Multiparty Voice over IP (MVoIP) Peer-based System for Large-scale Conference Support

Wajdi ELLEUCH and Alain C. HOULE

Abstract— Even if both traditional centralized and decentralized models can support small conferences, their deployment for large number of participants is less obvious. In fact, while server-based centralized conference facilitates control and administration operations, supporting centralized media processing in large scale conference causes system overhead on the mixer. On the other hand, decentralized solutions that use multi end-system media processors will introduce an overhead to control conference floors and users membership. Our solution introduces a new model that enable multi-host media process support while the conference control and management is kept simplified and centralized around the administrator. To do that, we build two different meshed networks to enable both voice audio distribution between participants and general conference control. In this short paper we introduce our system topology and indentify components that enable conference creation/destruction, user addition/removal, media assignment and speech floor control.

***Index Terms*—VoIP, MVoIP, streaming applications, multi-party communication, conference models, tree-based distribution, conference control, floor control, media distribution, SIP protocol**

I. INTRODUCTION

In contrast to network and server infrastructure-based stream content delivery networks, Application Layer Multicast (ALM) end-point based solution has recently received attention [1]. These ALM solutions tend to produce self-organizing, efficient and self-improving overlay meshes that can be dynamically adapted to different network variations. In these systems, a participant node can decide about its parents and migrate within the constructed tree to minimize the redundant transmission on physical links. Unfortunately, we found these solutions are mostly focused on the network resources aspect instead of user preferences and activity. For example, active participants that talk on MVoIP (Multiparty Voice over IP) conference should be handled distinctively from a participant that is just in listen mode. Also, implementing complex control functionalities on all end-host might restrict access for handheld user devices with limited resources. Moreover, designer of large scale MVoIP based on media processing peers should consider different aspects related to conference creation, user membership management, control media distribution and speech floor control for each participant. Implementing these control functionalities on all peers, in a distributed manner, is excessively complex since a global and unified view of the conference floor and unique control decisions about user's membership and media distribution are required. Maintaining an updated version of such information, and sharing it between peers, for large scale conferences is also an excessive overhead.

To resolve these challenges, we introduce an innovative peer-based system that supports large scale conferences by making separation between media network and control network. While media processing uses dynamic tree-shaped decentralized approach, the control network is based on a topology centralized around the conference administrator. The proposed system enables conference creation/destruction and user addition/removal. Moreover, each user that joins a conference can offer its media processing service to the administrator that controls media distribution using the Third Party Call Control mechanism by remotely establishing, updating and removing media sessions between participants. This system also enables speech floor control to manage participant talking privileges and consequently adapts the media mixing/distribution network topology. Users can then migrate within the media tree depending on their capabilities and activities. Media tree optimization and protocol implementation will not be treated in this work that focuses more on a general presentation of the system topology in section II and model components in section III.

II. SYSTEM TOPOLOGY

Our topology builds two different meshed networks that enable both voice audio distribution between participants and general conference control as shown on Figure 1. The first network ensures media delivery in tree-shaped topology. This network uses two different media roles components: the Mixer/Distributor Node (MDN) and the Leaf Node (LN). The MDN role is affected to nodes that maintain more than one audio session on the conference. Over and above processing media for themselves, MDN nodes can mix or even distribute media for others. On the other hand, LF role concerns nodes that hold only one audio session. Their departure should not

Manuscript received August 15, 2008

W.Elleuch B.Ing, M.Sc.A is a Ph.D candidate at the department of Electrical Engineering and Computer Engineering at Université de Sherbrooke, Sherbrooke (Québec) CANADA. (e-mail: Wajdi.Elleuch@USherbrooke.ca).

A.C.Houle, ing. Ph.D. is a professor in the department of Electrical Engineering and Computer Engineering at Université de Sherbrooke, Sherbrooke (Québec) CANADA. (e-mail: Alain.Houle@USherbrooke.ca).

978-0-7695-3393-3/08 $25.00 © 2008 IEEE
DOI 10.1109/WiMob.2008.122

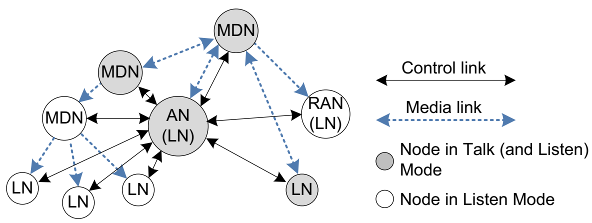

Figure 1: System control and media networks topologies

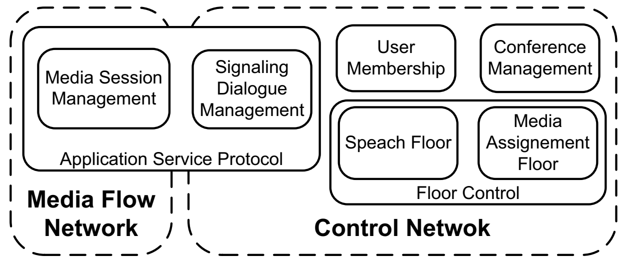

Figure 2: System components

prevent any other participant from continuing conferencing. Such role is adapted for light handheld devices with limited resources, for conventional IP phones or for participants that would not offer any part of their computational power and bandwidth to others. The second network establishes star-shaped control network constituted by connection link from each participant to the Administrator Node (AN). For each new added participant, new control link is established. These links are used to deliver control information between the AN and all remaining participants. As we do not wish to rely on a single non-failing entity to control conference, the AN should elect replica node, that we call Replica Administrator Node (RAN), among participants to replace it in cases of departure or failure.

III. System Components

A. The Conference Management

This component mainly supports conference announcement, creation, modification and destruction. Each conference is uniquely identified by an URI address created by the AN. This address is communicated to public or private users by displaying it on dedicated web pages or shared between users on discussion forum, instant messaging, e-mail or other third communication system. General parameters related to the conference title/subject, maximum number of supported participants, maximum simultaneous speaker, etc. should be defined in this component.

B. Membership Management

This component is responsible of user addition and removal operations. Requests that enable user to join or to leave can be initiated whether by the users themselves (dial-in mode) or by the administrator (dial-out mode). Before accepting a user join request or before adding new users, user membership component should comply with conference management rules and restrictions. Conference access rules (pre-authorized participants, black-list, etc) can be used to facilitate membership management.

C. Floor control

IP-based Conference may have any number of floors, depending on the features supported by the conference i.e. audio, video, white board, mouse pointed, etc. Since we limit our actual system to MVoIP service support, we focus the conference floor on the available participant shared bandwidth (media assignment floor) and the talking rights (speech floor). While the media assignment floor is controlled by the media flow distribution network, the speech floor involves speech states use combined with associated queues. A participant that joins conference is automatically enqueued on the “waiting to listen” queue. As soon as a media distributor is available, the participant will be able to listen to the conference. If the participant decides to talk, in that case, he will be automatically enqueued on the “waiting to talk” queue and affected to the first available media mixer.

D. Application session management

This component implements communication protocol that enables both voice call establishment between users and conference control/administration. Call establishment requires from protocol to support session initialization, modification and termination. On the other hand, exchanging control and administration data between participants can be achieved by a subscription/notification mechanism. Such functionalities can be supported by the Session Initiation Protocol (SIP) [2]. AN can use Third Party Call Control (3PCC) technique [3] to establish signaling dialog between two remote participants and enable media communication between them.

E. Media flow distribution Network

Since traversing multi-mixers adds packetization and play-out delay in audio stream, our media processing tree should use the minimum number of mixers. The AN should enable priority to active participant to be directly connected to the central mixer. This priority can be implemented on the speech floor control queues. When joining a conference, nodes should notify the AN about their offer of service by supplying their *input-degree* and *output-degree* values. The central mixer is selected by the AN, among the set of participants, according mainly to its offered degrees. Since the number of participant and their status (listen or talk) can dynamically change, the media tree dynamically grows or shrinks to adapt the audio processing load change over time.

IV. Bibliography

[1] Xiaohui, G., Zhen, W., Yu, P.S, Zon-Yin, S., "peerTalk: A Peer-to-Peer Multiparty Voice-over-IP System" Parallel and Distributed Systems, IEEE Transactions, vol.19, no.4, pp.515-528, April 2008.

[2] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Sparks, R.,Handley, M., Schooler, E., “SIP: Session Initiation Protocol”, IETF RFC 3261, June 2002.

[3] Rosenberg, J., Peterson, J., Schulzrinne, H., Camarillo, G., "Best Current Practices for Third Party Call Control (3pcc) in the Session Initiation Protocol (SIP)", IETF RFC 3725, April 2004.