A Framework and Data Model for Centralized Conferencing
draft-barnes-xcon-framework-02

Status of this Memo

This document is an Internet-Draft and is subject to all provisions of Section 3 of RFC 3667. 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 become aware will be disclosed, in accordance with RFC 3668.

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 August 18, 2005.

Copyright Notice

Copyright (C) The Internet Society (2005).

Abstract

This document defines the framework for Centralized Conferencing (XCON), which is applicable to participants using different call control signaling protocols and exchanging media over networks with
potentially different characteristics. The XCON Framework defines
the XCON data model, the XCON logical entities, the naming
conventions, and outlines the standard conferencing protocols
(complementary to the call control signalling protocols) for building
advanced conferencing applications. The framework binds all the
derived components together for the benefit of conferencing system
builders.

Table of Contents

1. Introduction .................................................. 3
2. Conventions and Terminology ................................. 3
3. Overview ........................................................ 6
4. XCON Data Model and Identifiers ............................. 8
   4.1 Conference Object Identifiers ............................ 9
   4.2 Conference Object Type ................................ 10
   4.3 Common Conference Information .......................... 11
   4.4 Conference Template ..................................... 12
   4.5 Conference policies ....................................... 12
5. Conferencing System Realization ............................. 13
   5.1 Cloning Tree ................................................ 14
   5.2 Ad-hoc Example ............................................. 16
   5.3 Advanced Example ......................................... 17
   5.4 Scheduling a ‘Conference Object for Reservation’ .......... 19
6. Conferencing Mechanisms ....................................... 22
   6.1 Call Control Signalling .................................. 22
   6.2 Notifications ................................................. 22
   6.3 Conference Control Protocols ............................. 23
       6.3.1 CPCP .................................................. 23
       6.3.2 CCCP .................................................. 24
   6.4 Floor Control ................................................ 24
       6.4.1 User Identifier ....................................... 26
7. Conferencing Scenario Realizations ........................... 27
   7.1 Participant Manipulations ................................. 27
   7.2 Media Manipulations ....................................... 29
   7.3 Sidebar Manipulations ..................................... 29
   7.4 Whispering or Private Messages .......................... 29
   7.5 Conference Announcements and Recordings .................. 29
8. Relationships between SIPPING and XCON Frameworks .......... 29
9. Security Considerations ........................................ 30
10. IANA Considerations ............................................ 32
11. Acknowledgements .............................................. 32
12. References ...................................................... 32
   12.1 Normative References .................................... 32
   12.2 Informative References ................................. 32
       Authors’ Addresses ........................................ 34
       Intellectual Property and Copyright Statements ........... 36
1. Introduction

This document defines the framework for Centralized Conferencing (XCON), which is applicable to participants using various call signalling protocols (such as SIP, H.323, Jabber, HTML, PSTN, etc.) and exchanging media over networks with potentially different characteristics.

The XCON Framework defines the XCON data model, the XCON logical entities, the naming conventions, and outlines the standard conferencing protocols (complementary to the call control signalling protocols) for building advanced conferencing applications. The framework binds all the defined components together for the benefit of conferencing system builders.

The XCON Framework is compatible with the functional model presented in the SIPPING Conferencing Framework [9]. Section 8 of this document discusses the relationship between the XCON Framework and the SIPPING Conferencing framework, in the context of the XCON architecture.

2. Conventions and 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 [1] and indicate requirement levels for compliant implementations.

The XCON Framework document both generalizes, when appropriate, the SIPPING Conferencing Framework [9] terminology and introduces new concepts as listed below.

Active Conference: The term Active Conference refers to a Conference Object that’s been created (for example, using a "blueprint") and "activated" via the allocation of its identifiers (i.e. Conference Object Identifier, Conference Identifier) and the associated Focus.

Call (Control) Signalling: The protocol used between a participant and a Focus. In this context, the term "call" means a channel or session used for media streams establishment. Protocol examples include SIP, H.323, MSRP, Jabber, HTML and PSTN signalling. This interface is not limited to the listed protocols, however, other than references to general functionality required (e.g. establishment and teardown), details of these protocols are outside the scope of this document.
Common Conference Information: The data type (i.e. the XML schema) which is used to represent the fixed information part of a Conference Object. It includes a common set of definitions for basic conference features, such as conference identifiers, membership, signalling, capabilities, media, etc.

Conference Control Protocol (CCP): A protocol used by clients to manipulate a Conference Object [Section 6.3].

Conference Factory: A logical entity that, upon request, generates unique URI(s) to identify and represent a conference Focus. The Conference Factory is typically used in the conference creation process via a signalling interface and non-signalling methods such as Conference Control Protocol [Section 6.3] and proprietary mechanisms.

Conference Identifier (ID): The call signalling protocol-specific URI that uniquely identifies a Conference Instance and a conference Focus. Note that a Conference Instance can have more than a single Conference Identifier, for example, for each call signalling protocol supported.

Conference Instance: Represents the internal implementation of a conference; addressable using the Conference Identifier. There is a one-to-one mapping between a Conference Instance and a conference Focus.

Conference Object: A Conference Object is a logical representation of a Conference Instance at a certain stage (e.g. description upon conference creation, reservation, activation, etc.), which a Conferencing System maintains in order to describe the system capabilities and to provide access to the available services provided by the Conferencing System. All Conference Objects are of type 'Conference Object Type' which is comprised of two distinct sub components; the 'Common Conference Information' and the 'Conference Template' (see definitions in this section for details).

Conference Object Identifier (ID): A [TBD schema name] URI which uniquely identifies a Conference Object and is being used by a Call Control Protocol [Section 6.3].

Conference policies: A term which is used to collectively refer to a virtual set of rights, permissions and limitations pertaining to operations being performed on a certain Conference Object. The term is described in more details in Section 4.5.


Conferencing System: The term used to refer to a conferencing solution (system) that is based on the set of specifications and is built using the protocols and the data model defined by the XCON working group within the IETF.
Conference Template: The data type (i.e. the XML schema) which used to represent the variable information part of the Conference Object. It can be included in the Conference Object to represent enhanced conferencing capabilities (e.g. media mixers, etc.) and/or user interface abstraction.

Floor: A term which is used to refer to a set of data or resources associated with a Conference Instance, for which a conference participant (or group of participants) is granted temporary input access.

Floor Chair: A Floor control protocol compliant client (human participant or automated entity) who is authorized to manage access to one floor (grants, denies, or revokes a floor). The floor chair does not have to be a participant in the Conference Instance.

Focus: Focus is a logical entity that maintains the call signalling interface between each participating client (i.e. participant) and the Conference Instance. As such, the Focus acts as an endpoint for each of the supported signalling protocols and is responsible for all primary conference membership operations (e.g. join, leave, update the Conference Instance) and for media negotiation/maintenance between a conference participant and the Focus. There is a one-to-one mapping between the Focus and its Conference Instance. Focus is addressed by explicitly associated unique conference URIs for each signalling protocol, supported for its Conference Instance.

Media Graph: The logical representation of the flow of media for a conference.

Media Mixer: A logical entity that has the capability to combine media inputs of the same type (or/and transcode the media) and distribute the result(s) to a single or multiple outputs. In this context, the term "media" means any type of data being delivered over the network using appropriate transport means, such as RTP/RTCP (defined in RFC 3550), Message Session Relay Protocol (defined in [25]), etc.

Registered Template Definition: ‘Registered Template Definition’ is the term used to refer to an official standards document (RFC) that defines and registers a Conference Template schema with the appropriate standards body (IANA). A ‘Registered Template Definition’ also includes any complimentary textual information in relation to the conference template schema and its meaning.

Role: Represents differing membership categories that a conference participant can assume within a conference. Each category has a difference set of conference operations that a participant can perform. A default (e.g. Standard Conference Participant) ‘Role’ will always exist which provides a standard user with a set of basic conference operations. Any user with appropriate authentication and authorization may assume a role that has a wider set of conference operations that are otherwise not
available (to a standard Conference Participant) - e.g. A 'Role' called 'Conference Moderator' may exist that has additional conference operations such as 'modify conference end time'.

Sidebar: TBD.
Whisper: TBD.

3. Overview

A centralized conference is an association of endpoints (called conference participants) with a central endpoint (called a conference Focus) where the Focus has direct peer-wise relationships with the participants by maintaining a separate call control signalling interface with each. Consequently, in this tightly coupled model, the call control signalling graph is always a star.

The most basic conference supported would be an ad-hoc unmanaged conference, which would not necessarily require any of the functionality defined within this framework. For example, it could be supported with basic SIP signalling functionality with a participant serving as the Focus; the SIPPING Conferencing Framework [9] together with the SIP Call Control Conferencing for User Agents[15] documents address these types of scenarios.

An XCON-compliant Conferencing System, in addition to the basic features, can offer richer functionality including dedicated conferencing applications with explicitly defined capabilities, reserved reoccurring conferences, and the standard protocols for managing and controlling different conference aspects.

The core requirements for tightly managed centralized conferencing are outlined in [10]. These requirements are applicable for conferencing systems using various call control signalling protocols, not restricted to SIP alone. Additional conferencing requirements are provided in [12], [13], and [14].

The XCON architecture is built around a fundamental concept of a Conference Object. A Conference Object is a logical representation of a Conference Instance. For conference creation, a Conference Object provides a "blueprint" representing the System Capabilities. A conference object also provides the logical representation of a conference during each of the various stages of a conference (e.g. reservation, active, completed, etc.). A Conference Object is accessible using XCON protocols (e.g. a Conference Control Protocol [Section 6.3]).

A Conferencing System can support any subset of the conferencing functions depicted in the Conferencing System logical decomposition in Figure 1 and described in this document. Nevertheless, typically
advanced functions could not be implemented without implementing others. For example, the Notification Service is an essential component required for implementing almost any advanced functionality and being used, among other things, for correlating of information (such as list of participants with their media streams) between otherwise distinct functions.
The media graph of a conference can be centralized, de-centralized, or any combination of both and potentially differ per media type. In the centralized case, the media sessions are established between the focus and each one of the participants. In de-centralized (i.e. distributed) case, the media graph is a (multicast or multi-unicast) mesh among the participants. Consequently, the media processing (e.g. mixing) can be performed either by the focus alone or by the participants. The concepts in this framework document clearly map to a centralized media model. They can also apply to the de-centralized media case, however, the details of such are left for future study by XCON charter.

Section 4 of this document describes the usage of the Conference Object in more details. Section 5 of this document describes how a Conferencing System is built using the defined data model and how the Conference Object trees are maintained. Section 6 describes the fundamental conferencing mechanisms and provides a high level overview of the XCON protocols. Section 7 then provides realizations of various conferencing scenarios, detailing the manipulation of the Conference Objects using XCON defined protocols. Section 8 of this document summarizes the relationship between the XCON Framework and the SIPPING Conferencing Framework.

4. XCON Data Model and Identifiers

This section first provides details of the identifiers associated with the XCON constructs (e.g. Conference Object and Conference Instance). The constructs of the XCON Data Model are then described in detail.

(Editor’s Note: The exact XML schema of the Conference Object (i.e. the organization of the Common Conference Information and the Conference Template) is an open issue, which has been summarized as the three potential alternatives below:

1. The top-level information is the template section, and it contains a subsection that is the common conference information. This has the advantage that the schema can all be well defined: because the common conference information is known at the time the template is developed, the appropriate schema definition can be inserted into the template schema. The downside is that this setup requires navigation of the template information to get to the common information, which is likely to be manipulated most frequently.

2. The top-level information is the common conference information, which contains the template information. This has the advantage that
clients don’t even need to care about the template being used to allow rendering and control over basic conference functionality (which will suffice for many clients (e.g. those with limited screen). The downside is that the common conference information schema must include an extension point to allow new templates to hook into the schema. This may make schema validation more difficult.

3. The template information and common conference information are conveyed as two separate objects (e.g. using multipart MIME). This provides the benefits of allowing completely separate schema, straightforward schema validation, and easy access to the common conference information. The downside is that any mechanism for separating the schema is going to add some amount of protocol complexity and the need for synchronization between the two data objects. Note that it has been argued that it is increasingly difficult to find a potential client of the XCON protocols that doesn’t already support multipart MIME).

The model put forth in this document is the result of the consensus reached during the XCON second interim meeting in Boston and depicts option 2 above. With slight adaptations it can also support option 3. }

4.1 Conference Object Identifiers

In order to make a certain Conference Object accessible, the Conferencing System allocates a unique URI for each distinct Conference Object in the system. A Conference Control Protocol [Section 6.3] can be used for manipulating a particular Conference Object and for retrieving its state.

[Editor’s Note: The URI schema name TBD and registered with IANA.]

A Conference Instance is an internal implementation of a conference, which is accessible by call signalling means. A single Conference Instance can be internally mapped to a single or multiple Conference Objects; each of them accessible by a Call Control protocol. The mapping details are an implementation decision and out of scope of this framework.

There is a one-to-one mapping between a Conference Instance and a conference Focus. The Focus is addressed by explicitly associating unique conference URIs for each signalling protocol, supported for its Conference Instance. The state of a Conference Instance may be retrieved by a call signalling dependent mechanism. For example, the Conference Package mechanism [11] can be used for SIP. Note, that in this case, the subscription to the state information is performed based on the Conference Focus SIP URI.
4.2 Conference Object Type

The XCON data model defined in this framework has no strict separation between conference membership, conference media information and the related policies (i.e. the capabilities and limitations for each). This approach meets the requirement in many conference control operations to enable synchronized access to the conference policies as a whole, to the conference state as a whole, and for receiving notifications about changes to either.

A Conference Object is of type ‘Conference Object Type’ which is comprised of two distinct components: the ‘Common Conference Information Type’ and the ‘Conference Template Type’, as illustrated in Figure 2. Each of these types is a placeholder for including potentially multiple sub-types.
Since, in an XCON-compliant system the same Conference Object Type is used for representation of a conference for different purposes, such as expressing Conferencing System capabilities, reserving conferencing resources or reflecting the state of ongoing conferences, each of the two components (i.e., the Common Conference Information and the Conference Template) is syntactically optional to be included in a particular Conference Object. Section 5 describes the usage semantics of the Conference Objects.

4.3 Common Conference Information

The common conference information section contains the core information that is utilized in any conference and can be abstracted regardless of the specific conference media nature (e.g. the mixing
algorithms performed, the advanced floor control applied, etc.). Typically, participants with read-only access to the conference information would be interested in this common information only.

The basic common conference information can be represented using the conference-info-type schema as defined in [11]. This schema contains the definitions for representation of the Conference Object capabilities, membership, roles, call control signalling and media statuses relevant to different stages of the conference life-cycle.

New XCON specifications can extend the basic conference-info-type and introduce additional schemas to be used within the common conference information type placeholder.

4.4 Conference Template

The concept of a "Conference Template" is introduced to abstract the complexity and the details of the "mixer" and other conferencing features and to allow for easy user interface abstraction for advanced Conferencing Systems.

Each Conference Template needs to be registered with IANA. The IANA registration needs to point to an RFC having the text description of the feature behavior and optionally the XML definition allowing the feature presentation, configuration, and management. RFCs of this kind are referred by XCON documents as a "Registered Template Definitions".

Typically, a conference template would contain the information about the specific media mixing details, the associated client roles and the available floor controls. This information would allow authorized clients to manipulate mixer’s behavior and the resultant distribution of the media to all or individual participants. By doing so, a client can change its own state and/or state of other participants in the conference.

A conference template can also include an abstract user interface definition in terms of sliders, radio boxes, etc. for simplifying user interaction with a specific non-trivial feature.

4.5 Conference policies

Conference policies is the term used to collectively refer to a virtual set of rights, permissions and limitations pertaining to operations being performed on a certain Conference Object.

The XCON architecture identifies three different levels of Conference policies:
Data Access Rights: The Data Access Rights is the data object describing the read/write access privileges for accessing the Conference Object as a whole. This access would usually be granted and defined in terms of giving the read-only or read-write access to clients with certain roles in the conference. For details see [TBD].

Conference Control Limits and Permissions: The Conference Control Limits and permissions is specified as an integral part of the Conference Object Type. This term collectively references the data objects containing the allowed ranges for other data objects (e.g., maximum number of participants) and lists of clients allowed to perform certain operations on a conference object. For example, the "allowed to join" list of participants is consulted to decide who is allowed to join. The entries in the list can specify the identity of an individual user (joe@example.com), a role, a domain (*.example.com), etc. For details see [TBD].

Common Conference Rules: The Common Conference Rules are envisioned as a more general rule mechanism, beyond the functionality provided by the Conference Control Limits and Permissions. These Common Conference Rules are left for future study by XCON charter.

5. Conferencing System Realization

XCON-compliant implementations can range from systems supporting ad-hoc conferences, with default behavior only, to sophisticated systems with the ability to schedule re-occurring conferences (each with distinct characteristics), being integrated with external resource reservation tools, and providing snapshots of the conference information at any of the stages of the conference life-cycle.

A Conference Object is the logical representation of a Conference Instance at a certain stage, such as capabilities description upon conference creation, reservation, activation, etc., which a Conferencing System maintains in order to describe the system capabilities and to provide access to the available services provided by the Conferencing System.

Consequently, the XCON specifications don’t mandate the actual usage of the Conference Object, but rather defines the general cloning tree concept and the mechanisms required for its realization, which is described in detail in Section 5.1.

Adhoc and advanced conferencing examples are provided in Section 5.2 and Section 5.3, with the latter providing additional description of the Conference Object in terms of the stages of a conference, to support scheduled and other advanced conference capabilities.

The scheduling of a conference based on these concepts and mechanisms
is then detailed in Section 5.4

5.1 Cloning Tree

The concept defined in this section is a logical representation only, as it is reflected through the XCON mechanisms: the URIs and the protocols. Of course, the actual system realization can differ from the presented model and doesn’t require physical copying of the information, etc.

Any Conference Object in a Conferencing System is created by either being explicitly cloned from an existing parent object or being implicitly cloned from a default system object. This document uses the "cloning" metaphor instead of the "inheritance" metaphor because it more closely fits the object replication and extension concept, rather than data type re-usage and extension concept.

By default, all the data existing in the parent object is copied to the newly created object.

The cloning operation needs to specify whether the link between the parent and the child needs to be maintained in the system or not. If no link between the parent and the child exists, the objects become independent and the rest of the cloning process doesn’t apply.

Once the new object is created, it can be addressed by a unique Conference Object URI assigned by the system. The Conference Object URI is defined by [TBD].

By default, the newly created object can expand the data it contains, within the schema types supported by the parent, just as an independent object can. It can also restrict the read/write access to its objects, but unlike an independent object, cannot relax it relative to its parents access.

Any piece of data in the child object can be independently accessed and, by default, can be independently modified without affecting the parent data.

On the other hand, the parent object can enforce a different policy by marking certain data elements as "parent enforceable". The values of these data objects cannot be changed by directly accessing the child object; neither can they be expanded in the child object alone.

If required in the future, XCON specifications may introduce additional (to the "parent enforceable") logical relationships between elements being cloned from one another.
Figure 3: The Cloning Tree.

Using the defined cloning model and its tools, the following sections show examples of how different XCON-compliant systems can be realized.
5.2 Ad-hoc Example

Figure 4 illustrates how an ad-hoc conference can be created and managed in a conferencing system.

A client can create a conference by establishing a call control signalling channel with a Conference Factory as specified in Section 2. The Conference Factory can internally select one of the system supported Conference Object blueprints based on the requesting client privileges and the media lines included in the SDP body.

The selected blueprint with its default values is copied by the server into a newly created Conference Object, referred to as an ‘Active Conference’. At this point the Conference Object becomes independent from its blueprint. A new Conference Object Identifier, a new Conference Identifier and a new focus are allocated by the server.

During the conference lifetime, an authorized client can manipulate the Conference Object (such as adding participants) using the Conference Control Protocol [Section 6.3].
Figure 4: Conference Ad-hoc Creation and Lifetime.

5.3 Advanced Example

Figure 5 illustrates how a recurring conference can be specified according to system capabilities, scheduled, reserved, and managed in a conferencing system.

Firstly, a client would query a Conferencing System for its capabilities. This can be done by requesting a list of the Conference Object "blueprints" the system supports. Each blueprint can contain a specific combination of capabilities and limitations of the conference server in terms of supported media types (audio, video, text, or combinations of these), participant roles, maximum number of participants of each role, availability of floor control, controls available for participants, availability and type of sidebars, the definitions and names of media streams, etc.

[Editor’s Note: Currently, there is an open issue around Querying a server for a particular template, per Discussion Point 3 on the mailing list.]
The selected blueprint with its default values is copied by the client into a newly created Conference Object, referred to as a 'Conference Object for Reservation’, that specifies the resources needed from the system for this Conference Instance. At this point the ‘Conference Object for Reservation’ becomes independent from its blueprint. The client can also change the default values (within the system ranges) and add additional information (such as the list of participants and the conference start time) to the ‘Conference Object for Reservation’.

At this point the client can ask the conference server to create a new conference reservation by attaching the ‘Conference Object for Reservation’ to the request. As a result, the server can allocate the needed resources, create the additional Conference Objects for each conference occurrence (referred to as a ‘Conference Object for Occurrence’) and allocate the Conference Object identifiers for all the ‘Conference Object for Reservation’ and for each ‘Conference Object for Occurrence’.

From this point on, any authorized client is able to access and modify each of the Conference Objects independently. By default, changes to an individual ‘Conference Object for Occurrence’ will affect neither the ‘Conference Object for Reservation’ nor its siblings.

On the other hand, some of the conference sub-objects, such as the maximum number of participants and the participants list, can be defined by the system as parent-forcible. As a result, these objects can be modified by accessing the ‘Conference Object for Reservation’ only. The changes to these objects can be applied automatically to each of the ‘Conference Object for Occurrence’s (subject to local policy).

```
+---+-----------------------+
 | p |                       |
 | o |   SELECTED           |
 | l |   CONFERENCE         |
 | i |   BLUEPRINT          |
 | c |                       |
 +---+-----------------------+
 |
 \|/
 /\   
 /  
/   
```
5.4 Scheduling a ‘Conference Object for Reservation’

Scheduling conferences forms an important part of the Conferencing System solution. The concept of a Conference Data Object representing an individual Conference instance is introduced in Section 4.2. A basic Conference Instance represents a single occurrence that has a specified ‘start’ and ‘end’ time. In most advanced conferencing solutions it is possible to not only schedule an individual conference instance, but also schedule a series of related conferences (e.g. A weekly meeting that occurs on Thursday at 09.00).

[Editor’s Note: Currently, there is an open issue around Referring to sets of meetings, per Discussion Point 1 on the mailing list.]

To be able to achieve such functionality, a Conferencing System needs to be able to appropriately schedule and maintain Conference Instances that form part of a recurring conference schedule. The
mechanism proposed in this document makes use of the 'Internet Calendaring and Scheduling Core Object' specification defined in RFC2445[8] in union with the concepts introduced in Section 4 for the purpose of achieving advanced conference scheduling capability.

Figure 6 illustrates a simplified view of a Client interacting with a Conference System. The Client is using the 'Conference Control Protocol'[ref - TBD] protocol to add a new Conference Instance to the Conference System by interfacing with the Conference Control Server. A Conference Control Protocol request contains a valid 'Conference Object for Reservation' and reference by value to an 'iCal' object which contains scheduling information about the conference instance (e.g. start time, end time).
Figure 6: Resource Scheduling

A successful creation of a Conference Instance using the Conference Control Protocol results in a new 'Conference Object for Reservation'. A Conference Control Protocol request is validated, including the associated iCal object, and the resultant 'Conference Object for Reservation' is created. The Conference Object is uniquely represented within the Conference System by Conference Object Identifier (e.g. xcon:hd87928374) as defined in [TBD] and discussed in Section 4.1. The unique URI is returned to the client and can be used to reference the 'Conference Object for Reservation' representation if any future manipulations are required (e.g. Alter start time) using a Conference Control Protocol.

The previous example explains how a client creates a basic 'Conference Object for Reservation' using an iCal reference in association with a Conference Control Protocol. Figure 6 can also be applied when explaining how a series of Conferences are scheduled in the system. The description is almost identical with the exception that the iCal definition that is included in a Conference Control Request represents a series of recurring Conference Instances (e.g. conference start time, end time, occur weekly). The Conference system will treat this request the same as the first example. The protocol request will be validated, along with the associated iCal object, and the 'Conference Object for Reservation' will be created. The 'Conference Object for Reservation' and the Conference Object ID created for this example represent the entire series of recurring Conference Instances rather than a single Conference. If the client uses the Conference Object ID provided and a Conference Control Protocol to adjust the 'Conference Object for Reservation', every Conference Instance in the series will be altered.

A Conferencing System that supports the scheduling of a series of Conference Instances should also be able to support manipulation within a series range. A good example might be that a 'Conference Object for Reservation' has been scheduled to occur every Monday at 09.00 GMT. For the next three weeks only, the meeting has been altered to occur at 10.00 GMT in an alternative venue. With Figure 6 in mind, the client will construct a Conference Control request whose purpose is to modify the existing 'Conference Object for Reservation' representing the recurring Conference Instance. The Client will include the Conference Object ID provided by the Conference System to explicitly reference the 'Conference Object for Reservation' within the Conference System. A Conference Control
request will contain all the required changes to the ‘Conference Object for Reservation’ (e.g., Change of venue). The Conference System matches the incoming request to the existing ‘Conference Object for Reservation’ but identifies that the associated iCal object only refers to a range of the existing series. The Conference System creates a child clone of the original ‘Conference Object for Reservation’ to represent the altered Conference Instances within the Series. The cloned ‘Conference Object for Reservation’ representing the altered series of Conference Instances has its own associated Conference Object ID which is returned to the Client using a Conference Control Protocol. This Conference Object ID is then used by the Client to make any future alterations on the newly defined sub-series. This process can be repeated any number of times as the newly returned Conference Object ID representing an altered (cloned) series of Conference Instances, can itself be manipulated using the Conference Control Protocol and the newly created Conference Object ID representation. This provides a flexible approach to the scheduling of recurring Conference instances.

6. Conferencing Mechanisms

6.1 Call Control Signalling

The focus is the central component of the conference. Participants interface with the focus using an appropriate Call Control Signalling protocol. Participants request to establish or join a conference using the Call control signalling interface. A focus then either accepts (after checking the policies), sends a progress indication (e.g., for a parked call while awaiting moderator approval to join) or rejects that request using the call control signalling interface.

During an active conference, a Conference Control Protocol [Section 6.3] can be used to affect the conference state. For example, Conference Control Protocol requests to add and delete participants will be communicated to the focus and checked against the conference policies. If approved, the participants will be added or deleted using the call signalling to/from the focus.

6.2 Notifications

A Conferencing System is responsible for implementing a Conference Notification Service. The Conference Notification Service provides updates about the Conference Instance state to authorized parties (e.g., participants) according to [11].

User identity and his/her Role are used by the conferencing system to filter the notifications such that they contain only information that is allowed to be sent to that user.
6.3 Conference Control Protocols

The XCON working group will develop a protocol or a set of protocols for controlling the state of a Conference Object and for retrieving its state. The nature of this protocol is currently under discussion in the XCON working group. The proposals span from data manipulation (management-like) protocols to semantic-oriented. Two of the proposed protocols are detailed in the sections below. Among other mentioned candidates are SOAP and HTML FORMS. The following paragraphs summarize the fundamental issues around the selection of the protocol(s). [Editor’s Note: Discussion Point 5 on the XCON WG mailing list].

It is recognized that semantic manipulations make for a cleaner protocol design, with the disadvantage that extensions to the underlying data model require extensions to the protocol used to manipulate it. Syntactic manipulations allow for extensions to the data model without requiring protocol extensions, with the disadvantage that the server generally has to infer intent from data manipulations instead of having intent explicitly signaled.

It is worth noting that one portion of the data to be manipulated, the Common Conference Information, will not be extended, and would naturally lend itself to semantic manipulation. Another part of the data, the Conference Template, is intended to be extended, and would naturally lend itself to syntactic manipulation. However, there has been a stated desire to use a single protocol (and presumably a single mode of operation within this protocol) to manipulate all conference object state (common and template).

The third statement in the paragraph above makes the first two solution options mutually exclusive; the XCON working group needs to determine which of the three statements above is least important to the working group.

6.3.1 CPCP

A Conference Policy Control Protocol (CPCP) is a data manipulation protocol being proposed as a standard means of storing and manipulating the conference policy. According to CPCP, the conference policy is comprised of the rules associated with a specific Conference Instance. Requirements for the CPCP are defined in the CPCP Requirements document [13]. The Conference Policy Control Protocol document [17] defines the XML Schema for the conference policy data elements.

The privileges as to which users are allowed to read and/or manipulate a specific Conference Instance are defined in a separate
Extensible Markup Language (XML) Schema[19]. This schema is based on the common policy model being used for geographic privacy information and for presence information.

A separate document [18] proposes XCAP as one protocol mechanism for storing and manipulating this conferencing policy data. XCAP is a HTTP 1.1 based protocol that allows clients to read, write, modify and delete application data stored in XML format on a server. One of the main advantages of XCAP is that it maps XML document elements and attributes to HTTP URIs that can be directly accessed by HTTP.

6.3.2 CCCP

A Centralized Conferencing Control Protocol [20] is a semantic-oriented protocol defined to allow participants or otherwise authorized entities to directly manipulate an active Conference Instance. CCCP is defined as a set of transactions issued over a reliable transport protocol. A transaction consists of a Request carrying the required information in an XML body and a corresponding Response carrying an appropriate XML body.

CCCP requests are submitted to the CCCP server and can be prioritized and queued, based on the CCCP client Role and the requested primitives. CCCP requires no single lock per document, and the CCCP server can locally implement an optimization strategy independent of CCCP client behavior.

6.4 Floor Control

[Editor’s Note: This text still requires further updating to reflect new model and pending new WG agreements. Amongst the things TODO are:

1. Need to be able to fetch from the Conference Object the credentials required for BFCP. This includes the conference id, user id, and password.

2. Clarification of identifiers based on agreement with content in section Section 6.4.1, etc.

3 Trimming the level of detail in this section since some of it is beyond the scope of the framework – it’s here for now to establish the context for discussion. New draft TBD if mechanism details need to be specified.

4. Evaluation of an alternative proposal [TBD as a stand alone draft as well] for using Templates as the means for correlating Floor Control identifiers.]
A mechanism for floor control within a Conferencing System is defined in the ‘Binary Floor Control Protocol’ specification [16]. Floor control is not a mandatory mechanism for a Conferencing System implementation but provides advanced media input control features for participants.

An XCON compliant client supporting floor control needs to obtain information for connecting to a floor control server to enable it to issue floor requests. This connection information can be retrieved using information provided by mechanisms such as negotiation using the SDP[2] offer/answer[5] exchange on the signalling interface with the focus.

As well as the Client --> Floor Server connection information, a client wishing to interact with a Floor Control server requires access to additional information. This information associates Floor Control interactions with the appropriate Floor instance. Once a connection has been established and authenticated (see [16] for authentication details), a specific floor control message requires detailed information to uniquely identify a user, a floor and a conference. This information, defined in the next set of bullet points, can be obtained using methods such as negotiation using the SDP offer/answer exchange on the signalling interface with the focus:

- Conference Object ID - Each Conference Object has a unique identifier per Section 4.1, obtained from the Conference server, which MUST be included in all Floor messages. When using SDP offer/answer exchange to negotiate a Floor control connection with the focus using the call control signalling interface, the unique conference identifier is contained in the ‘confid’ SDP attribute, as defined in [24] - e.g. a=confid:2874.

- User Identifier - Each user has a unique identifier within the Conference Object per Section 6.4.1, obtained from the Conference server, which MUST be included in all Floor messages. When using SDP offer/answer exchange to negotiate a Floor control connection with the focus using the call control signalling interface, the unique conference identifier is contained in the ‘userid’ SDP attribute, as defined in [24] - e.g. a=userid:8937.

- Floor ID - A media session with a Conferencing System can have any number of ‘Floors’ (0 or more) that are represented by a unique identifier within the Conference Instance (as represented by Conference ID). When using SDP offer/answer exchange to negotiate a Floor control connection with the focus using the call control signalling interface, the unique conference identifier is contained in the ‘floorid’ SDP attribute, as defined in [24] e.g.a=floorid:1 m-stream:10 . Each ‘floorid’ attribute,
representing a unique floor, has an ‘m-stream’ tag containing one or more identifiers. The identifiers represent individual SDP media sessions (as defined using ‘m=’ from SDP) using the SDP ‘Label’ attribute as defined in [23].

Clients can authenticate a BFCP server using the TLS certificates. Requests submitted on a successfully created BFCP connection may additionally require digest authentication within the BFCP protocol to authenticate the Floor control client to the server. For this to take place a shared secret needs to be exchanged between the BFCP client/server entities. This can be achieved out of band using a mechanism such as the ‘k=’ SDP attribute. The shared secret can also be exchanged using un-specified ‘out of band’ mechanisms. When using Digest authentication of BFCP client messages the exchange of an active ‘Nonce’ is also required. This can be achieved using a BFCP request response interaction as defined in BFCP (A request is submitted without aNonce TLV and the server generates an error response with either an Error Code 7 (DIGEST TLV Required) or 6 (Invalid Nonce) containing the valid nonce). The BFCP ‘Nonce’ value can also be obtained ‘out of band’ using information provided in the Offer/Answer exchange. As with the other SDP Floor attributes referenced in this section and defined in [24], the ‘nonce’ attribute can be inserted in the SIP response e.g. a=nonce:dhlsa8hd0dwqj.

Editors Note: Should we be referencing SDP-new? The authors believe so, but the SDP BFCP draft does NOT, so we need to agree and align all the documents to be consistent.

6.4.1 User Identifier

[Editor’s Note: This section is included to capture the information for discussion. It’s likely this topic should be covered in a separate document once there’s agreement on the concept. It will need to include more details such as use of User ID as ‘username’ challenge in SIP Digest etc.]

An important concept when considering a Conference System is the identity of users. This document introduces the concept of a Conferencing System User Identifier, as defined in [*TBD - may fit with new Conference Object ID concept doc]. A Conference System user has an associated unique identifier (in the context of the Conference System) that forms a generic internal representation.

It is intended that the generic User Identity (User ID) for a Conference system can be used in all relevant supporting protocols (e.g. Floor Control, Media Policy, CCP, signaling interface). The mapping of User ID from various XCON protocols should be discussed and referenced in relevant protocol documents. Section 6.4 provides
more details regarding the mapping of User ID with floor control.

It is important to emphasize that the User Identity being described is in the context of the Conference System. This is due to the requirement for identity of Conference System users who may not be participating in a Conference Instance. Examples being a non participating 'Floor Control Chair' or 'Media Policy' Controller. Users can then be associated with Conference Objects as defined in Section 4.2. This association enables the Conference System to associate and enforce user level policies at both a system level and Conference Object level.

7. Conferencing Scenario Realizations

This section addresses how advanced conferencing scenarios, many of which have been described in [14], are realized using this XCON framework. The objective of this section is to further illustrate the model, mechanisms and protocols presented in the previous sections and also serves to validate that the model, mechanisms and protocols are sufficient to support advanced conferencing scenarios.

Section 5.2 and Section 5.3 provide examples of the creation of a basic adhoc conference and a more advanced scheduled conference.

[Editor’s note: This whole section needs additional work based upon the new agreed model, new terminology and needs to reflect the rework done in the other WG documents. In the end, some XML snippets would also likely be useful. Section 7.1 is a 1st pass at a detailed diagram. WG feedback on this approach (versus traditional ladder diagrams) is appreciated.]

7.1 Participant Manipulations

There are different ways to affect a participant state in a conference.

A participant can join and leave the conference using call signalling means only, such as SIP. A participant can change its own media streams by, for example, sending re-INVITE with new SDP content using SIP only. This kind of operations is called "1st party signalling" and they do not affect the state of other participants in the conference.

Limited operations for controlling other conference participants (a so called "3rd party control") through the Focus using call signalling only may also be available for some signalling protocols. For example, "Conferencing for SIP User Agents" [15] shows how SIP with REFER can be used to achieve this functionality.
In order to perform richer conference control a user client needs to implement a conference control protocol client. By using a conference control protocol, the client can affect its own state, state of other participants, and state of various resources (such as media mixers) which may indirectly affect the state of any of the conference participants.

Figure 7 provides an example of one client "Alice" impacting the state of another client "Bob". This example assumes an established conference. In this example, the Client, "Alice" whose Role is "moderator" of the conference, wants to mute "Bob" on a medium-size multi-party conference, as his device is not muted (and he's obviously not listening to the call) and background noise in his office environment is disruptive to the conference.

It should be noted that only entities impacted by the request are shown in the diagram.

Figure 7: Client Manipulation of Conference - Mute a party

Upon receipt of the Conference Control Protocol request to "mute" a party ("Bob") in the specific conference as identified by the
Conference Object ID, the Conference Server ensures that "Alice" has the appropriate authority based on the policies associated with that specific Conference Object to perform the operation. "Bob's" status is marked as "recvonly" and the Conference Template of the Conference Object (if included) is updated to reflect that "Bob's" media is not to be "mixed" with the conference media.

Depending upon the policies, other participants (including "Bob") may be notified of this change via the Notification Service.

7.2 Media Manipulations

7.3 Sidebar Manipulations

[Editor's Note: TBD. Once we get full agreement on terminology and the basic ideas in the other sections, we'll tackle this. Sidebars are in many ways normal conferences but additional authorization aspects need to be taken into consideration. Such as that only members of the parent conference may be allowed to interact with the sidebar.]

7.4 Whispering or Private Messages

[Editor's Note: TBD. Once we get full agreement on terminology and the basic ideas in the other sections, we'll tackle this.]

7.5 Conference Announcements and Recordings

[Editor's note: TBD. Use Cullen’s comments on the previous version of the doc.]

8. Relationships between SIPPING and XCON Frameworks

The SIPPING Conferencing Framework [9] provides an overview of a wide range of centralized conferencing solutions known today in the conferencing industry. The document introduces a terminology and logical entities in order to systemize the overview and to show the common core of many of these systems. The logical entities and the listed scenarios in the SIPPING Conferencing Framework are being used to illustrate how SIP [4] can be used as a signalling means in these conferencing systems. SIPPING Conferencing Framework does not define new conference control protocols to be used by the general conferencing system. It uses only basic SIP [4], the SIP Conferencing for User Agents [15], and the SIPPING Conference Package [9] for basic SIP conferencing realization.

The XCON framework defines a particular centralized Conferencing
System and the logical entities implementing it. It also defines a particular XCON Data Model and refers to the standard protocols (beyond call signalling means) being defined by the XCON WG to be used among the XCON logical entities for implementing advanced conferencing features. The purpose of XCON working group and this framework is to achieve interoperability between the XCON entities from different vendors for controlling different aspects of advanced conferencing applications.

The logical entities defined in the two frameworks are not intended to be mapped one-to-one. The two frameworks differ in the interpretation of the internal conferencing system decomposition and the corresponding operations. Nevertheless, the basic SIP [4], the SIP Conferencing for User Agents [15], and the SIPPING Conference Package [9] are fully compatible with both Framework documents.

9. Security Considerations

There are a wide variety of potential attacks related to conferencing, due to the natural involvement of multiple endpoints and the many, often user-invoked, capabilities provided by the conferencing system. Examples of attacks include the following: an endpoint attempting to listen to conferences in which it is not authorized to participate, an endpoint attempting to disconnect or mute other users, and theft of service, by an endpoint, in attempting to create conferences it is not allowed to create.

There are several issues surrounding security of this conferencing framework. One set of issues involves securing the actual protocols and the associated authorization mechanisms. This first set of issues should be addressed in the specifications specific to the protocols described in Section 6. The protocols used for manipulation and retrieval of confidential information MUST support a confidentiality and integrity mechanism. Similar requirements apply for the floor control protocols.

There are also security issues associated with the authorization to perform actions on the conferencing system to invoke specific capabilities. Section 4.5 discusses the policies associated with the Conference Object to ensure that only authorized entities are able to manipulate the data to access the capabilities. The final set of issues involves the privacy and security of the identity of a user in the conference.

Many policy authorization decisions are based on the identity of the user or the Role that a user may have. There are several ways that a user might authenticate its identity to the system. The conferencing system may know about specific users and assign passwords to the
The users may also be authenticated by knowing a particular Conference ID and a PIN for it. Sometimes, a PIN is not required and the Conference ID is used as a shared secret. The call signalling means can provide a trusted form of the user identity or it might just provide a hint of the possible identity and the user still needs to provide some authentication to prove it has the identity that was provided as a hint in the call signalling. This may be in the form of an IVR system or other means. The goal for the Conferencing System is to figure out a user identity and a Role for any attempt to send a request to the Conferencing System.

When a Conferencing System presents the identity of authorized users, it may choose to provide information about the way the identity was proven to or verified by the system. A user may also come as a completely unauthenticated user into the system - this fact needs also be communicated to interested parties.

When guest users interact with the system, it is often in the context of a particular conference. In this case the user may provide a PIN or a password that is specific to the conferences and authenticates the user to take on a certain role in that conference. The guest user can then perform actions that are allowed to any user with that Role.

The term password is used to mean the usual, that is to say a reasonable sized, in number of bits, hard to predict shared secret. Today users often have passwords with more than 30 bits of randomness in them. PIN is a special password case - a shared secret that is only numeric and often contains a fairly small number of bits (often as few as 10 bits). When Conferencing Systems are used for audio on the PSTN, there is often a need to authenticate using a PIN. Typically if the user fails to provide the correct PIN a few times in a row, the PSTN call is disconnected. The rate of making the calls and getting to the point to enter a PIN makes it fairly hard to do an exhaustive search of the PIN space even for 4 digit PINs. When using a high speed interface to connect to a Conferencing System, it is often possible to do thousands of attempts per second and the PIN space could quickly be searched. Because of this, it is not appropriate to use PINs for authorization on any of the interfaces that provide fast queries or many simultaneous queries.

This conferencing system has an idea of the identity of a user but this does not mean it can reveal this identity to other users, due to privacy considerations. Users can set select various options for revealing their identity to other users. A user can be "hidden" such that other users can not see they are involved in the conference, or they can be "anonymous" such that users can see that another user is there, but not see the identity of the user, or they can be "public"
where other users can see their identity. If there are multiple "anonymous" users, other parties will be able to see them as independent "anonymous" parties and will be able to tell how many "anonymous" parties are in the conference. Note, that the visibility to other participants is dependent on their Roles. For example, users' visibility (including "anonymous" and "hidden") may be displayed to the moderator or administrator, subject to a Conferencing System's local policies. "Hidden" status is often used by robot participants of a conference and is also used in many call center conference situations.

10. IANA Considerations

This is an informational draft, with no IANA considerations required.

11. Acknowledgements

This document is a result of architectural discussions among IETF XCON working group participants. The authors would like to thank Henning Schulzrinne for the "Conference Object Tree" proposal and Cullen Jennings for providing input for the "Security Considerations" section.

12. References

12.1 Normative References


12.2 Informative References


Authors’ Addresses

Mary Barnes
Nortel
2380 Performance Drive
Richardson, TX

Email: mary.barnes@nortel.com
Chris Boulton
Ubiquity Software Corporation
Building 3
Wern Fawr Lane
St Mellons
Cardiff, South Wales  CF3 5EA

Email: cboulton@ubiquitysoftware.com

Orit Levin
Microsoft Corporation
One Microsoft Way
Redmond, WA  98052

Email: oritl@microsoft.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.