to establish and modify media connections to a media server.
- REQ-MCP-08 - It should be possible to support a single conference spanning multiple media servers.
- Note: It is probable that spanning multiple MSs can be accomplished by the AS and does not require anything in the protocol for the scenarios we have in mind. However, the concern is that if this requirement is treated too lightly, one may end up with a protocol that precludes its support.
- REQ-MCP-09 - It must be possible to split call legs individually, or in groups, away from a main conference on a given media server, without performing re-establishment of the call legs to the MS (e.g., for purposes such as performing IVR with a single call leg or creating sub-conferences, not for creating entirely new conferences).
- REQ-MCP-10 - The MS control protocol should be extendable, facilitating forward and backward compatibility.
- REQ-MCP-11 - The MS control protocol shall include an authentication component to ensure that only an authorized AS can communicate with the MS, and vice versa.
- REQ-MCP-12 - The MS control protocol shall use some form of transport protection to ensure the confidentiality and integrity of the data between the AS and MS.
- REQ-MCP-13 - Different application servers may have different privileges for using an MS. The protocol should prevent the AS from doing unauthorized operations on a MS.
- REQ-MCP-14 - The MS control protocol requires mechanisms to protect the MS resources used by one AS from another AS since the solution needs to support multiple ASs controlling one MS.

REQ-MCP-15 - During session establishment, there shall be a capability to negotiate parameters that are associated with media streams. This requirement should also enable an AS managing conference to specify the media streams allowed in the conference.

REQ-MCP-16 - The AS shall be able to instruct the MS to perform stream operations like mute and gain control.

REQ-MCP-17 - The AS shall be able to instruct the MS to play a specific announcement.

REQ-MCP-18 - The AS shall be able to request the MS to create, delete, and manipulate a mixing, IVR, or announcement session.

REQ-MCP-19 - The AS shall be able to instruct the MS to play announcements to a single user or to a conference mix.

REQ-MCP-20 - The MS control protocol should enable the AS to ask the MS for a session summary report. The report may include resource usage and quality metrics.

REQ-MCP-21 - The MS shall be able to notify the AS of events received in the media stream if requested by the AS. (Examples - STUN request, Flow Control, etc.)

### 3.2. Media mixing Requirements

REQ-MCP-22 - The AS shall be able to define a conference mix; the MS may offer different mixing topologies. The conference mix may be defined on a conference or user level.

REQ-MCP-23 - The AS may be able to define a custom video layout built of rectangular sub-windows.

REQ-MCP-24 - For video, the AS shall be able to map a stream to a specific sub-window or to define to the MS how to decide which stream will go to each sub-window.

REQ-MCP-25 - The MS shall be able to notify the ASs of who the active sources of the media are; for example, who the active speaker is or who is being viewed in a conference. The speaker and the video source may be different, for example, a person describing a video stream from a remote camera managed by a different user.

REQ-MCP-26 - The MS shall be able to inform the AS which layouts it supports.

REQ-MCP-27 - The MS control protocol should enable the AS to instruct the MS to record a specific conference mix.

### 3.3. IVR Requirements

REQ-MCP-28 - The AS shall be able to instruct the MS to perform one or more IVR scripts and receive the results. The script may be in a server or contained in the control message.

REQ-MCP-29 - The AS shall be able to manage the IVR session by sending requests to play announcements to the MS and receiving the response (e.g., DTMF). The IVR session flow, in this case, is handled by the AS by starting a next phase based on the response it receives from the MS on the current phase.

REQ-MCP-30 - The AS should be able to instruct the MS to record a short participant stream and play it back. This is not a recording requirement.

### 3.4. Operational Requirements

These requirements may be applicable to the MRB, but they can be used by an AS if it has a one-to-one connection to the MS.

REQ-MCP-31 - The MS control protocol must allow the AS to audit the MS state during an active session.

REQ-MCP-32 - The MS shall be able to inform the AS about its status during an active session.

## 4. Security Considerations

This document discusses high-level requirements for MCP. The MCP has some specific security requirements, which will be summarized here at a very high level.

All of the operations and functions described in this document need to be authorized by an MS or an AS. It is expected that MS resources will be governed by a set of authorization rules defined as part of the AS / MS policy. In order for the policy to be implemented, the MS needs to be able to authenticate requests. Normal SIP mechanisms, including Digest authentication and certificates, can be used as specified in [RFC 3261](#) [RFC3261]. These MCP security requirements will be discussed in detail in the framework and protocol documents.

## 5. Acknowledgments

This RFC represents the work from two previous personal works in progress, "Media Control Protocol Requirements" and "Requirements for a Media Server Control Protocol". The authors would like to acknowledge the work of Gary Munson from AT&T Labs, and James Rafferty from Cantata who helped write "Media Control Protocol Requirements", on which this work is based.

## 6. Informative References

- [CARCH] Rosenberg, J., "A Framework for Conferencing with the Session Initiation Protocol (SIP)", [RFC 4353](#), February 2006.
- [MEDIACTRL-FW] Melanchuk, T., Ed., "An Architectural Framework for Media Server Control", Work in Progress, February 2008.
- [RFC3261] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., and E. Schooler, "SIP: Session Initiation Protocol", [RFC 3261](#), June 2002.
- [XCON-FRMWRK] Barnes, M., Boulton, C., and O. Levin, "A Framework for Centralized Conferencing", Work in Progress, November 2007.

## Authors' Addresses

Martin Dolly  
AT&T Labs  
200 Laurel Avenue  
Middletown, NJ 07748  
USA

Phone:  
EMail: [mdolly@att.com](mailto:mdolly@att.com)  
URI:

Roni Even  
Polycom  
94 Derech Em Hamoshavot  
Petach Tikva 49130  
Israel

EMail: [roni.even@polycom.co.il](mailto:roni.even@polycom.co.il)



## Full Copyright Statement

Copyright (C) The IETF Trust (2008).

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, THE IETF TRUST 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).