

XCON
Internet-Draft
Expires: March 10, 2006

R. Even
Polycom
N. Ismail
Cisco Systems, Inc.
September 6, 2005

Conferencing Scenarios
draft-ietf-xcon-conference-scenarios-05.txt

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 March 10, 2006.

Copyright Notice

Copyright (C) The Internet Society (2005).

Abstract

This document describes multimedia conferencing scenarios. It describes both basic and advanced conferencing scenarios involving voice, video, text and interactive text sessions. These conferencing scenarios will help with the definition and evaluation of the protocols being developed in the centralized conferencing XCON working group.

Table of Contents

1.	Introduction	3
2.	Basic Conferencing scenarios	3
2.1.	Ad-hoc conference	4
2.2.	Extension of a Point to point calls to a multipoint call	4
2.3.	Reserved conference	4
3.	Advanced Conferencing scenarios	5
3.1.	Extending a point-to-point call to a multipoint call	5
3.2.	Lecture mode conferences	5
3.3.	Conference with conference aware and unaware participants	6
3.4.	A reserved or ad-hoc conference with conference-aware participants.	6
3.5.	Advanced conference features	6
4.	Scenarios for media policy control	9
4.1.	Video mixing scenarios	9
4.2.	Typical video conferencing scenario	10
4.3.	Conference Sidebar scenario	11
4.4.	Coaching scenario	11
4.5.	Presentation and QA session	12
4.6.	Presence-enabled ad-hoc conference	12
4.7.	Group chat text conferencing	12
4.8.	Interactive text	13
4.9.	Moderated group chat	13
4.10.	Text sidebars	13
4.11.	Conference announcements	13
5.	Security Considerations	14
6.	IANA Considerations	14
7.	Acknowledgements	14
8.	Informative References	14
	Authors' Addresses	15
	Intellectual Property and Copyright Statements	16

1. Introduction

This document describes multimedia conferencing scenarios. The development of these conferencing scenarios is intended to help with definition and evaluation of the requirements for the centralized conferencing (XCON) working group. Although this document uses some definitions and conventions described in the SIP Conferencing Framework document[1], these scenarios are not SIP-specific. The document describes basic and advanced conferencing scenarios. The advanced scenarios assume that the user agents support the set of XCON protocols, identified in the Framework and Data Model for Centralized Conferencing [3], in order to take advantage of the conference functionality. However, note that many of these features can be implemented today using an IVR or web interface to control the conferencing application.

The entities comprising the Conferencing System are the conference that is the center point for signaling and the participants. The participant who initiated the conference is referenced as the initiating participant.

The scenarios described demonstrate different conferencing services. These conferencing services can be offered in a multimedia environment that benefit from having some support in the user agents that enable more robust and easier to use conferencing services. It is up to the conferencing system manufacturers and the conferencing service provider to decide what services can be built and which services are offered to the end users.

The scenarios describe multimedia examples but they are applicable to audio only as well as for audio and video conferences.

Multimedia conferences may include any combination of different media types like audio, video, text, interactive text, or presentation graphics. The conference scenarios are similar but the media handling may be dependent on the media type.

2. Basic Conferencing scenarios

These scenarios enable a conference unaware participant to create, join and participate in a conference. The participant may use out of band signaling to participate in a conference but this is not a mandatory requirement. The Conferencing System has all the functionality it needs in order to supply the service offered to the participants. A typical minimum requirement is that the participant support DTMF tones/signal or provide voice responses to an IVR system.

2.1. Ad-hoc conference

A participant has a service provisioned to him that enables him to start an ad-hoc conference when he calls the Conferencing System. When the participant wants to start a conference he calls the conference service. The participant may be identified by different means including request destination, authenticated identity, or an IVR system using DTMF. The conference is created automatically with the predefined functionality. The participant who has such a service notifies the other participants how to call the conference via external means such as instant message or email. The participant may have the functionality of a Conferencing System and thus can create ad-hoc conference using his own user agent functionality. An example of such a conference is an audio conference initiated by one of the participants who has a conference service that enables him to start a conference when he calls a specific URI. The conference may be created by the first person calling this URI or it may be created only after the owner is authenticated using an IVR system. In the latter case, the other participants may get an announcement and are placed on hold if they call the conference before the owner.

2.2. Extension of a Point to point calls to a multipoint call

This is a basic case. The initiating participant (PA) is in a point to point call with another participant (PB). PA wants to add a third participant (PC) to the call. The initiating participant (PA) cannot provide the Conferencing System functionality on his user agent nor can the other participant (PB). PA and PB do not supports call transfer. PA has a conferencing service using the methods described in 2.1. PA conveys the conference information to PB in the point-to-point call. Both participant disconnect and call the Conferencing System. The Conferencing System may support dial out, for example via DTMF, allowing the initiating participant to call the third party through the Conferencing System.

2.3. Reserved conference

The reservation for this type of conference is typically done by an out of band mechanism and in advance of the actual conference time. The conference identification, which may be a URI or a phone number with a pin number, is allocated by the reservation system. It is sent to all participants using email, IM, etc. The participants join using the conference identification. The conference identification must be routable enabling the allocation of a conference with free resources at the time when the conference actually run. The Conferencing System can also dial out to the conference participants. The participants may not be informed that they are in a conference since their User Agent is not conference aware. The participants may

know, via announcement from the Conferencing System, that they are in a conference and who the other participants are.

3. Advanced Conferencing scenarios

These scenarios assume user agents that support at least call transfer service and a way to communicate information on events from the Conferencing System to the user agent. The Conferencing System may have the ability to discover the capabilities of the participants, for example, to identify if they support call transfer. This section specifies in each scenario the dependencies. An advanced conference can be initiated only by an user agent that has advanced features, but some user agents in the conference may have less functionality.

3.1. Extending a point-to-point call to a multipoint call

The initiating participant is in a point-to-point call and wants to add a third participant. The initiating participant can start a multipoint call on a conferencing bridge known to him. The extension can be without consultation, which means that he moves the point-to-point call to the Conferencing System and then adds the third party (this can be done in various ways). Alternatively the extension can be done with consultation, which means that he puts his current party on hold, calls the third party and asks him to join the conference, and then transfers all the participants to the Conferencing System.

3.2. Lecture mode conferences

This conference scenario enables a conference with a lecturer who presents a topic and can allow questions. The lecturer needs to know who the participants are and to be able to give them the right to speak. The right to speak can be based on floor control or an out of band mechanism.

In general, the lecturer is seen/heard by the conference participants and often shares a presentation or application with the other participants.

A participant joining this type of conference can get the identity of the lecturer and often the identities of the audience participants.

This type of conference may have multiple media streams. For example, if simultaneous language translation is available, a participant has the option of selecting the appropriate language audio stream. Multiple video streams could include the speaker's face and a whiteboard/demonstration stream.

3.3. Conference with conference aware and unaware participants

A conference can include participants that are a mix of conference aware and unaware participants. Those participants may be conference unaware participants using a proxy function that proxies the advanced functionality between the different protocols and the Conferencing System. For example, an IVR system or a web page interface can be used to provide additional functionality.

3.4. A reserved or ad-hoc conference with conference-aware participants.

The initiating participant calls the Conferencing System using, for example, a unique identifier in order to start the conference. The Conferencing System may use some authenticating method to qualify the participant. The other participants may call the Conferencing System and join the conference. The Conferencing System is able to find the capabilities of the participants. In case of a reserved conference the Conferencing System starts the conference at the scheduled time. The participants may join by calling the conference URI or the Conferencing System may call them. The conference may have privilege levels associated with a specific conference or participant. The privileges are for the initiating participant and for a regular participant; the initiating participant may delegate privileges to the other participants. The privileges allow functionality as defined in the next section.

3.5. Advanced conference features

The following features can be used in all the advanced conferencing scenarios. In the examples given in this section, when referring to a participant that has a functionality it means a participant with the right privileges. These scenarios may be available in the advanced conferencing scenarios and are common in many conferencing applications. This is not a requirement list, rather some examples of how specific functions may be used in a conference.

- o Add Participants - A participant may add a new participant to the conference. This can be done, for example, by instructing the Conferencing System to call the participant or by the first participant calling the new participant and pointing him to the conference.
- o Delete Participant - A participant may delete participants from the conference if he can identify them.
- o Changing User Agent/Modes - During the course of a conference, a participant may switch between user agents with different capabilities while still remaining part of the conference. For example, a participant may initially join using a mobile phone and

then switch to a desk top phone. Or a participant may join with a phone, discover that the conference has video streams available, and switch to a video phone.

- o Changing Media - During the conference a participant may be able to select different media streams than the one he had when he joined the conference. An example is a participant that initially joined the conference as an audio participant. The participant is unable to understand the conversation properly and he learns that there is also an interactive text available, he will ask to receive also the text stream.
- o Authenticate participants - A participant can authenticate other participants who want to join the conference. This can be done, for example, in a video conferencing session by creating a sidebar between the two participants allowing the authenticating participant to talk with the new participant and verify his identity.
- o Authorize participants - A participant can authorize other participants in order to allow them to join the conference. This can be done implicitly by assigning a password to the conference or to each participant and letting the Conferencing System decide if the new participant is allowed to join. the authorization can be done explicitly by directing the entered password to the initiating participant who will authorize each participant. The conferencing system may use an authentication mechanism to authenticate the participants.
- o Controlling the presentation of media - During the conference the participant may be able to manage whose media is being sent to each participant. For example, the participant may be able to decide that he wants to be the speaker and all the rest are listeners; he may also specify whose media he wants to receive. The participant may be able to mute a media stream during the conference.
- o Giving privileges - The participant may want, during the conference, to give a privilege to another participant. The assigning of privileges may be implicit when requested or explicit by asking the participant to grant a privilege.
- o Side conferences or sidebars - The participant may want to create a side conference that include some of the main conference participants. When the side conference is done the participants return to the main conference. A sidebar may have the same functionality as the main conference. There can be several sidebars scenarios:
 1. Basic sidebar is based on the capabilities of two participants to have two calls at the same time, with a point to point call in parallel to the main conference. It is user agent implementation specific whether to automatically mix both call's streams or allow the participant to manually switch

between them.

2. Conferencing System based sidebar uses the Conferencing System to create the sidebar and compose the relevant sidebar stream mixes. These mixes can include the main conference as an incoming stream to the mix. Mechanisms to signal the creation of the sidebar, invite participants and control the mixes should be available.

For example, participants in an audio sidebar may not be heard by the rest of the conference. However, the main conference audio may be mixed in the sidebar, but at a lower volume, or in a different channel. Another example, a sidebar can have a different media type from the main conference; a video call can have an audio sidebar where the other participants can see the sidebar participants talking but can not hear them; or an audio or video conference may have a text sidebar.

- o Conference information - When a participant joins the conference he is announced to the participants. An announcement may be available when he leaves the conference. The participants may query the conferencing system for the current participants of a specific conference. This conference information may include other information, for example, the media streams available in the conference.
- o Extending of a conference - Reserved conferences and ad-hoc conferences may have a time limit. The Conferencing System informs the participants when the limit is approaching and may allow the extension of the conference.
- o Adding and removing a media type to the conference - A participant may want to start a data presentation during a conference. He may want to distribute this new media to all the participants. The participant asks the Conferencing System to start the new media channel and to allow him to send data in the new channel.
- o Audio-only participants - In a multimedia conference some of the participants who want to join may have no way to send and receive all the media types. Typically they can send and receive audio. Such participants join the conference as audio-only participants. The general case is that participants may send and receive only part of the media streams available in the multi media conference.
- o Passive participants - In a conference some participants may be listeners to all or part of the media streams, but be invisible to all the other participants.
- o Recorders - A recorder can be added to the conference. A recorder can record all streams or a subset of the streams. Recorders may be turned on and off during the conference. Recorders may be used for "role call" scenario in order to record a participant name. This name can be announced at a later stage automatically or based on a participant request. A recorder is a case of a passive

participant.

- o Whisper/Private Message - A participant can send a one way message (text, audio, or even some other media) to another participant that is immediately rendered. This differs from a sidebar in that it is immediate and creates no long-lived session.
- o Human operator - A participant may ask for assistance from a human operator during the conference.

4. Scenarios for media policy control

During a conference media streams may be controlled by authorized participants using either a media control protocol or a third party application. This section describes some typical media control scenarios. The conference can be of any size. Some of the media control scenarios are typical to specific conference sizes. As a general rule larger conferences scenarios tend to be more centrally managed or structured.

The mixing of media in a conference may start when the conference starts or when the initiating participant joins. In the later case, early participant may be put on hold and get "music on hold".

The scenarios apply to audio conferences as well as to multimedia conferences. There are some specific information about the mixed video layout and about interactive text discussed below.

4.1. Video mixing scenarios

For video the participant selects one of a set of pre-defined video presentations offered by the server. Each video presentation is identified by a textual description as well as an image specifying how the presentation appears on the screen. In this scenario by choosing a video presentation the participant chooses how many video streams (participants) are viewed at once and the layout of these video streams on the screen.

The contents of each sub-window can be defined by a conference policy and/or controlled by authorized participants. It may also be possible to have multiple mixes per conference, possibly as many as there are participants. (Note that the same flexibility may be afforded to audio mixes as well.).

The following are a list of typical video presentations; there are other layouts available today in commercial products:

- Single view: This presentation typically shows the video of the loudest speaker

- Dual View: This presentation shows two streams. If the streams are to be multiplexed in one image (typical of centralized servers) the multiplexing can be:

1. Side by side with no altered aspect ratio and hence blanking of parts of the image might be necessary if the streams are to be combined as one image.
2. Side by side windows with altered aspect ratios and hence blanking parts of the image is not necessary. The mixer handles the cropping of the images.
3. One above the other windows with no altered aspect ratio
4. One above the other windows with altered aspect ratio

- Quadrate view: This presentation shows 4 streams. If the streams are multiplexed into one image (centralized server) they are arranged in a 2x2 style. Note that in this style the aspect ratios are maintained.

- 9 sub-picture view: This presentation shows 9 streams. If the streams are to be multiplexed in one image they are arranged in a 3x3 style. In the multiplexing case cropping is performed under the discretion of the mixer.

- 16 sub-picture view: This presentation shows 16 streams. If the streams are to be multiplexed into one image they are arranged in a 4x4 style. In this style the aspect ratios are maintained and no cropping or blanking is needed.

- 5+1 sub-picture view: This presentation shows 6 streams. If the streams are to be multiplexed into one image then the pictures are laid so that one sub-window occupies 4/9 of the screen while the other five occupy 1/9 of the screen each.

4.2. Typical video conferencing scenario

This scenario is known as voice activated video switch. Every participant hears the N loudest participants but he does not hear himself. All the participants see the loudest speaker; the loudest speaker may see the previous loudest speaker. This mode is typical to small conference.

A participant with proper authorization can exclude one or more participants from the audio or video mix. An indication might be displayed to the affected participants indicating that they are not being seen/heard.

A participant with proper authorization can manipulate the gain level associated with one or more audio streams in the mix.

4.3. Conference Sidebar scenario

An authorized participant creates a side bar. The participant selects whether the sidebar should include the media from the main conference or not and the audio gain level associated with the main conference audio.

A participant invites participants to the sidebar and upon acceptance they start receiving the sidebar media as specified by the sidebar creator. If the new participant is not a participant of the conference, but just the sidebar, the participant only receives the sidebar media without the media of the main conference.

A participant with the right authorization can move another participant into the sidebar with no indication, in which case the participant suddenly start receiving the sidebar media.

Sidebar participants with the right authorization can select to hear or not hear the main conference audio mixed with the sidebar audio

A participant can be a participant to more than one sidebar but can only actively participate in one.

A participant can jump back and forth between the main conference and one or more sidebars.

4.4. Coaching scenario

This is a call center or a remote training session where there is a supervisor who can monitor the conference. There are the supervised participants that may be the call center operators or the teachers. A participant in the conference may be a supervised participant or a "customer".

The supervisor is a hidden participant and is not part of the participant roster.

The supervised participants might get an announcement/tone indicating that the supervisor has joined. The other participants do not hear the announcement.

The supervisor listens to / sees the session but can only be heard / seen by the supervised participant.

The supervisor can become a normal participant, in which case the

participants see the supervisor as part of the roster and start hearing and seeing him.

4.5. Presentation and QA session

An example is an earning call scenario in which a group of presenters deliver material to a group of people. After the presentation is finished a QA session is opened.

The conference is created as a panel and the panel participants are identified. Only their streams are mixed.

After the end of the presentation the session chair changes the conference type to normal and now streams from all participants may be mixed. Alternatively a floor control protocol can be used. The chair can grant the right to speak by adding the participant, whose turn it is to ask a question, to the conference mix.

4.6. Presence-enabled ad-hoc conference

A presence-enabled ad-hoc conference, sometimes described as "walkie talkie" service, is a scenario in which a participant sends media to the other participants of the conference after receiving a confirmation of the other participants' availability. For example, a participant presses a talk button, which checks the presence of the participants to see if they are available for communication. If they are, a confirmation tone is played and the participant can then talk, which results in the media being sent to the other participants in the conference. These types of conferences tend to be long lived, hence the need for presence to ensure that the other participants are still available. The ad-hoc nature of the conference means that the participant list can be changed at any time. Floor control can be used to allow other participants to speak, as the conference is usually half-duplex in nature.

4.7. Group chat text conferencing

Group chat is a common scenario for text messaging in which a participant joins (or enters) a chat room in which text messages from participants are rendered in a single window and attributed to the participant that sent the message. Changes in conference membership are often announced in the text window itself (e.g. "Alice has just entered the room. Bob has just departed."). Note that a real-time transcription/closed captioning service can provide a similar window in which audio media is converted into interactive text. "Nick names" or aliases are often chosen by participants or assigned by the Conferencing System and used as handles within the room.

4.8. Interactive text

Interactive text is using RTP to carry text one character at a time providing real-time interactivity, as described in [RFC2793\[2\]](#). The interactive text session may be the main conference itself, or it may be used in conjunction with other media types. Interactive text may be used to represent the audio in the conference using some translation services. There can be more than one such stream where each text stream is in a different language. These text streams may be used as subtitles to the audio stream. The translation from text to speech and back is done by transcoders. Those transcoder have similar functionality to transcoders between different audio or video algorithms.

The conference participants should be able to select to receive those text streams with the conference audio or without it.

4.9. Moderated group chat

A moderated group chat scenario for text messaging is similar to group chat but with all text messages sent to the group being filtered/approved by a moderator. Note that the moderator can be a human or an application. The moderator also often has the ability to remove participants and provide feedback on their submissions (e.g. provide warnings before removal).

4.10. Text sidebars

Interactive text or instant messaging sidebars are perhaps the most common sidebars in conferences today. Often the text sessions are separate from the conference. However, there are some advantages to having text sessions be a sidebar and as a result a part of the main conference. For example, a conference which is providing anonymity/aliases to participants can also provide anonymous/alias sidebars. A text sidebar can also benefit from other security/logging/recording services provided by the Conferencing System.

Another use of a text sidebar is a text-only conversation/discussion between two or more conference participants who at the same time are following the main conference.

4.11. Conference announcements

The conference moderator may be able to play announcements to all the conference participants. The announcement may be pre-recorded or composed by the moderator before sending them. The announcements may be text, audio or audio visual. An example is a conference with several audio break-out sessions going on. At some point in the

time, the moderator wants to record an audio message like "in 5 minutes, everyone please come back to the main meeting" and then play that message to all of the breakout sessions.

5. Security Considerations

Conferences generally have authorization rules about who may or may not join a conference, what type of media may or may not be used, etc. This information, sometimes called the conference policy or common conference information, is used by the Conferencing System to admit or deny participation in a conference. For the conference policy to be implemented, the Conferencing System needs to be able to authenticate potential participants. The methods used depend on the signaling protocols used by the conference. This can include a challenge/response mechanism, certificates, shared secret, asserted identity, etc. These conference-specific security requirements are discussed further in the XCON requirements and framework documents.

6. IANA Considerations

There are no IANA considerations associated with this specification.

7. Acknowledgements

Thanks to Brian Rosen for contributing conferencing scenarios.

Thanks to Alan Johnston for going over the document and adding some more scenarios; to Keith Lantz, Mary Barnes and Dave Morgan for carefully reading the document.

8. Informative References

- [1] Rosenberg, J., "A Framework for Conferencing with the Session Initiation Protocol", [draft-ietf-sipping-conferencing-framework-04](#) (work in progress), October 2003.
- [2] Hellstrom, G., "RTP Payload for Text Conversation", [RFC 2793](#), May 2000.
- [3] Barnes, M., Boulton, C., and O. Levin, "A Framework and Data Model for Centralized Conferencing", [draft-barnes-xcon-framework-02](#) (work in progress), February 2005.

Authors' Addresses

Roni Even
Polycom
94 Derech Em Hamoshavot
Petach Tikva 49130
Israel

Email: roni.even@polycom.co.il

Nermeen Ismail
Cisco Systems, Inc.
170 West Tasman Drive
San Jose 95134
CA USA

Email: nismail@cisco.com

Intellectual Property Statement

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.

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 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.

