SIP Conferencing

Jörg Ott <jo@tzi.org

17 January 2003

Upperside SIP 2003 - Paris

IETF Conferencing

- Packet multimedia experiments since the 1980s
  - Audio/video tools + protocols for A/V over IP
  - Conference announcement and control protocols
- First IETF Audiocast (1992)
  - Mbone-based audio transmission from selected IETF working groups
- Since then: IETF sessions on the Mbone
  - Audio + video (+ sometimes slides)
  - Enabling remote participation (even talks)
- Other uses of Mbone conferencing
  - Broadcasting NASA missions, concerts, ...
  - Lectures, seminars, project meetings, ...
Traditional IETF Conferencing Concept

- Multicast-based
- Loosely-coupled conferences
  - no membership control
  - inexact information about participants
    - provided on a voluntary basis
  - security by encryption
- Public announcements and invitations
  - Convey session parameters, then get out of the way
    - Session Announcement Protocol (SAP)
    - Session Initiation Protocol (SIP)
- Conference control
  - Some need perceived; several attempts; insufficient real interest

Some Historic (IETF) Conferencing Protocols...

- MMCC / CCP (early 1990s)
  - Control for a distributed packet-based conferencing system
  - State synchronization in multicast groups
  - Communicating state and events between local and remote entities
  - Support for modular systems (media engines, controllers, UIs, …)
- SCCP (1996; revision 2001)
  - Distribution of conference + media session state
  - Makes use of reliable multicast
- Some other protocols…
Some Historic (IETF) Conferencing Protocols…

- MMCC / CCP (early 1990s)
  - Control for a distributed packet-based conferencing system
  - State synchronization in multicast groups
  - Communicating state and events between local and remote entities
  - Support for modular systems (media engines, controllers, UIs, …)
- SCCP (1996; revision 2001)
  - Distribution of conference + media session state
  - Makes use of reliable multicast
- Some other protocols…

All protocols were based on IP Multicast!

SIP and Conferencing over Time…

- Origin: MMUSIC: Multiparty Multimedia Session Control
- From Invitation… to initiation, modification, and termination
- From Multiparty… to point-to-point-focused
- From Multimedia… to voice-centric

Now: Multiparty & multimedia rediscovered

But: Don’t believe in multicast (anymore)!

- SIP signaling relationships
  - Centralized (bridge, endpoint) vs. mesh of SIP dialogs
- Media distribution
  - Unicast vs. multicast
- Media mixing
  - Centralized (bridge, endpoint) vs. decentralized
- Conference creation
  - ad-hoc vs. scheduled
  - “dial-in” vs. “dial-out”
- How much functionality to provide in SIP?
  - Join / leave
  - Membership management? ...?
Current SIP Conferencing Framework

- **Motivators:** n-way calling, video conferencing, …
  - Tightly coupled conferences
- **Recent conferencing models preserved**
  - Except for “fully meshed” conference: complexity just not worth it!
- **Terminology**
  - Trying to avoid already overloaded terms as much as possible
- **Functional entities in a “system model”**
  - Plus implementations examples
- **Set of protocols**
  - Definition of basic building blocks (SIP and other)
  - Sample combinations to implement conferencing services
Conference Participants

- MUST work with basic SIP support only
  - No awareness of conference
  - Just a point-to-point call with minimal means for control

- Member types (SIP UA)
  - Basic participant: plain-old SIP device
  - Complex participant: supports conferencing features
  - Focus: (one) center of a conference
  - Anonymous: Visible but unidentified participant
  - Invisible: Participant whose presence is not known

Conferences

- Conference types
  - Basic: just plain SIP, no further means for control
  - Complex: some conferencing features provided
  - Cascaded: several foci concatenated in a conference
  - Sidebar: conference as (logical) part of another

- Focus
  - Signaling center of a conference

- Conference URI
  - Identifies a focus
  - (isFocus parameter may indicate this)

- Factory URI
  - for automated conference creation
  - Yields a dynamically generated conference URI in return
Conferencing Scenarios (1)

- Simple conferencing scenarios
  - Plain SIP only (RFC 3261, 3264)

- Extend point-to-point call
  - Works only with local focus; otherwise, new call required

- Ad-hoc conference
  - Automated creation at focus
  - IVR / DTMF for control
  - Audio for information about the conference and its members

- Reserved conference
  - Same as ad-hoc
  - Use external means for reservation and configuration (e.g. web)
Conferencing Scenarios (2)

- Advanced conferencing scenarios:
  - Support for Call Transfer
  - Support means to communicate information from focus to UA
  - Optional: means to manipulate conference and media policy

- Extend point-to-point call
- Join / create a conference based upon an existing dialog
- Ad-hoc conference
- Reserved conference

- Make use of additional conferencing features

Sample Conferencing Features

- Invite participants (dial-in, dial-out), expel participants
- Authenticate new participants by members
- Obtain conference and media policy information
- Manipulate conference policy
  - Participant privileges, participant management (black list, white list, …)
  - Floor control,
- Explicit media control (media policy)
  - Configure media distribution
  - Add / remove media sessions
- Create, control, and terminate sidebars
  - Separate conference vs. media policy
- …
Conference Policy

Define, retrieve/notify, modify, and act upon...

- Formal rules for the conference
  - Conference creation, termination, (policy) modification
  - Access control: black list, white list, rules for authentication
  - Privileges of individual participants
  - Visibility of the conference and its members
  - Access to floor and media policy (defined separately)

- General conference attributes
- Participant management
  - Invite, expel
- ...

Media Policy

- Mixer model
  - Input switch: collecting & selecting input streams from participants
    - Possibly transcoding and other per-media functions
  - Mixing topologies describing mixing policies
  - Output switch: selecting & distributing output streams to participants
    - Possibly transcoding and other per-media functions
- Mixer may be centralized or not

- Media policy defines how incoming streams are processed, combined, and then distributed
  - Individual mixing functions may be defined per participant
  - Common mixing functions may be defined for the conference
  - Mixing function may take into account “events” from other components
SIP Signaling Building Blocks

- **Membership control**
  - Initiation of conferences: INVITE
  - Inviting / adding to conferences: INVITE, REFER
  - Leaving a conference: BYE
  - Expelling from a conference: REFER (method="BYE")

- **Conference control**
  - State change notifications: SUBSCRIBE / NOTIFY
    - Dialog package, conference package, ...
  - Conference / media policy control
    - Might use data manipulation framework discussed in SIMPLE context

- **Other**
  - Determine focus URI OPTIONS
  - ...

---

Signaling Building Blocks using other Protocols

- Conference Policy Control Protocol (CPCP)
- Media Policy Control Protocol (MPCP)
- Floor control protocol

- **Common baseline:**
  - Request-response protocol (RPC-style)
  - May be strictly client – server
  - Asynchronous notifications provided by SIP
  - HTTP/HTML access to a web page for human interaction
  - SOAP RPCs e.g. over HTTP
**Call Flow Example: Conference Creation**

Alice  
INVITE sip:conf-factory  
180 Ringing  
200 OK Contact:conf-id  
ACK  
Media sessions  
SUBSCRIBE sip:conf-id  
200 OK  
NOTIFY  
200 OK

Focus  
Bob

Carol

**Call Flow Example: User Joining**

Alice  
Media sessions  
NOTIFY  
200 OK  
Media sessions

Focus  
INVITE sip:conf-id  
180 Ringing  
200 OK Contact:conf-id  
ACK

Bob  
Media sessions

Carol  
Media sessions  
SUBSCRIBE sip:conf-id  
200 OK  
NOTIFY  
200 OK
Call Flow Example: Adding a User

1. **INVITE** Contact:conf-id
   - Bob

2. **Ringing**
   - Carol

3. **200 Accepted**
   - Bob

4. **NOTIFY**
   - Bob

5. **ACK**
   - Bob

6. **REFER sip:conf-id**
   - Refer-To: carol
   - Focus

7. **202 Accepted**
   - Focus

8. **NOTIFY**
   - Focus

9. **200 OK**
   - Focus

10. **SUBSCRIBE sip:conf-id 200 OK**
    - Media sessions

11. **NOTIFY 200 OK**
    - Media sessions

12. **NOTIFY 200 OK**
    - Media sessions

Call Flow Example: Removing a User

1. **BYE sip:carol**
   - Carol

2. **200 OK**
   - Media sessions

3. **REFER sip:conf-id Refer-To: carol?m=BYE**
   - Focus

4. **202 Accepted**
   - Focus

5. **NOTIFY 200 OK**
   - Focus

6. **200 OK**
   - Focus

7. **NOTIFY 200 OK**
   - Media sessions

8. **NOTIFY 200 OK**
   - Media sessions

9. **NOTIFY 200 OK**
   - Media sessions
Call Flow Example: User Leaving

Alice — Focus — Bob — Carol

Media sessions — Media sessions — Media sessions — Media sessions

BYE sip:conf-id

200 OK

ACK

Terminate conference
e.g. by non-SIP means

A Bit of Status

- Framework: done
- Scenarios: agreed upon
  - Issue: sidebar conferences
- Requirements: nearing completion for many areas
- SIP Support: largely “complete”
  - Still some minor issues
- Conference and media control: moving ahead
  - Requirements well advanced; protocol ideas progressing
- Media control (e.g. “FIR”, Picture Freeze, …)?
- …
Conclusion

- After ~10 years of consideration, IETF conferencing seems to finally make it.
- Shift towards a more realistic target
- SIP as strong, commercially accepted base protocol
- Further control protocols developed at aggressive schedule
- Looking for further standardized applications

- Work in Progress
- Further info:
  http://www.softarmor.com/sipping/teams/conf/