

中継機能とミキシング機能の分散構成による 適合型会議通話システム

橋本 浩二^{1,a)} 白崎 博生² 柴田 義孝¹

受付日 2014年8月12日, 採録日 2015年1月7日

概要: 複数の遠隔地をつなぎ、映像と音声を用いて双方向通信をするための会議通話システムは様々な用途で用いられていが、各種通信イベントの実施を支援するためには、通信イベントの主旨に応じて中継機能とミキシング機能を汎用の機材で一時的に適合し構成できる会議通話システムが有用である。本論文では、従来のビデオ会議システムにおけるサーバシステムの機能をネットワーク上で分散し、通信イベントの実施において利用可能な端末とその数に応じた機能の構成を可能とする会議通話システムを提案する。提案システムでは、通信イベントを支援するために必要となる中継機能とミキシング機能を必要に応じてバックエンドで分散構成する。中継機能のみを必要とする会議通話と、ミキシング機能を必要とする会議通話の両方に対応した会議サーバをバックエンドに複数台用意することで、複数会議通話の同時実施に対応するとともに、会議サーバから利用可能なミキシング用の端末を用意することで、利用者端末ごとのミキシング機能へも対応する。汎用 PC を用いて実装したプロトタイプシステムによる評価実験では、HD 品質の会議通話を実施し、複数の会議サーバによる機能の分散が可能であることと、3つの構成（中継機能のみ利用/共用ミキサ利用/専用ミキサ利用）を利用した会議通話が可能であることを示す。特に、利用者端末ごとに専用ミキサを用いる会議通話では、バックエンドにミキシング用の端末を複数台配置することで台数効果が得られることを示す。本論文では、通信イベントの主旨に応じてバックエンドシステムとして稼働させる端末の構成を適合させ、利用者端末における会議通話の高機能化を図る分散システムの設計と実装およびその評価について報告する。

キーワード：会議通話システム, 分散構成, AV ミキサー

Adaptable Conference Communication System by Distributed Configuration of Relay and Mixing Functions

KOJI HASHIMOTO^{1,a)} HIROO SHIRASAKI² YOSHITAKA SHIBATA¹

Received: August 12, 2014, Accepted: January 7, 2015

Abstract: Recently, conference communication systems that are able to connect multiple sites using audio and video streams are widely used for a variety of purposes. However, it will be more useful conference communication system, if there is a mechanism which can be configured relay and mixing functions for the purpose of each communication event. In this paper, we propose a distributed backend system for conference communications. The system distributes each function of the server side for current video conference systems over computer networks, and can configure communication functions in response to available terminals on the backend. In our proposed system, stream relay functions and mixer functions are configured automatically in the backend system for conference communications. At first, by using multiple conference servers which implement both of relay and mixing function, the proposed system realizes functional distribution for each conference communication. Next, by using multiple shared terminals which offer mixing functions to conference servers, the whole system realizes individual mixing functions per user station. In experiments on our prototype system that was implemented using a general-purpose PCs, it was confirmed that the functional distribution method and three types of conference communication (with relay function, with shared mixer function or with private mixer function) works by using the configurable backend system. Especially, in the configuration which uses private mixers per user station, by placing shared terminals in the backend, the prototype system can increase the number of user stations joining a conference communication. The proposed system supports communication events by distributable relay and mixing functions. In this paper, we describe a high functionality distributed conference communication system, implement the prototype system and report the results of functional evaluations.

Keywords: conference communication system, distributed configuration, AV mixer

1. はじめに

情報通信技術の進歩と計算機システムの性能向上により、インターネット上で映像と音声を用いて遠隔地間とリアルタイムに双方向通信を行うことは日常的に可能となった。近年では、スマートフォンやタブレットでも利用できるビデオチャットシステムや、汎用 PC を用いた Web 会議システムが普及の一途をたどっている。インフォーマルかつ簡単なコミュニケーション手段としては、スマートフォンやタブレットを利用したビデオチャットが適しており、日常業務において電子的なデータを共有しつつ遠隔地と協調作業するのであれば、汎用 PC を用いた Web 会議システムを利用できる。一方、専用端末を用いたビデオ会議システムは HD (High Definition) 品質の高解像度な映像を扱えるようになり、企業における複数拠点をつないだ定例会議等の用途で利用されている。また、遠隔地にいる相手と対面しているような臨場感を生み出すテレプレゼンスシステム [1] といった高機能なビデオ会議システムも実現されている。様々な用途に応じて、我々は映像と音声を用いた通信システムを選択して利用できるようになった。

一方、たとえば年に数回場所を変えて不定期に行われる各種の催し事 (シンポジウム/ワークショップ/セミナー等) を、複数の地点間をつないで実施する場合には、通信システムの構築および利用する機能の構成と拡張性に関して検討すべき課題がある。ここで、複数の遠隔地間で行われる映像と音声を用いた双方向通信のことを会議通話と呼び、会議通話を行う各種の催し事を通信イベントと呼ぶことにする。最初に、通信イベントを支援する観点から、会議通話システムへの要件をまとめる。

まず、通信イベントを実施するためには、会議通話用の機材を一時的に揃えたり、専用回線や商用回線を組み合わせた通信イベント用のネットワーク環境を一時的に準備したりする必要がある。しかし、たとえばビデオ会議システムの専用端末を普段利用していても、利用上または物理的な制約により、設置場所から移動できなかつたり、通信イベント会場となる複数の地点すべてでは揃えられなかつたりすることがある。また、インターネット上で Web 会議システムを普段利用していても、インターネットとは物理的または論理的に分離されたネットワークを通信イベントで利用する場合には、Web 会議システムを利用できないことがある。実際、日常的に利用している会議通話システムを通信イベント用途には適用できない状況がある。

また、通信イベント用のネットワーク環境における可用

帯域幅が地点ごとに異なることを想定して、利用する機能を構成する必要がある。近年の映像符号化技術の進歩により HD 品質の映像と音声も、利用する符号化方式の種類と設定次第では 1 本につき数 Mbps で転送可能となったが、要求される品質に応じて必要帯域幅は数 Mbps～数十 Mbps の範囲となる。高品質な映像と音声を用いて地点ごとに複数本の映像と音声を送受信する場合、各地点で帯域幅を十分に確保できるとは限らない。したがって、多くの可用帯域幅を確保できる会場へ各地点からの映像と音声を集め、ミキシングした映像と音声を各地点へ中継する会議通話の形態が適している。その中継機能とミキシング機能をネットワーク環境に応じて適応的に構成できる仕組みが必要とされる。

さらに、同時進行で行われる複数会議通話への対応や、地点ごとのミキシング機能への対応が必要となる。たとえば、複数のセッションを同時進行で行う各種のシンポジウムや、複数のプログラムを同時進行させるワークショップやセミナーの実施を支援するためには、複数の会議通話を同時に実施できる仕組みが必要である。また、各地点において表示される映像の配置位置や再生される音声のボリュームを地点ごとに調整するためには、地点ごとに独立したミキシング機能が必要となる。これらの機能は、通信イベントの主旨や通信イベント用のネットワーク環境に応じて拡張的に利用できる仕組みが有効と考えられる。

以上をふまえ、通信イベントを支援する会議通話システムへの要件を次のとおりまとめる。

- 要件 1: 制約のある通信イベント用のネットワーク上で会議通話システムを一時的に構築できること。
- 要件 2: ネットワーク環境に応じて映像と音声の中継機能とミキシング機能を適応的に構成できること。
- 要件 3: 同時に実施可能な会議通話数を増やせる機能と、地点ごとのミキシング機能を、必要に応じて拡張的に利用できること。

通信イベントの支援を目的とし、これらの要件を満たす会議通話システムを、普段利用している機材を用いて準備するのは容易ではない。そこで本論文では、通信イベントの主旨に応じて様々な構成を可能とする分散型の会議通話システムを、汎用 PC 上で実現する。以前よりデジタルビデオの伝送 [2] や、HD 品質の映像伝送 [3] が汎用 PC 上で実現されているが、端末の性能向上により、携帯可能な小型の端末でも HD 品質の映像と音声をリアルタイムに圧縮して送信したり、受信して再生表示したりすることが可能となった。しかし、高解像度な映像のミキシング処理に加えて圧縮処理には多くの計算機資源が必要 [4] となるので、通信イベントを支援するうえでは、一時的に構築可能で台数効果のあるバックエンドシステムが有効であると考えられる。

以下、2 章では既存システムと関連研究について述べ、提案システムを位置づける。3 章では提案する会議通話用

¹ 岩手県立大学
Faculty of Software and Information Science, Iwate Prefectural University, Takizawa, Iwate 020-0693, Japan

² 株式会社ストラトスフィア
Stratosphere Inc., Chiyoda, Tokyo 102-0071, Japan

a) hashii@iwate-pu.ac.jp

バックエンドシステムを複数の端末を用いて構成する際の機能分散と、機能分散を考慮したプロトコルについて述べる。続く4章では、会議通話において映像と音声を複数地点間で双方向通信するためのモデルを示す。5章ではプロトタイプシステムの実装について述べ、6章でその有用性を評価し、7章で本論文をまとめる。

2. 既存システム・関連研究と提案システム

上記の要件を満たすシステムの実現に向けて、まず、既存システムと関連研究の現状をまとめ、次に、提案システムの機能の概要を述べる。

2.1 既存システムと関連研究における課題

先に述べた要件を満たすためには、ビデオ会議システムのサーバ型MCU（多地点接続装置：Multipoint Control Unit）における、複数の会議通話を同時に実施する機能および地点ごとのミキシング機能を、選択的に構成できる仕組みが必要である。MCUにはビデオ会議システムの端末内蔵型MCUとサーバ型MCUがあり、内蔵型MCUを用いるとHD品質の映像による4地点～10地点間[5]の通信が可能である。しかし内蔵型MCUには、地点ごとのミキシング機能は搭載されておらず、複数の会議通話を同時に実施することはできない。地点ごとのミキシング機能を使ったり複数の会議通話を同時に実施したりするためには、サーバ型の高機能なMCUが別途必要となる。しかし、通信イベント実施にあたり、一時的に集めやすい機材によるシステム構築を考慮した場合、サーバ型MCUの利用は前提にできない。ソフトウェアによるMCUを搭載した汎用PCで動作するシステムも存在するが、専用装置として提供されているサーバ型MCU同様、バックエンドで常時サーバとして稼働させることを想定しており、一時的な通信イベントにおける利用を前提としたものではない。

一方、中継機能とミキシング機能を一時的なバックエンドシステムで構成するにあたり、参考となる関連研究やサービスの事例がある。映像や音声データの転送プロトコルであるRTP[6]には、トランスレータとミキサの機能が定義されている。トランスレータはRTPストリームにおけるフォーマット変換等をとともなう中継システムであり、ミキサは複数の送信元からのRTPストリームを1つに合成して出力する中継システムである。これらの機能は、本提案システムの中継機能およびミキシング機能と同様の機能であるが、その分散構成方法はRTPでは定義されていない。トランスレータやミキサの構成に関する過去の研究としては、メディアストリームに対し、トランスレータやミキサの処理をアプリケーションレベルで行うメディアゲートウェイシステム[7]や、複数のマルチキャストセッションを統合した双方向の通信セッションを実現する方法[8]が提案されている。しかしこれらは会議通話におけ

る複数の構成を考慮した分散構成方法を述べたものではなく、上述した要件を満たしていない。また、映像の圧縮/展開を考慮した機能適合に関する研究[9]や、インターネットにおける分散システムの最適なサーバ割当てに関する研究[10]、クラウドにおける複数サーバの構成[11]や非集中型の適合方法に関する研究[12]等、バックエンドシステムの機能の分散構成に関して参考となる研究が報告されている。さらに、クラウドを想定した会議通話システムも研究開発されており[13]、[14]、[15]、クラウドサービスとしての事例もある[16]。しかしながら、通信イベント用のネットワーク環境における会議通話を前提とし、中継機能とミキシング機能の拡張を必要に応じて可能とするバックエンドシステムの分散構成方法は確立されていない。

さらに、中継機能とミキシング機能をバックエンドで分散するためには、適切なストリームパスを複数の端末で動的に構成する機能が必要となる。これまでに、ビデオ会議におけるストリームのルーティングに関する研究[17]や、P2Pのビデオ会議に有効なストリームのモデルの提案[18]があり、帯域幅の制約があるヘテロジニアスな環境におけるライブビデオ配信の研究[19]も行われている。これらの研究におけるストリームのモデルは汎用のモデルとしては有用なものだが、会議通話の主な構成をバックエンドで容易に実現するための分散型ストリームモデルの提案はなされておらず検討の余地がある。

2.2 提案システムの概要

既存システムと関連研究をふまえ、提案システムの概要を述べる。まず、要件1に対しては、一時的に準備しやすい汎用PCを用いて、会議通話システムを容易に構築できるシステムの実現を目指す。要件2に対しては、通信イベント用ネットワーク環境でバックエンドシステムを一時的に構築し、中継機能とミキシング機能を構成できる仕組みを実現する。要件3については、バックエンドシステムにおける端末の数を必要に応じて増やすことで、同時に実施可能な会議通話数を増やしたり地点ごとのミキシング機能を利用したりできる仕組みを実現する。

システム実現にあたり、会議通話の主な構成と構成ごとに必要となるストリームパスはパターン化することができる。そのパターンに従ったストリームパスを選択的に構成できれば、中継機能とミキシング機能を複数の端末上で分散する際の複雑さが軽減され、結果として実装の容易さと運用のしやすさにつながる。そこで本論文では会議通話に必要なストリームパスをパターンに分けてモデル化する。そのパターンを用いて、映像に関しては、各地点共通の画面イメージを生成する共用ミキサと、各地点別々の画面イメージを生成する専用ミキサを選択的に利用できる仕組みを実現する。音声に関しては、会議通話の各地点における音声出力をそれぞれ調整できるよう、音声の専用ミキ

サを地点ごとに利用できる仕組みを実現する。また、異なるストリームパスのパターンを組み合わせることにより、ネットワークの末端で利用可能な帯域幅や要求されるミキサの用途に応じて、1つの会議通話の中でも、地点ごとに適切な品質の映像と音声や、共用ミキサまたは専用ミキサの機能を提供することが可能となる。

以上、提案システムの概要を述べた。通信イベントの支援を目的とし、一時的に容易に構築可能であり、バックエンドで機能分散できる会議通話システムを実現する。最小構成1台から、必要に応じて台数効果のあるバックエンドシステムを実現することが目標となる。

汎用PCを用いるので、利用する端末の性能と入出力装置によって処理可能な映像と音声の品質は変わるが、普及しているHD品質の映像を処理可能なプロトタイプシステムを実装する。近年では備え付けのプロジェクトヤディスプレイも高解像度な表示装置が増え、映像入力装置としてのビデオカメラもHD品質が主流となっている。そこで、入出力装置の規格として比較的揃えやすいHD-720(720p)相当の映像を基準とし、プロトタイプシステムを構築する。

本論文では、プロトタイプシステムで使用した端末における台数効果を確認するため、複数の端末でバックエンドシステムを構成する環境を用意し実験を行った。既存の主なHD対応型ビデオ会議専用端末を用いる場合と同程度の地点間における会議通話を、汎用PCを用いたバックエンドシステムでも実現できることを実験結果から示す。さらに、高機能なサーバ型MCUを必要とする地点ごとのミキシング処理を、複数台の汎用PCで分散処理することにより実現できることも実験結果から示す。

3. 会議通話用バックエンドシステムの構成

中継機能とミキシング機能を複数台の汎用PCによりバックエンドで分散構成するためには、会議通話を実施する利用者の端末における状態に合わせて機能分散を行うためのプロトコルが必要となる。ここでは、まず会議通話用バックエンドシステムの構成を示し、会議通話を実施する利用者端末内の状態遷移について述べる。そして、会議の開催および会議への参加の際に必要な手順を示す。

提案システムを構成する端末の概要を図1に示す。会議通話のフロントエンドとなる利用者端末(U: User Terminal)は、会議サーバ(C: Conference Server)を利用して利用者へ会議通話の機能を提供する。会議サーバは1台でも会議通話の機能を利用者端末へ提供できるが、複数の会議通話を同時に行う際必要に応じてその会議通話数を増やせるよう、複数の会議サーバを配置することができる。複数の会議サーバを配置する場合、各利用者端末は、どの会議サーバを利用するかポータル会議サーバ(P: Portal Conference Server)へ問い合わせ、ポータル会議サーバは適切な会議サーバを利用者端末へ割り当てる。また、会議情報データ

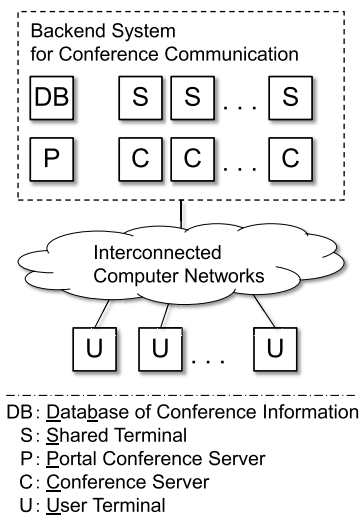


図1 会議通話用バックエンドシステムの構成

Fig. 1 Configuration of backend system for conference communication.

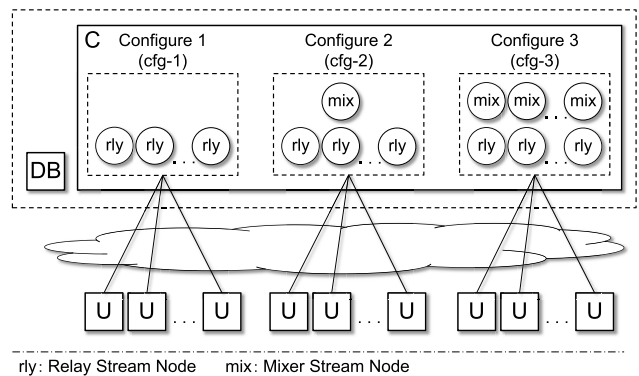


図2 提案システムで想定する3つの機能構成

Fig. 2 Three functional configurations supported by proposed system.

ベース(DB: Database of Conference Information)には、会議通話を実施する際に参照すべき会議室情報や利用者のアカウント情報が格納されており、ポータル会議サーバまたは会議サーバがその情報を参照する。さらに、複数の会議サーバから共通の計算機資源として利用でき、会議通話で要求されるミキシング機能をバックエンドシステム内で分散構成するための共用端末(S: Shared Terminal)を、必要に応じて配置することにより、会議サーバはミキシング機能を分散させることができる。一方、会議情報データベースを会議サーバに内包することで、会議サーバ1台による通信イベントの支援が可能となることも考慮する。

3.1 想定する中継機能とミキシング機能の構成

提案システムにおいて想定している中継機能とミキシング機能の構成を図2に示す。会議サーバはこれら3つの構成を組み合わせることで会議通話を実施する。ここで、中継ストリームノード(rly: Relay Stream Node)は、映像・音声ストリームの中継処理をするプログラムであり、複数の

ストリームを中継する際は複数の中継ストリームノードが利用される。また、ミキサストリームノード (mix: Mixer Stream Node) は、映像または音声ストリームのミキシング処理をするプログラムであり、必要に応じて複数のミキサストリームノードが利用される。

構成1 (cfg-1: Configure 1) は、中継機能のみの構成である。利用者端末が映像または音声を送信すると、他の利用者端末へ会議サーバが中継する。構成2 (cfg-2: Configure 2) は、中継とミキシング機能の両方を利用する構成である。利用者端末が映像または音声を送信すると、中継ストリームノード経由でミキサストリームノードの入力となり、ミキシング処理が行われる。ミキシング結果は1つであり、そのコピーが複数の利用者端末へ送信される。本論文では、構成2におけるミキサを共用ミキサと呼ぶ。共用ミキサの操作は各利用者端末から排他的に行うことが可能である。一方、構成3 (cfg-3: Configure 3) では、構成2と同様に中継とミキシング機能の両方を利用するが、共用ミキサを利用するのではなく利用者端末ごとのミキサを利用している。本論文では、利用者端末ごとのミキサを専用ミキサと呼ぶ。専用ミキサの操作は利用者端末ごとに行うことが可能であり、利用者端末ごとに異なる映像または音声のミキシング結果を受信して再生表示することが可能となる。

会議サーバは、中継機能とミキシング機能の3つの構成に対応する。会議通話の際、映像ミキサを共用ミキサとして用いて、音声ミキサは専用ミキサを利用するというような用途に対応することも考慮し、映像については構成2、音声については構成3という複合的な構成の会議通話が可能となる。

次に示す図3は、バックエンドシステムとして複数の会議サーバを配置した例である。1台の会議サーバでも会議通話を実施できるが、複数の会議通話を同時に実施する際に要求されるネットワークの帯域幅や、ミキシング処理の負荷を考慮すると、複数の会議サーバによる機能分散が必要となる。図3は、cfg-1~cfg-3の構成による通信をそれ

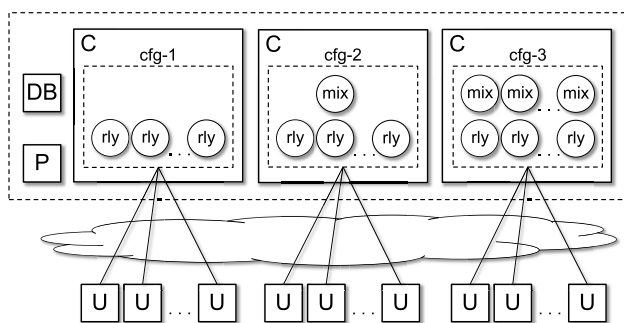


図3 複数の会議サーバを配置したバックエンドシステム

Fig. 3 Backend system that is composed of multiple conference servers.

ぞれ別の会議サーバで実施している図となっているが、各会議サーバは図2同様に cfg-1~cfg-3 いずれにも対応できるものとする。また、複数の会議サーバを配置する場合は、新しく会議通話を開始する際の会議サーバを選定する機能と、実施中の会議通話に対応しているのはどの会議サーバであるかを確認する機能が必要となるので、その機能を実現するためにポータル会議サーバを配置している。

さらに図4は、共用端末を配置してミキシング機能を分散させた例である。構成2 (cfg-2s: Configure 2 with Shared Terminal) においては、利用者端末への映像または音声ストリームの中継機能と、共用ミキサによるミキシング機能を分散させることが可能となる。また、構成3 (cfg-3s: Configure 3 with Shared Terminals) においては、専用ミキサによるミキシング機能を複数の共用端末で分散させることが可能となる。負荷の高い映像のミキシング処理も、複数の共用端末を配置することで、利用者端末ごとの専用ミキサを使うことが可能となる。

会議通話を実施する際、バックエンドにおける中継機能とミキシング機能を利用するためには、利用者端末の状態に応じたバックエンド側での機能分散処理が必要となる。引き続き、利用者端末内の状態と会議通話に必要な手順について述べる。

3.2 会議オーナーと会議メンバの状態

上述した中継機能とミキシング機能の3つの構成を利用した会議通話の実現にあたり、提案システムでは、会議通話を開始する利用者端末のプログラムを会議オーナーと呼び、会議通話に参加する利用者端末のプログラムを会議メンバと呼ぶ。各構成に対応可能なシステムを実現するために必要となる会議オーナーと会議メンバの状態を、状態マシンとして示したのが図5と図6である。

図5に示すとおり、会議オーナーは Closed 状態から open-session イベントにより会議通話開始処理に入る。共用ミキ

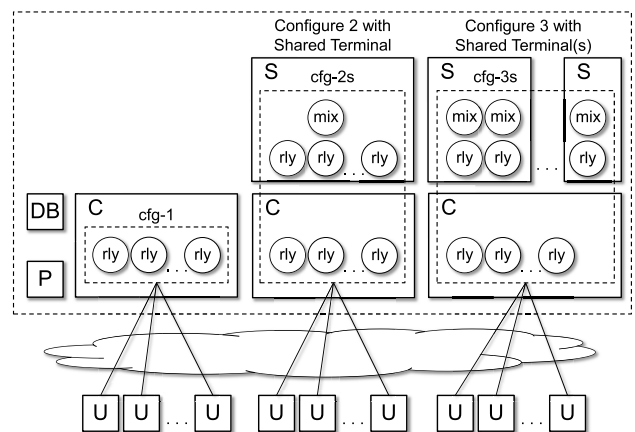


図4 共用端末によるミキシング機能の分散

Fig. 4 Functional distribution of mixing processes using shared terminals.

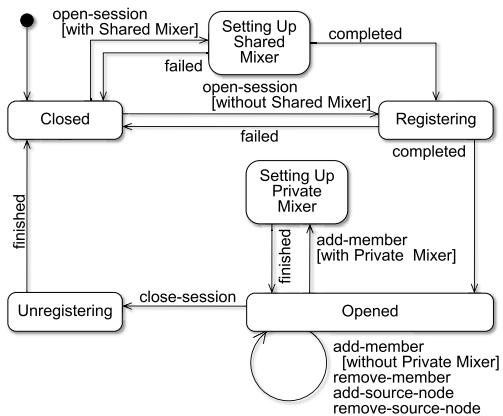


図 5 会議オーナーの状態マシン図

Fig. 5 State machine diagram of conference owner.

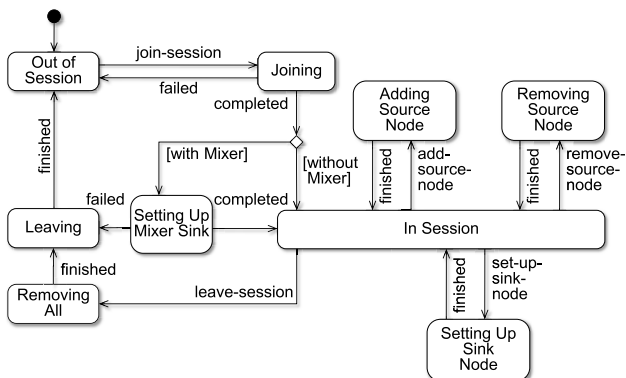


図 6 会議メンバーの状態マシン図

Fig. 6 State machine diagram of conference member.

サを必要としない場合は Registering 状態に遷移し、利用する会議サーバに会議室情報を登録する。共用ミキサを必要とする会議通話の場合は Setting Up Shared Mixer 状態の中で共用ミキサをバックエンドシステム内に準備したうえで Registering 状態に遷移する。会議サーバへの会議室情報の登録が成功すると Opened 状態へ遷移する。

Opened 状態では、会議メンバーの追加/削除 (add-member/remove-member)、会議メンバーによるソースストリームノードの追加/削除イベント (add-source-node/remove-source-node) に対応する処理を実行する。会議メンバーが会議通話参加の際に専用ミキサの利用を要求する場合は、Setting Up Private Mixer 状態へ遷移し、専用ミキサをバックエンドシステム内に準備したうえで再度 Opened 状態へ戻る。ここで、会議通話に必要なミキサの確保と解放は会議オーナー主導で行われる。これは会議通話に必要なバックエンドシステムの資源を会議通話ごとに一元管理することで、バックエンドシステム内の資源割当てを効率良く行うためである。

また、Opened 状態で処理されるイベントはすべて、会議通話に参加中の会議メンバーへも通知される。たとえば、会議メンバーの端末で映像または音声の送信を新たに開始する際には、ソースストリームノードを追加するためのメッ

セージが会議メンバーから会議オーナーへ送信され、これが add-source-node イベントとなる。会議オーナーはイベントに対し、必要に応じてストリームノードの接続を行い、その結果はすべての会議メンバーへ通知される。

最後に会議通話を終了する際は、close-session イベントが発行され、Unregistering 状態へ遷移する。この状態では、利用していた会議サーバへ登録していた会議室情報を削除する。また、バックエンドシステム内で利用していた共用ミキサはここで解放される。

会議メンバーは、会議通話への参加と退出、ソースストリームノードの追加と削除が可能であり、図 6 はその状態マシンを示している。

図 6 のとおり、会議メンバーは Out of Session 状態から join-session イベントにより会議参加処理状態 (Joining) へ遷移する。専用ミキサを利用する場合は、会議オーナーがそのセットアップ処理を実行する。そして、共用か専用かにかかわらずミキサをバックエンドで利用する場合は、Setting Up Mixer Sink 状態に遷移し、バックエンドのミキサストリームノードから利用者端末へのストリームノードまでの接続を行う。

会議通話に参加すると、In Session 状態へ遷移する。この状態では、ソースストリームノードの追加/削除 (add-source-node/remove-source-node) が可能である。また、他の利用者端末において追加されたソースストリームノードの情報を会議オーナーから受け取った場合、これを set-up-sink-node イベントとして扱い Setting Up Sink Node 状態へ遷移し、他の利用者端末で送信が始まった映像・音声を受信するためのストリームノードを生成する。ここで必要となるストリームノードは、中継機能とミキシング機能の構成によって、利用者端末内であるかバックエンドシステムの端末内であるかが決まる。その詳細は 4 章で述べる。

会議通話から退出する際は、leave-session イベントが発行され Removing All 状態へ遷移する。ソースストリームノードはすべてこの状態の中で会議通話から削除され、専用ミキサを利用していた場合もこの状態の中でバックエンドシステムにおける資源の解放処理が実行される。最後に Leaving 状態では、会議オーナーへ退出メッセージを送信してその返信を受け、会議通話からの退出が完了する。

ここまで、会議オーナーと会議メンバーそれぞれの状態マシンについて述べた。引き続き、会議オーナーによる会議通話の開始手順、および会議オーナーと会議メンバーが連携動作する会議通話への参加の手順について述べる。

3.3 会議開始と参加の手順

会議通話を行うためには、バックエンドシステムとして配置されている複数の端末とフロントエンドの利用者端末間におけるメッセージ交換が必要となる。本論文における会議通話用のバックエンドシステムは、通信イベント実施

の際一時的に構築されることを想定しており、構成端末の台数は可変である。しかし、会議情報データベースを内包する会議サーバ1台のみでバックエンドが構成されているか、複数の端末で構成されているかにかかわらず、利用者端末側では同じ手順で会議通話を実施できることが望ましい。すでに述べた会議オーナーと会議メンバの状態マシンは利用者端末内の状態であるが、バックエンドシステムでは利用者端末内の状態とバックエンドを構成する端末数に応じて適切な会議サーバと共用端末を会議通話に適用する必要がある。これらを考慮した、会議通話を開始する際のメッセージフローを図7に、会議通話へ参加する際のメッセージフローを図8に示す。

会議通話を行う際には、あらかじめ会議情報データベースに格納されている会議室情報を参照する。会議室情報には、会議室名と参加可能な利用者IDが含まれている。一方、会議通話中に変化しうる情報は会議オーナーが管理する。本提案システムではこれを会議通話情報と呼ぶ。会議通話情報には、現在会議通話へ参加している利用者および利用者の端末情報や、送受信されている映像と音声ストリームのパス等が含まれる。

会議通話を開始する際、会議オーナーとなる利用者端末

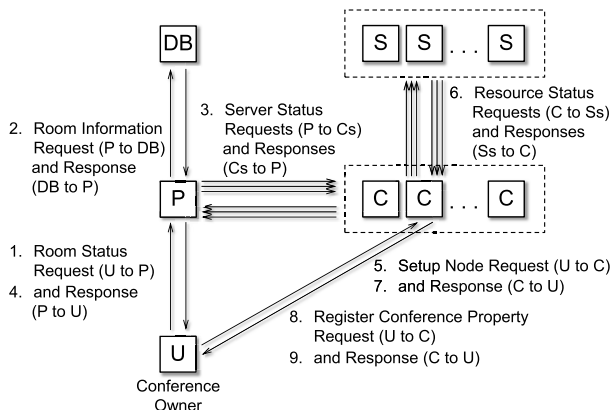


図7 会議通話開始時のメッセージフロー

Fig. 7 Message flow for setting up conference communication.

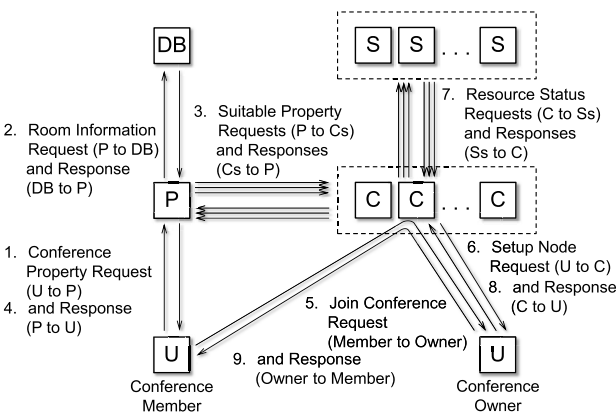


図8 会議通話参加時のメッセージフロー

Fig. 8 Message flow for joining to conference communication.

(図7における Conference Owner) は、最初に Room Status Request メッセージをポータル会議サーバへ送信する。ポータル会議サーバは、必要に応じて利用者認証を行った後、会議室情報を会議情報データベースへ問い合わせる (Room Information Request and Response)。次に複数の会議サーバへ Server Status Request メッセージを送信し、その応答を受ける。この応答には会議サーバの資源利用状況として CPU 利用率と使用帯域が含まれている。ポータル会議サーバは負荷の低い会議サーバを選び、そのサーバの IP アドレスを含む Room Status Response メッセージを会議オーナーの利用者端末へ送り返す。ここまでは、利用者端末における Closed 状態 (図5) の処理となる。引き続き、会議オーナーの利用者端末は、選択された会議サーバ上で会議通話を開始するための準備を進める。開始する会議通話が共用ミキサを用いる場合は、会議サーバに対して Setup Node Request メッセージを送信する。会議サーバは利用可能な共用端末の資源利用状況を問い合わせ (Resource Status Requests and Responses)、負荷の低い共用端末上で共用ミキサを動作させる準備を進める。これは、利用者端末における Setting Up Shared Mixer 状態 (図5) の処理となる。その後、図5における Registering 状態へ遷移した利用者端末は Register Conference Property Request メッセージを会議サーバへ送信し、開始する会議室情報を登録する。登録後、会議オーナーの利用者端末は Opened 状態 (図5) へ遷移し、会議参加要求に対応できる状態となる。ここで、共用ミキサを用いない場合は、図7における5~7のメッセージの送受信は行われない。

会議通話へ参加する際は、会議メンバとなる利用者端末 (図8における Conference Member) が、最初に Conference Property Request メッセージをポータル会議サーバへ送信する。会議通話開始時の流れ同様、必要に応じて利用者認証が行われ、会議室情報を会議情報データベースへ問い合わせる。次に複数の会議サーバへ Suitable Property Request メッセージを送信し、その応答を受ける。この応答には、各会議サーバで現在実施されている会議室情報が含まれている。ポータル会議サーバは、会議メンバが参加できる会議通話を実施している会議サーバの会議室情報をリスト化し Conference Property Response メッセージを会議オーナーの利用者端末へ送り返す。ここまでは、利用者端末における Out of Session 状態 (図6) の処理となる。

引き続き、会議メンバの利用者端末は、参加する会議通話を実施している会議サーバを経由して会議オーナーへ Join Conference Request メッセージを送信する。ここで利用者端末の状態は図6の Joining 状態に遷移する。会議通話に参加する会議メンバが専用ミキサを利用する場合、会議オーナーは Setting Up Private Mixer 状態 (図5) へ遷移し、専用ミキサを共用端末上で動作させる準備を進める (図8の6~8)。専用ミキサの準備が終了すると、会議

オーナーは Opened 状態 (図 5) へ遷移し、会議メンバへ Join Conference Response メッセージを送る。その後、会議メンバ側では、必要に応じてミキサからの出力を受けるシンクストリームノードを準備し、会議通話参加状態となる。

以上、3章では、提案システムの構成を示し、利用者端末内のプログラムの状態マシン、および会議通話開始と参加に必要なメッセージフローを示した。続く4章では、映像と音声の双方向通信を行うためのストリームノードの接続について、構成ごとに整理してモデルを示す。

4. ストリームノードの接続モデル

会議通話者が映像または音声の送信を開始したり終了したりするタイミングで、ストリームの適切な接続処理が必要となる。中継機能とミキシング機能を構成するためのストリームパスを整理すると、構成ごとにパターン化することができる。提案システムでは、パターン化されたストリームパスを必要に応じて選択的に構成することで会議通話を実現する。以下、構成ごとにパターン化されたストリームノードの接続モデルについて述べる。

4.1 各構成におけるストリームノードの接続

cfg-1 (図 2) は中継処理のみを必要とする構成であり、図 9 はそのストリームノードの接続モデルを示している。図 9 は、会議サーバを利用して3台の利用者端末間で会議通話を行っている例であり、映像または音声データ送信の始点となるソースストリームノードを src と表し、映像または音声データを受信して表示または再生する終点となるシンクストリームノードを snk と表している。

図 9 は、利用者端末 U₁ のみが映像または音声データを1本送信している状態を示している。利用者端末上でソースストリームノードが追加されると、会議サーバ C の中には中継ノードが生成され、他の利用者端末 (U₂ と U₃) ではそれぞれシンクストリームノードが生成される。ソースストリームノードが削除される際は、そのソースストリームノードを始点とするストリームノードのパス全体が削除される。

cfg-2 (図 2) は、バックエンドにおいて共用ミキサを利用する構成であり、図 10 (a) はそのストリームノードの接続モデルを示している。共用ミキサを利用する場合、会議通話開始時に会議オーナーが共用ミキサを準備し、会議メンバは会議通話参加時に共用ミキサから利用者端末までのストリームノードをあらかじめ接続しておく。利用者端末においてソースストリームノードが追加/削除されるタイミングで、会議サーバ上でも中継ストリームノードが追加/削除される。会議サーバ上の中継ストリームノードは、共用ミキサの入力となる。

一方、図 10 (b) は共用端末を利用する構成 cfg-2s (図 4) におけるストリームノードの接続モデルを示している。共

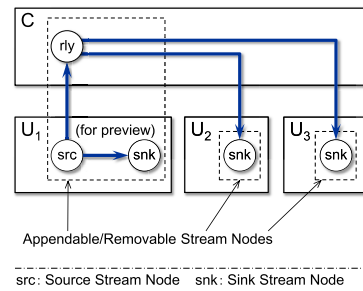


図 9 cfg-1 におけるストリームノードの接続モデル
Fig. 9 Connectivity model of stream nodes in cfg-1.

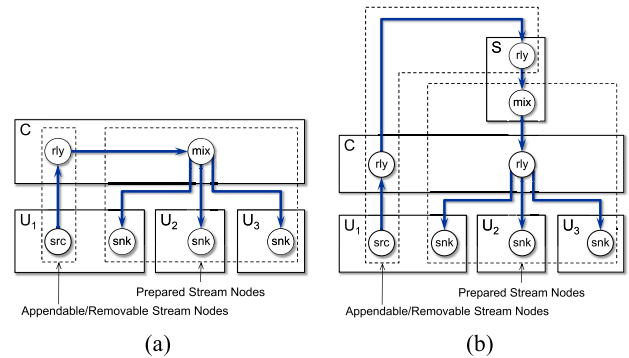


図 10 cfg-2 および cfg-2s におけるストリームノードの接続モデル
Fig. 10 Connectivity models of stream nodes in cfg-2 and cfg-2s.

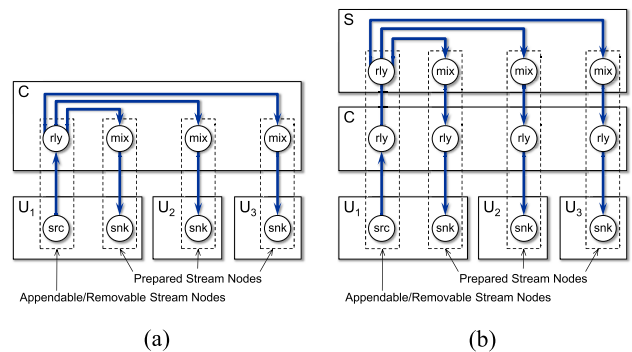


図 11 cfg-3 および cfg-3s (単一共用端末) におけるストリームノードの接続モデル
Fig. 11 Connectivity models of stream nodes in cfg-3 and cfg-3s (1 shared terminal).

用端末を利用していない場合と同様、共用ミキサから利用者端末までのストリームノードはあらかじめ接続される。利用者端末においてソースストリームノードが追加/削除されるタイミングで、会議サーバと共用端末間も中継ストリームノードによってストリームのパスが追加/削除される。

会議通話 cfg-3 (図 2) は、バックエンドにおいて専用ミキサを利用する構成であり、図 11 (a) はそのストリームノードの接続モデルを示している。会議メンバは会議通話へ参加する際に会議オーナー経由で専用ミキサを準備する。また、会議メンバは会議通話参加時に専用ミキサから利用

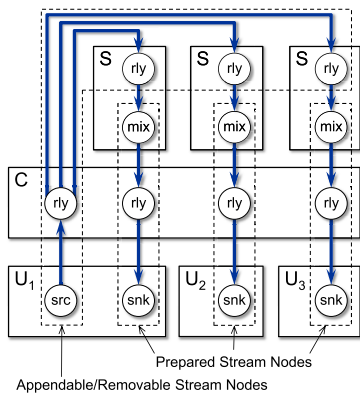


図 12 cfg-3s (M 台の共用端末) におけるストリームノードの接続モデル

Fig. 12 Connectivity models of stream nodes in cfg-3s (M shared terminals).

者端末までのストリームノードをあらかじめ接続しておく。利用者端末ごとのミキサノードへ入力を与える中継ストリームノードは、ソースストリームノードから 1 本のストリームを受信すると、端末内で複数のミキサノードへストリームデータを分配する。図 11 (b) は共用端末を 1 台利用した場合の構成 cfg-3s (図 4) におけるストリームノードの接続モデルである。

そして図 12 は、構成 cfg-3s において、複数台の共用端末を利用した場合のストリームノードの接続モデルを示している。この図では、利用者端末と同数の共用端末が配置され、それぞれの共用端末内で利用者端末ごとのミキシング処理を行っている。利用者端末上でソースストリームノードが追加されると、会議サーバに中継ストリームノードが生成される。共用端末 1 台の場合と異なり、生成された中継ストリームノードは各共用端末のミキサストリームノードへストリームデータを中継する。

ここまで、各構成におけるストリームノードの接続モデルを示した。パターン化されたモデルを用いることで、中継機能とミキシング機能を複数の端末上で分散させる際の複雑さの軽減が期待できる。一方、機能を分散することにより、バックエンドを構成する端末間で送受信されるストリームの数は増える。ストリームモデルを実際の会議通話へ適用することを考えると、バックエンドシステムにおける台数効果にストリーム数が関係するので、次に、各構成において必要となる入出力ストリーム数を整理する。

4.2 各構成と入出力ストリーム数

表 1 のとおり、バックエンドで必要となる入出力ストリーム数は構成ごとに異なる。ここで、N 台の利用者端末が 1 つの会議通話に参加していることを想定し、各利用者端末からは 1 つのソースストリームノードによる映像または音声ストリームが送信されているものとする。また、各構成におけるストリームノードの接続モデルに従って各端

表 1 各構成における入出力ストリームの数
Table 1 Number of I/O streams in each configuration.

		User Station	Conference Server	Shared Terminal	
cfg-1	in	$N - 1$	N	/	
	out	1	$N \cdot (N - 1)$		
cfg-2	in	1	N	/	
	out				
cfg-2s	in		$N + 1$	N	
	out		$2N$	1	
cfg-3	in		1	N	/
	out				
cfg-3s (1 shared terminal)	in	$2N$		N	
	out				
cfg-3s (M shared terminals)	in	$2N$		N	
	out	$N \cdot (M + 1)$		1	

末は中継とミキシング処理を行っているものとする。

表に示すとおり、cfg-1 における利用者端末と会議サーバの入力ストリーム数は利用者端末の数に比例するが、会議サーバの出力ストリーム数は利用者端末数の二乗に比例する。cfg-2 以降の利用者端末における入出力ストリーム数は、入力出力ともに 1 本であるが、その分、バックエンドシステム側のストリーム数は増える。会議通話を構成する端末ごとのストリームの本数のみを考慮すると、N 台の利用者端末と同数の共用端末をバックエンドに配置した場合 cfg-3s (M shared terminals: $M = N$) における会議サーバの出力ストリーム数が最も多くなる。

cfg-1 において、各利用者端末が受信する映像・音声ストリームの数は、その利用者端末以外の利用者端末から送信されている映像・音声ストリームの数と同じである。映像・音声のミキシング処理をバックエンドで行わないので、ミキシング処理が必要となる構成 2 (cfg-2, cfg-2s) および構成 3 (cfg-3, cfg-3s) に比べるとバックエンドの負荷は低い。一方で、利用者端末では複数本の映像・音声ストリームを受信し、再生表示する必要があるため、特に高解像度の映像を用いる場合は、利用者端末の負荷が高くなる構成といえる。

共用ミキサを利用する構成 2 (cfg-2, cfg-2s) では、バックエンドシステム内でミキシング処理を行うため、バックエンドの負荷は構成 1 に比べて高くなるが、利用者端末の負荷は構成 1 に比べて低くなる。

そして、専用ミキサを利用する構成 3 (cfg-3, cgf-3s) では、複数のミキシング処理を必要とするのでバックエンドシステム内では構成 2 より負荷がかかるが、利用者端末の負荷は構成 2 と同様となる。

映像または音声ストリームが必要とするビットレートは、要求される映像と音声の品質や会議通話を構成する各端末の性能、そして利用するネットワーク環境に依存する。表 1 は入出力ストリームの数のみを示しているが、利用す

るストリームのビットレートが決まると、会議通話で必要となる帯域幅も求められる。しかし、必要となる帯域幅が求められたとしても、その構成での会議通話が可能であるかは、バックエンドシステムを構成する端末の性能に依存する。提案システムは、バックエンドシステムを構成する端末の性能を限定するものではないが、提案システムを用いて各構成の会議通話が可能であることを示し、バックエンドシステムに配置する端末の数に応じて機能的な拡張が可能となることを示すために、プロトタイプシステムを用いて機能的な評価実験を行う。続く5章ではそのプロトタイプシステムの実装について述べる。

5. プロトタイプシステムの実装

中継処理とミキシング処理の分散構成における会議通話の実現に関して、実機を用いた機能的な検証を行うため、Windows 7を搭載した汎用PC上にプロトタイプシステムを実装した。図13は利用者端末、会議サーバおよび共用端末上で動作するソフトウェアの構成概要を示している。

まず、各端末で動作するソフトウェアの共通コンポーネントとして、本システム内部で発生する各種のイベントを処理するためのSystem Event Handlerを、複数のスレッドプールを用いて実装した。独立した各種のイベントを複数のスレッド上で処理させる一方、時間的に直列した処理を必要とするメッセージの配送等には単一のスレッドを用いている。また、メッセージ用とストリーミング用の複数のTCPコネクションを管理し、端末間のメッセージ送受信処理を行うためのInter-System Message Handlerを実装した。図7と図8で示したメッセージのためのパケット送受信はここで行われる。SIP同様にテキストベースのメッセージプロトコルを実装した。

会議通話を実施するために必要となる会議オーナーと会議メンバの機能は、利用者端末でのみ動作するConference Componentsとして実装した。これらのソフトウェアコンポーネント（Conference Owner および Conference Mem-

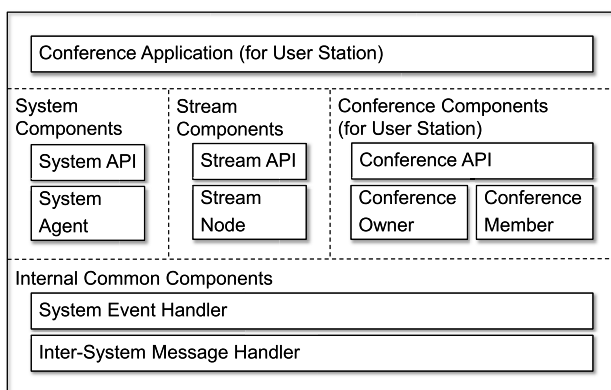


図13 プロトタイプシステムにおけるソフトウェアコンポーネント
Fig. 13 Software components of prototype system.

ber)が、図5と図6に示した状態遷移マシンを実装し、図7と図8に示した会議通話ごとのメッセージの処理を行う。アプリケーションに対しては会議通話に関する各種の操作をまとめたConference APIを提供する一方、Conference MemberはStream ComponentsのStream APIを通して映像と音声の送受信を行う。また、入出力ストリームのビットレートやCPU利用率の測定、および会議通話ごとのメッセージ以外のシステムメッセージを処理するためにSystem Componentsを実装した。

映像と音声のストリーミング機能は、既存システム[20]のストリーミング機能を改良し、図13におけるStream Nodeとして実装した。Stream Nodeの内部におけるメディア処理のフレームワークとしてはDirect Show[21]を用いており、映像と音声の伝送にはRTPを基礎としたプロトコルをTCP上で用いている。ストリーム転送用のパケットヘッダには、シーケンスナンバとペイロードタイプ、および、ビデオフレーム（またはオーディオサンプルデータ）を分割した場合のオフセット値と、分割の最後を示すマーカビットが含まれている。

一方、文献[20]におけるストリーミング機能には、複数の端末によるストリームの多段接続機能が実装されているので、プロトタイプシステムではその機能を使い、パターン化されたストリームノードの接続モデル（図9～図12）を実現した。加えて、ストリームノードごとに識別子(UUID)を割り当て、ストリーム用のコネクション確立時にその識別子を送信元へ伝えることにより、ストリーム用のコネクションとストリームノードを管理している。これにより、ストリームノードが端末をまたがり多段結合した際に、送信元では端末内に送信ツリーを構成し、中継状況を把握できる仕組みを実装している。

また、映像および音声ミキサも文献[20]におけるミキサ機能を改良して実装した。文献[20]で実装されているミキサでは、入力を動的に追加/削除することができない。そこで本プロトタイプシステムでは、図14に示すとおりミキサの入力として動的に追加/削除可能なリングバッファを中継ストリームごとに用意し、中継ストリームノードの出力先を、そのリングバッファへつなぐ仕組みを新たに実

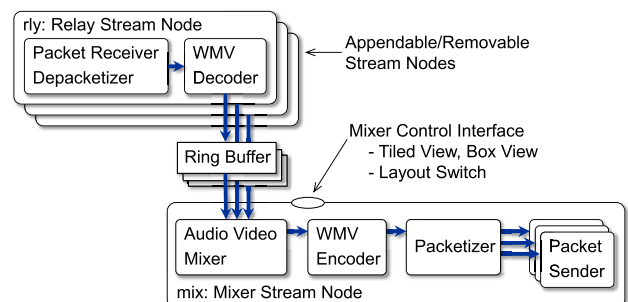


図14 中継およびミキサストリームノード
Fig. 14 Relay stream nodes and mixer stream node.

装した。映像ミキサの機能としては、複数の入力映像をタイル状に並べる Tiled View と、縦または横一列に並べる Box View の 2 つの表示を組み合わせたレイアウトを可能とし、そのレイアウトを複数用意して切り替える機能も実装した。音声ミキサの機能としては、PCM データのミキサと音量調整機能を実装した。そして、これらの機能を遠隔から利用可能なインタフェースを用意し、Inter-System Message Handler によるメッセージ通信機能を使って遠隔からのミキシング操作を実現している。

ミキサの入力となるリングバッファからは、ミキサノードが一定間隔（評価実験では、映像は 1/30 sec., 音声は 1/25 sec.）でストリームデータを取得している。その際、バックエンドシステム内における処理の時間的な揺らぎを吸収するため、ミキサの入力となるリングバッファでバッファリング（映像は 2/30 sec, 音声は 2/25 sec）をしているので、その分遅延する。一方、ミキシング処理と宛先ごとの送信処理を異なるスレッドに分けることにより処理の並列性を高め、中継時の遅延を抑える実装となっている。文献 [22] では、Web 会議システムにおいて発話衝突ならびに精神的ストレスが生じないようにするにはエンド間の音声遅延量を 200 ms 以下に抑えることが必要であるとしているが、プロトタイプシステムは、共用端末を用いた専用ミキサを適用した場合でも、この制約内での会議通話を実現している。

一方、会議情報データベースおよび利用者認証には、企業内ネットワーク等で利用されている LDAP (Lightweight Directory Access Protocol) サーバを適用し、プロトタイプシステムでは Windows 用の Open LDAP [23] を用いた。Open LDAP のプログラムは、会議サーバやポータル会議サーバの端末内で動作させることも可能であるので、最小構成として会議サーバ 1 台でのバックエンドシステムも実現することができる。評価実験においても、会議サーバまたはポータル会議サーバの端末上で Open LDAP を動作させて実験を行った。

また、プロトタイプシステムでは、利用者端末がポータル会議サーバへメッセージ通信用の TCP コネクションを確立した直後に LDAP を用いた利用者認証を行っている。一方、利用者端末が会議通話を実施する際には会議サーバへもメッセージ通信用の TCP コネクションを確立するので、その際も LDAP を用いた利用者認証を行っている。利用者端末がバックエンドシステムを利用するためには、ポータル会議サーバまたは会議サーバへの接続が必要であるので、その両方への接続時に利用者認証を行っている。

6. 評価実験と運用事例

プロトタイプシステムを実装した汎用 PC を用いて、実験用のバックエンドシステムを構築し、評価実験を行った。利用者端末における映像と音声のキャプチャには、各利

用者端末に搭載されているビデオカメラとマイクロフォンを用いた。キャプチャ映像には、720p（ピクセル解像度：1280×720 pixel, フレームレート：30 fps）および PCM（チャンネル数：1, サンプリング周波数：44.1 kHz, 量子化ビット数：16 bit, キャプチャ間隔：1/25 sec.）のフォーマットを用いている。映像の圧縮には WMV (Windows Media Video) の CBR (Constant Bit Rate) 8 Mbps の設定を利用した。音声はキャプチャした PCM データを圧縮せずにそのまま 705.6 Kbps の音声データを利用している。

汎用 PC を用いているので、処理可能な映像と音声の品質は利用している端末の性能に依存する。しかし、台数効果による機能の拡張が可能であることを確認するためには、プロトタイプシステムで利用している端末における処理能力の上限を把握しておく必要がある。評価実験で用いた WMV では 720p の映像を 2 Mbps 程度まで圧縮することが可能だが、バックエンドにおける入出力スループットの上限も確認するために、評価実験では CBR 8 Mbps の設定を用いた。一方、共用端末において専用ミキサの分散機能を確認する実験においては、利用するフォーマットによる台数効果の違いを確認するために 360p（ピクセル解像度：640×360 pixel, フレームレート：30 fps）のキャプチャ映像と WMV CBR 2 Mbps の設定もあわせて用いた。

図 15 に評価実験環境を示す。ポータル会議サーバを利用した複数会議サーバによる機能分散機能の確認と、各構成の会議通話を 1 台の会議サーバで実施した場合に会議通話可能な利用者端末数を確認するために、図 15 (a) の環境を利用した。また、共用端末を利用したミキサの機能分散の評価を行うためには、図 15 (b) の実験環境を利用した。また、バックエンドを構成する端末として用いた汎用 PC のスペックとして CPU と NIC の一覧を表 2 に示す。会議サーバ C₁~C₄ と共用端末 S₁~S₄ は同じ端末であるが、図 15 (a) と図 15 (b) の実験環境それぞれで役割を変更して実験を行った。利用者端末としては、上述した映像と音声処理に対応できる性能、かつ、ビデオカメラとマイクロフォンを搭載している汎用 PC を用いた。

以下、各評価実験とその結果について述べる。

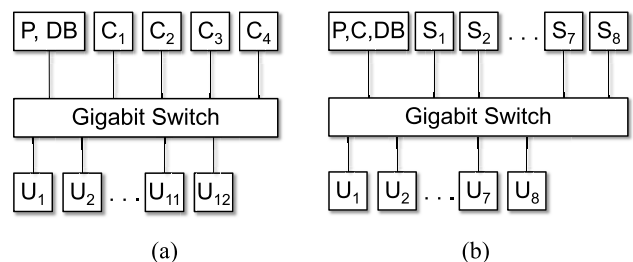


図 15 評価実験環境

Fig. 15 Experimental environments.

表 2 バックエンドを構成する端末のスペック

Table 2 Specifications of backend terminals.

	CPU	NIC
P,DB	Intel Core i7 2640M 2.80GHz	Realtek PCIe GBE Family Controller
C,DB	Intel Core i7 4770 3.40GHz	Intel Ethernet Connection I217-LM
C ₁ -C ₄ S ₁ -S ₄	Intel Core i7 3770 3.40GHz	Intel 82579LM Gigabit Network Connection
S ₅ -S ₆	Intel Core i7 2960XM 2.70GHz	
S ₇ -S ₈	Intel Core i7 3940XM 3.00GHz	

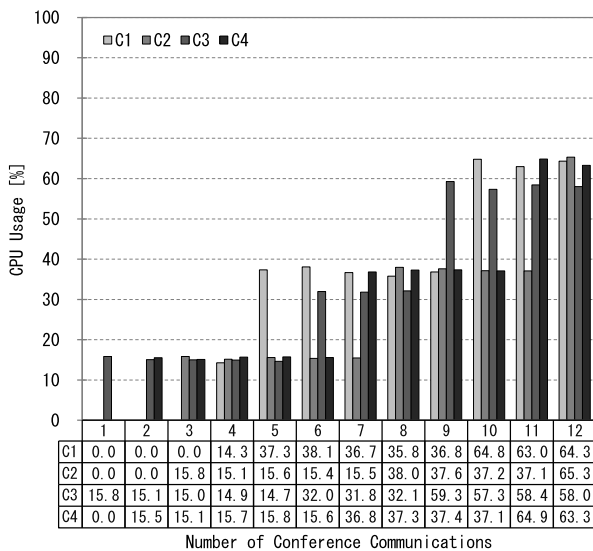


図 16 会議サーバ 4 台の CPU 利用率
Fig. 16 CPU usages of 4 conference servers.

6.1 ポータル会議サーバによる会議通話の分散

まず、ポータル会議サーバが複数台の会議サーバへミキシング処理を振り分けられることを確認する。

実験は図 15(a) の環境を用いて行い、映像ストリームには cfg-2 の共用ミキサを、音声ストリームには cfg-3 の専用ミキサを適用した。ポータル会議サーバは、図 7 および図 8 に示すとおり 1 台とし、会議サーバは同型の端末を 4 台用いた。ここで、1 つの会議通話につき 4 台の利用者端末が参加することを想定すると、映像に cfg-2 の共用ミキサを適用した場合、次節における会議サーバ 1 台の性能 (図 19) から 1 つの会議通話に対して約 32% の CPU 資源が必要となる。したがって、1 台の会議サーバで処理可能な会議通話数はせいぜい 3 つなので、4 台の会議サーバに対して 12 台の利用者端末を用いた。また、図 15(a) の実験環境におけるポータル会議サーバは、会議情報データベースを内包している。

図 15(a) の環境において、12 台の利用者端末が U₁ から U₁₂ の順番でそれぞれ会議通話を開始した場合の、4 台の会議サーバの CPU 利用率を測定した。それぞれの利用者端末が会議通話開始後、約 3 分間、1 秒ごとに測定した CPU

利用率の平均値をグラフにしたのが図 16 である。図 16 の縦軸およびグラフの下表は各会議サーバの CPU 利用率を示しており、横軸は会議通話数を示している。ポータル会議サーバが機能分散を行う際に利用する CPU 利用率の値としては、各会議サーバにおける過去 30 サンプル (30 秒間) の CPU 利用率の移動平均値を用いた。

図 16 のとおり、会議サーバの CPU 利用状況に応じて負荷の低い会議サーバが選択されることを確認できた。図 15 において、会議通話の数が 6 から 7 に増えたときのみ、C₂ (15.4%) よりも CPU 利用率が高い C₄ (15.6%) が選択されている。図 16 のグラフ下部に示している値は上述したとおり約 3 分間の平均値であり、ポータル会議サーバが各会議サーバへ CPU 利用率を問い合わせた際の CPU 利用率は、C₂ が 15.88%、C₄ は 15.47% だった。つまり問い合わせた時点で CPU 利用率の低かった C₄ が選択され、利用者端末に割り当てられたことになる。また実験を通して、映像ストリームは図 10(a)、音声ストリームは図 11(a) のとおりにストリームの接続が行われストリーミング処理が行われたことを確認した。

プロトタイプシステムで用いている会議サーバでは、720p の映像ミキサを 1 つ動作させるごとに 15~16% の CPU 資源を必要とする。当然、会議通話参加者が増えるごとに、より多くの CPU 資源が必要となる。CPU 負荷の分散にはラウンドロビンを用いる手法や、コネクション数で端末を選択する手法等があるが、ミキシング処理に多くの CPU 資源を必要とする会議通話のバックエンドシステムにおける機能分散方法としては、CPU 利用率による振り分け方法が効果的に機能するものと考えられる。

6.2 会議サーバ 1 台における各構成の会議通話

引き続き、会議サーバ 1 台における性能評価実験の結果をまとめる。cfg-1、cfg-2 および cfg-3 それぞれの構成で会議通話を実施した場合に、参加する利用者端末の数を増加させて会議サーバの性能評価を行った。

実験では、1 台の利用者端末で会議通話を開始し、その端末が会議通話へ参加する。以後、残りの利用者端末は順番に会議通話へ参加する。各利用者端末は会議通話へ参加した後、映像と音声ストリームを各 1 本ずつ送信する。映像データは約 8 Mbps、音声データは 705.6 Kbps である。映像および音声データの送信の際にはそれぞれ 1,200 byte と 1,000 byte にデータを分割した後 12 byte のヘッダを付加したパケットを、TCP を用いて送受信している。測定した入出力スループットは、この 12 byte のヘッダが付加された状態のデータストリームのものであり、TCP 以下の層で付加されるヘッダデータは測定値に含まれない。ここで、映像データの圧縮には CBR を用いているが入力映像によって圧縮後のビットレートは若干変動するので、結果的に映像 1 本と音声 1 本を合わせて 2 本分のストリームに

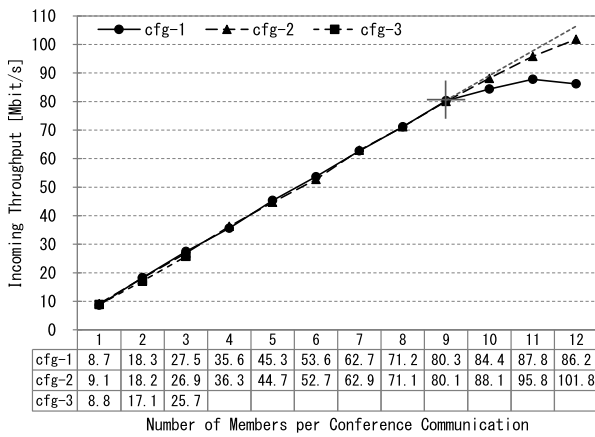


図 17 入力スループット (会議サーバ 1 台)

Fig. 17 Incoming throughput (1 conference server).

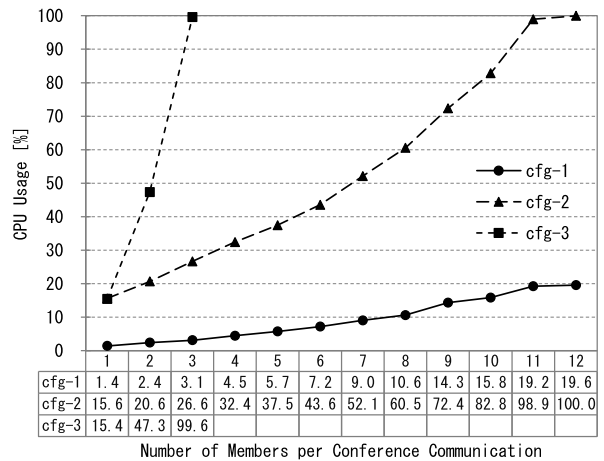


図 19 CPU 利用率 (会議サーバ 1 台)

Fig. 19 CPU usage (1 conference server).

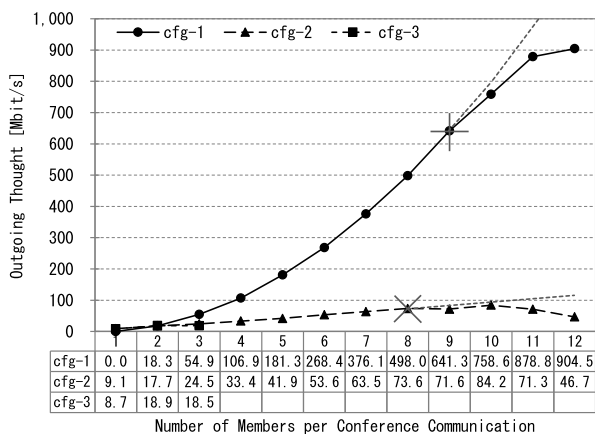


図 18 出力スループット (会議サーバ 1 台)

Fig. 18 Outgoing throughput (1 conference server).

必要とされるビットレートは約 9Mbps である。したがって、表 1 にまとめた各構成における入出力ストリーム数に 9Mbps を掛けた値が、スループットの計測値に対する比較の目安となる。

ここで表 1 より、会議通話に参加する利用者端末の数 (N) が 11 以上で、cfg-1 と cfg-3s (M shared terminals: $M = N$) における会議サーバのアプリケーションレベルの出力ビットレートは 1Gbps 以上必要となる。プロトタイプシステムの各端末の NIC およびネットワークスイッチには、普及している 1000Base-T を用いているので、会議通話に参加可能な利用者端末数の上限は明らかに 11 未満である。したがって、利用者端末 12 台で構成されている図 15 (a) の環境であれば 1 台の会議サーバにおける性能の上限を十分確認できるので、この実験でも図 15 (a) の環境を用いた。

会議サーバにおける入出力スループットと CPU 利用率の測定結果を図 17、図 18、図 19 のグラフにまとめる。これらの図の横軸は会議通話に参加している利用者端末数を示している。縦軸とグラフ下部の表は入出力スループット (図 17、図 18) または CPU 利用率 (図 19) の測定値

である。

cfg-1 の構成を利用する会議通話では、表 1 に示すとおり、会議サーバの入力スループットは利用者端末数 N に比例し、出力スループットは $N \cdot (N - 1)$ に比例する。実験では、利用者端末 9 台までは十分な入出力スループットが得られた。利用者端末 10 台目以降は、図 17 および図 18 において点線で示した理論値に対して十分なスループットが得られていない。一方、利用者端末 10 台の時点の CPU 利用率は図 19 から 15.8% であり、CPU 資源には余裕があるが、ネットワークの入出力処理がアプリケーションレベルの上限付近に達したと考えられる。結果として、プロトタイプシステムにおける cfg-1 の会議通話に参加可能な利用者端末の上限は 9 台となる。

cfg-2 (音声は cfg-3) の構成を用いた会議通話における会議サーバの入出力スループットは、音声ストリームに必要となるスループットに比べて映像ストリームに必要となるスループットが明らかに多く、表 1 の cfg-2 に示すとおり、ほぼ N に比例する。実験において、入力スループットは利用者端末 9 台まで十分に得られた。cfg-1 と同様に、利用者端末 10 台目以降は、図 17 において点線で示した理論値に対して十分なスループットが得られていない。また、図 18 より、出力スループットは利用者端末 8 台まで十分得られたが、9 台目以降は点線で示した理論値に対して十分でない。結果としては、cfg-2 (音声は cfg-3) の構成を利用する会議通話への参加可能な利用者端末数の上限は 8 台である。一方、図 19 より、利用者端末 9 台の時点での CPU 利用率は 72.4% と比較的高い値を計測しているが、まだ余裕がある。したがって、ミキシング処理およびその出力を各利用者端末へ送信する処理に改良の余地があると考えられる。

cfg-3 の構成を用いた会議通話は、利用者端末ごとに専用ミキサを利用するので会議サーバの負荷も高くなる。図 19 より、利用者端末数が 3 台になった時点で CPU 利用率は

99.6%であり、図 18 の出力スループットも、利用者端末数が 2 台から 3 台へ増加する際、増加しない。プロトタイプシステムにおける cfg-3 への参加可能な利用者端末数の上限は、720p の映像を用いた場合は 2 台となる。

ここまで、実験結果をもとにプロトタイプシステムにおける会議サーバ 1 台で会議通話可能な利用者端末数を確認した。結果として参加可能な利用者端末の数は、cfg-1 では 9 台、cfg-2 では 8 台、cfg-3 では 2 台である。cfg-1 と cfg-2 の構成を用いる場合は、既存の HD 対応型ビデオ会議専用端末を用いる場合と同程度の地点数における会議通話が、汎用 PC を用いたバックエンドシステムでも実現できることを確認した。一方、提案システムでは、ポータル会議サーバにより複数の会議サーバを用いることが可能なので、必要に応じて会議サーバを複数台用意すれば、同時に実施可能な会議通話数を増やすことも可能となる。

6.3 複数の共用端末による専用ミキサの機能分散

引き続き、共用端末を用いた専用ミキサの機能分散の実験結果をまとめる。cfg-3s (M shared terminal) の構成を利用する会議通話では、バックエンドに配置する共用端末の数を増やすことによって参加可能な利用者端末数の増やすことが可能となる。会議サーバ 1 台における cfg-2 (音声は cfg-3) の実験結果から、720p の共用ミキサを用いた場合、会議通話に参加可能な利用者端末数は 8 台であることが分かっている。したがって、会議サーバと同様の性能の共用端末を 8 台用意することにより、cfg-3s の会議通話へ参加できる利用者端末の数も 8 台まで増やすことができると期待できる。この台数効果を確認するため、図 15 (b) の実験環境で cfg-3s (M shared terminal) の構成を利用する会議通話を実施し、参加する利用者端末の数を 1~8 台へ増加させる実験を行った。図 15 (b) の実験環境は、会議情報データベースとポータル会議サーバの機能を内包する会議サーバ 1 台、共用端末と利用者端末がそれぞれ 8 台で構成されている。

また、利用するフォーマットによる台数効果の違いを確認するために 360p (ピクセル解像度: 1,280 × 720 pixel, フレームレート: 30 fps) のキャプチャ映像と WMV CBR 2 Mbps のストリームによる実験も行った。この実験では利用者端末 8 台に対して共用端末を 2 台用い、利用者端末ごとの専用ミキサが、2 台の共用端末で動作することを示す。

さらに、映像と音声のエンド間遅延も測定した。映像のエンド間遅延は、利用者端末においてビデオカメラからキャプチャしたビデオフレームのキャプチャ時刻に対して、そのビデオフレームが共用端末上でミキシング後に利用者端末で表示されるまでの時間を計測することで求めた。音声のエンド間遅延は、利用者端末のマイクロフォンのそばで 1kHz の音を 50ms 間発生させ、共用端末経由で戻ってきた同じ音が、利用者端末のスピーカから発せられるまで

の時間を計測した。IC レコーダを用いて 2 つの音を録音し、録音データから 1kHz 50ms の信号を抜き出し、その発音開始時刻の差を求めた。利用者端末の数を増加させるたびに、映像および音声とも計測を 10 回行った。

会議サーバにおける入出力スループットと共用端末の CPU 利用率の累計、および利用者端末におけるエンド間遅延の測定結果を 720p と 360p で分けてそれぞれ図 20、図 21、図 22 と図 23、図 24、図 25 に示す。各図の横軸は会議通話に参加した利用者端末数である。図 22 と図 25 に示すエンド間遅延は、計測値の平均を棒グラフで示し、その最大値と最小値もグラフに記載した。

この構成では、表 1 に示すとおり、会議サーバの入力スループットは 2N に比例し、出力スループットは $N \cdot (M + 1)$ に比例する。実験から、720p と 360p のそれぞれにおいて十分な入出力スループットが得られていたことを確認し、

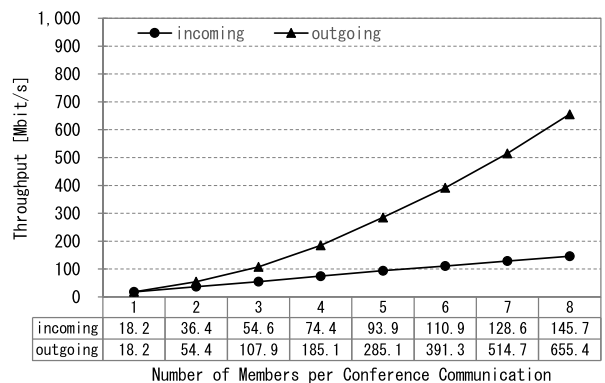


図 20 会議サーバの入出力スループット (720p)

Fig. 20 Incoming and outgoing throughput of conference server (720p).

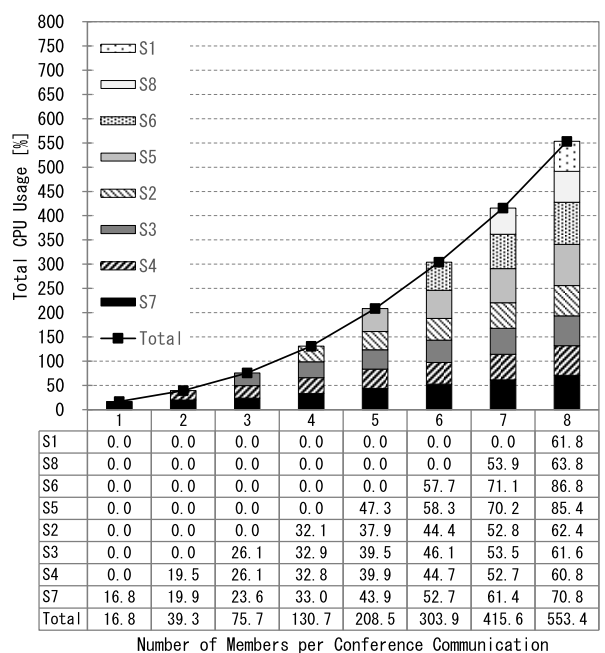


図 21 共用端末 8 台の CPU 利用率の累計 (720p)

Fig. 21 Total CPU usage of 8 shared terminals (720p).

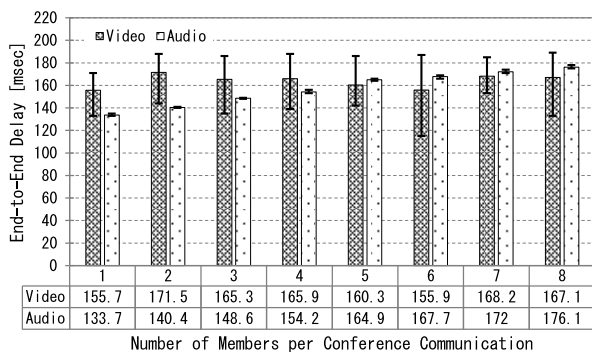


図 22 エンド間遅延 (720p)
Fig. 22 End-to-End delay (720p).

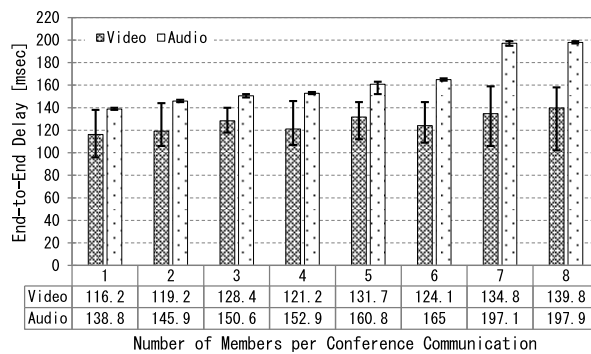


図 25 エンド間遅延 (360p)
Fig. 25 End-to-End delay (360p).

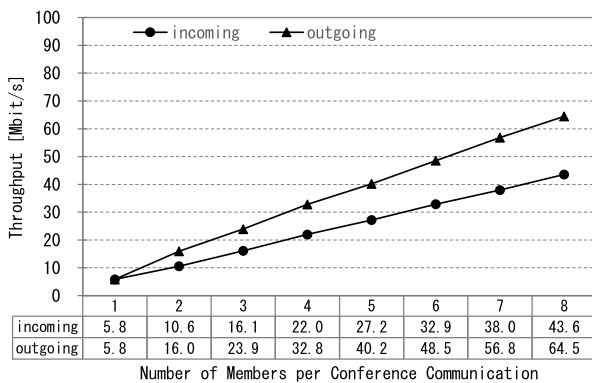


図 23 会議サーバの入出力スループット (360p)
Fig. 23 Incoming and outgoing throughput of conference server (360p).

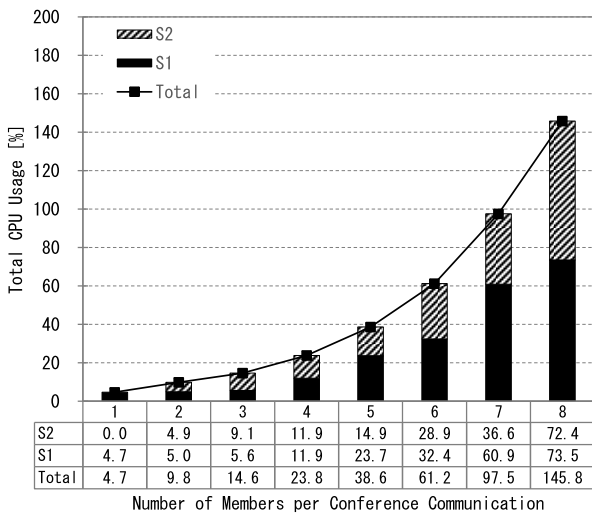


図 24 共用端末 2 台の CPU 利用率の累計 (360p)
Fig. 24 Total CPU usage of 2 shared terminals (360p).

利用者端末が会議通話へ参加する時点で CPU 利用率の低い共用端末が利用者端末へ割り当てられたことも確認できた。

映像ミキサでは、デコードされた映像フレーム (RGB データ) の矩形サイズ変換処理とアルファブレンド処理を行い、映像ミキサの出力用のメモリ領域へ処理された RGB データをコピーし、その後 WMV による圧縮処理を

している。今回の実験では、それぞれの専用ミキサにおいて、1つの入力映像を全体に表示させ、その他の映像は Picture-in-Picture として画面下部へ横に並べた。図 21 と図 24 は、共用端末の CPU 利用率の累計を示すとともに、会議通話に参加する利用者端末が増えた場合にそれぞれの共用端末が割り当てられたかも示している。ここで、CPU 資源の多くは圧縮処理に依存しているので、ミキシング結果の映像がフレーム内またはフレーム間の変化が少ない映像である場合、CPU 資源の消費は少なくなる。したがって、ミキサの処理内容によってはミキサ処理が割り当てられる共用端末も変わる可能性はある。現在のプロトタイプシステムでは、単純に CPU 利用率の低い共用端末に対してミキサ処理を割り当てており、ミキシング処理後の映像内容を考慮した共用端末の割当てに関しては、今後アルゴリズム検討の余地があると考えている。

また、エンド間遅延の測定結果より、映像と音声ともに 200 ms 以下の遅延 [22] で動作したことも確認した。図 22 および図 25 において、映像のエンド間遅延は音声のエンド間遅延に比べて最大値と最小値の差が大きいが、これは映像が 30 fps で処理されており、エンド間遅延の算出も 1/30 sec 単位で行ったためである。また、図 25 において、利用者端末数が 7 と 8 の場合の音声のエンド間遅延は高いが、これは 2 台の共用端末それぞれが、利用者端末 4 台分の専用ミキサを動作させ負荷の高い状態になったことが音声の遅延に影響していると考えられる。一方、リップシンクについては、ITU-R 勧告 BT.1359-1 [24] に記載されている許容範囲 (音が映像より先の場合 90 ms, 音が映像より遅れる場合 185 ms) を満たすことも確認できた。音が映像より遅れる時間は、音と映像の平均遅延値の差から、720p の実験において -31.1 ms (利用者端末数 2) ~ 11.8 ms (利用者端末数 6) であり、360p の実験では、22.2 ms (利用者端末数 3) ~ 62.3 ms (利用者端末数 7) であった。

実験を通し、cfg-3s (M shared terminals) の構成を利用する会議通話において、各利用者端末ではそれぞれ別の専用ミキサを利用できることが確認できた。720p の実験では、1 台の会議サーバを用いた場合に参加可能な利用者端

末が2台であったのに対し、共用端末の台数を増やすことで、その4倍の利用者端末間における会議通話が可能となった。また、360pの映像を用いた実験では、2台の共用端末で8台分の専用ミキサを利用できることが確認できた。

6.4 実際の通信イベントにおける運用事例

本プロトタイプシステムにより、これまで2回の通信イベントを支援したので、運用事例として概要をまとめる。

図 26 は、それらのネットワーク概要を示している。

2014年3月7日に実施された「第5回地域防災情報シンポジウム」[25], [26], [27]では、高知県立大学をメイン会場として岩手県立大学と静岡県立大学の3地点をJGN-X[28]とSINET4[29]でつないだ。図26におけるAPはJGN-Xのアクセスポイントを示している。岩手県立大学内では会場が2つに分かれており、実質4地点間の会議通話を実施した。岩手県立大学に一時的なバックエンドシステムを準備し、会議サーバ1台と共用端末を1台用いて、映像にはcfg-2sの構成で共用ミキサを、音声にはcfg-3s(1 shared terminal)の構成を用いて地点ごとの専用ミキサを適用した。映像には720pを用いたが、静岡県立大学側で使用帯域を抑える必要があったため、4Mbps(WMV CBR4Mbps)の映像ストリームを利用した。音声のキャプチャおよびストリームには評価実験と同じものを用いた。各会場ともに、持ち込み機材を中心としたイベント用の通信環境を用意す

る必要があったが、会場利用の都合により、機材のセットアップはイベント前日と当日に行われた。岩手県立大学側における前日(2014年3月6日)の準備は1時間で終了し、当日(2014年3月7日)、各会場間の通信テストや打ち合わせを含む3時間50分と本番4時間30分の間、システムを稼動させることができた。

一方、2014年5月28日に実施された「ICT推進フェア2014 in 東北」[30]では、仙台のメイン会場と高知工科大学および岩手県立大学の3地点を、JGN-Xを中心とするネットワークでつないだ。一時的なバックエンドシステムは先の利用事例同様、岩手県立大学に用意した。会議サーバ1台を用いて、映像にはcfg-2の構成を、音声にはcfg-3の構成を適用した。映像は720pを用いたが、メイン会場のせんだいメディアテークにおいて使用帯域を抑える必要があり、映像には3Mbps(WMV CBR4Mbps)のストリームを利用した。音声のキャプチャおよびストリームには評価実験と同じものを用いた。この通信イベントでも、各会場ともに持ち込み機材を中心としたイベント用の通信環境を用意する必要があったが、先の例と同様、セットアップはイベント前日と当日に行われた。前日(2014年5月27日)、岩手県立大学側におけるバックエンドシステムの準備は40分で終了した。前日は、メイン会場におけるネットワーク接続の調整に時間がかかったが、ネットワークレベルの接続が完了してからは、本システムを用いて事前の打ち合わせを行うことができた。また、当日(2014年5月28日)の準備に2時間、本番3時間50分の間、システムを稼動させることができた。

これらの通信イベントの準備と本番において、会議通話へ支障を来す通信トラブル等はなかった。しかし、特定の利用者端末における映像表示と音声再生の同期(リップシンク)に体感できるズレが生じることはあった。特定の利用者端末における音声再生モジュールで、レンダリングバッファ内に音声データが少しずつ蓄積されることが原因である。通信イベント実施時には、その利用者端末において会議通話中の無音期間にレンダリングバッファ内の音声データをフラッシュすることによりリップシンクのズレを回避することができた。一方、現在のプロトタイプシステムには、映像と音声の完全な同期機能は実装されておらず、遅延時間を増やさずにリップシンクをとる機能の実装は今後の課題と考えている。

これら2つの通信イベントを通して、本プロトタイプシステムを用いて一時的なバックエンドシステムをイベント用のネットワーク上に用意し、通信イベントを支援することが可能であることを確認した。

6.5 機能的な考察

通信イベントへの適用を想定し、機能評価実験と運用の結果をふまえて、提案システムと既存MCUを比較し、機

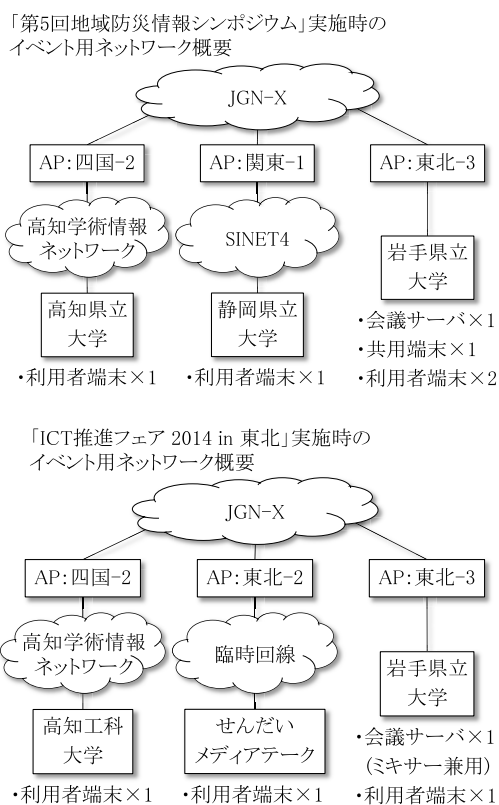


図 26 通信イベント実施時のネットワーク概要

Fig. 26 Network configurations at communication events.

能的な考察を以下にまとめる。

(1) HD 品質の会議通話機能

現在、複数のベンダから多くのビデオ会議専用システムが提供されている [31], [32], [33], [34]. ビデオ会議端末内蔵型の MCU の多くは 4 地点~9 地点間の通信を実現しており, 10 地点間 [5] を接続できる内蔵型の MCU もある. 一方, サーバ型 MCU を利用すると, そのシステム規模に依存するが数十地点間の通信も可能となる. 実装したプロトタイプシステムでは 1 台の会議サーバを用いて, cfg-2 の構成で 8 地点, cfg-1 の構成で 9 地点間の通信が可能であることを実験結果から示した. また, 必要となる帯域は利用するコーデックにも依存するが, プロトタイプシステムでは 720p の映像を 2Mbps 程度でも転送可能であり, 通信イベントのネットワーク環境に応じて調整することができる. 会議サーバ 1 台における HD 品質の会議通話接続数としては, 平均的な内蔵型 MCU と同等の地点数を汎用 PC で実現した.

(2) バックエンドにおける機能分散

通信イベントの実施にあたり, 複数の会議通話の同時実施や, 地点ごとに専用ミキサ機能を利用したい場合がある. 高機能なサーバ型 MCU には搭載されている機能であるが, 一時的に集めやすい機材を用いることを前提とすると, 高機能なサーバ型 MCU の利用は困難であることが多い. 一方, 汎用 PC を利用するという観点では, 既存のソフトウェア MCU の適用も考えられる. ソフトウェア MCU も複数のベンダから提供 [35], [36], [37] されているが, これらのソフトウェア MCU は, 日常的なビデオ会議の利用を前提としており, バックエンドで会議サーバを常時稼働させることを想定している. そのプラットフォームには, Intel Xeon シリーズ等, 主にサーバサイドで用いられる高性能な CPU が推奨されており, 専用機としてのサーバ型 MCU 同様, 一時的に準備するのは難しいことが想定される.

提案システムでは, 汎用 PC を用いた分散システムにより複数会議通話の実施に対応するとともに, 複数の共用端末を用いた分散型のミキサ機能も実現した. ミキサ機能のみを複数の共用端末で分散する仕組みにより, 利用するミキサの数に応じて機能を拡張できるバックエンドシステムが構築可能となる. サーバ型 MCU を利用しなくても, 汎用 PC を用いてこれらの機能を分散構成可能としたことは, 通信イベント実施時の一時的なバックエンドシステム構築をこれまでより容易にするものと考えている.

(3) 標準プロトコルへの対応

既存のビデオ会議システムにおける代表的な制御プロトコルは H.323 と SIP であり, 映像や音声データの送受信には RTP や HTTP が多く用いられている. 本システムのプロトタイプシステムは, 現在これらの標準プロトコルに準拠しておらず独自のプロトコルを用いている. したがって, 既存のビデオ会議システムとプロトタイプシステムの

表 3 提案システムと MCU の機能比較

Table 3 Comparison of proposed system and MCU systems.

	提案システム	端末内蔵型 MCU	サーバ型 MCU
HD品質の会議通話の接続数	会議サーバ1台で8/9地点	4~9地点が主流, 最大10地点	数10地点が可能
複数会議通話の同時実施	○	×	○
地点毎のミキサ機能	○	×	○
ミキサ機能の分散構成	○	×	×
標準プロトコル対応	×	○	○

互換はない. 一方, 通信イベントの実施にあたり, 既存のビデオ会議システムで利用されている標準プロトコルとの互換性があれば利用可能な端末の種類や利便性の向上にもつながるため, 標準プロトコルの導入は今後の課題となる.

(4) その他の機能

既存の MCU には, 会議のスケジューリング機能やストリーミング機能を搭載しているものもある. 通信イベントを支援するうえで, 複数の会議通話の開始と終了を円滑に進めるためにスケジューリング機能は有効であり, 通信イベント実施の様子を中継するためにはストリーミング機能が有効と考えられる. これらの機能については, 引き続き既存技術をふまえて実現を検討する. 一方, 本論文で提案したストリームパスのパターンを組み合わせることで, 1つの会議通話の中でも, 末端で利用可能な帯域幅に応じて会議通話地点ごとに適切な品質の映像と音声を用いた通信が可能となり, 利用者端末ごとに共用ミキサと専用ミキサを使い分けるといった会議通話も可能となる. 今後, 通信イベントにおける利用を通して, ストリームパスのパターンの組合せについても有用性を検証する.

以上をふまえ, ビデオ会議システムを提供している主要なベンダの既存 MCU と提案システムとの機能的な比較を表 3 に示す. 必要に応じて会議通話数や専用ミキサの機能を, 汎用 PC を用いて拡張的に利用できるのは提案システムのみとなる. 通信イベントの支援を想定し, そのイベントの主旨に応じて汎用 PC で機能を拡張できる適合型の会議通話システムを実現したと考えている.

7. まとめ

本論文では, 通信イベントにおける会議通話を支援するために必要となる通信システムの機能を整理し, 通信イベントの主旨に応じて一時的に構築可能な会議通話システムについて述べた. 通信イベントを支援する会議通話システムへの要件としては, (1) 制約のある通信イベント用のネットワーク上で会議通話システムを一時的に構築できること, (2) ネットワーク環境に応じて映像と音声の中継機能とミキシング機能を適応的に構成できること, (3) 同時に実施可能な会議通話数を増やせる機能と, 地点ごとのミ

キシング機能を、必要に応じて拡張的に利用できることがあげられる。これらの要件を満たすためには、既存のビデオ会議システムにおけるサーバ型 MCU の機能のうち、複数の会議通話を同時に実施する機能と地点ごとのミキシング機能を通信イベントごとに適合させる必要がある。そこで、中継機能とミキシング機能を必要に応じて複数の汎用 PC で分散させるためのプロトコルを設計・実装した。また、中継機能とミキシング機能を分散させるためには、適切なストリームパスを複数の端末で動的に構成する機能が必要となるので、会議通話で必要となるストリームの構成ごとにストリームモデルを用意し、それを選択的に組み合わせることができるよう会議通話システムを実装した。

また、評価実験を通して、既存の HD 対応型ビデオ会議専用端末を用いる場合と同程度の地点間における会議通話の実現できることを示し、同時に複数の会議通話を実施する機能と地点ごとのミキシング機能を、複数台の汎用 PC で分散処理することにより実現できることも示した。さらに、提案システムが実際に運用可能であることも 2 回の通信イベントの支援を通して確認できた。

今後は、プロトタイプシステムの実装における性能と機能の向上を図るとともに、実際の通信イベントでの利用を通し、より柔軟な運用を可能とする高機能な分散型会議通話システムの研究開発を進める。

謝辞 本提案システムの試用と運用にご協力いただいた、高知工科大学の福本昌弘教授ならびに静岡県立大学の湯瀬裕昭教授に感謝いたします。また、本研究の一部は、独立行政法人科学技術振興機構 A-STEP の研究課題（課題番号：AS242Z02573H）、および、独立行政法人情報通信研究機構のプロジェクト（「仮想化技術による大規模災害情報ネットワーク」、プロジェクト番号：JGNX-A11040）の一環として実施されました。ここに記して謝意を表します。

参考文献

[1] Polycom, Inc.: Best HD Video Conferencing Systems & Telepresence Services | Polycom, Inc., available from <http://www.polycom.com/products-services/hd-telepresence-video-conferencing.html> (accessed 2015-01-21).

[2] DVTS: DV (Digital Video) over IP [DVTS], available from <http://www.sfc.wide.ad.jp/DVTS> (accessed 2015-01-21).

[3] 熊谷誠治: ネットワーク仮想化技術と通信利用型放送: 3. ソフトウェアコーデックによる高品質 HD 映像の IP 伝送, 情報処理, Vol.50, No.11, pp.1124–1129 (2009).

[4] 八島由幸: 次世代画像圧縮技術研究の方向性, 情報処理, Vol.50, No.6, pp.548–557 (2009).

[5] パナソニック株式会社: ビデオ会議システム「HD 映像コミュニケーションシステム」の新モデルを発売 | プレスリリース | ニュース | パナソニック企業情報 | Panasonic, 入手先 <http://panasonic.co.jp/corp/news/official.data/data.dir/2014/07/jn140710-3/jn140710-3.html> (参照 2015-01-21).

[6] Schulzrinne, H., Casner, S., Frederick, R. and Jacobson,

V.: RTP: A Transport Protocol for Real-Time Applications, RFC3550 (2003).

[7] Ooi, W.T. and Renesse, V.R.: Distributing Media Transformation Over Multiple Media Gateways, *Proc. 9th ACM International Multimedia Conference* (2001).

[8] 橋本浩二, 柴田義孝: 利用者環境を考慮した相互通信のためのミドルウェア, 情報処理学会論文誌, Vol.46, No.2, pp.403–417 (2005).

[9] Bridges, P.G., Hiltunen, M. and Schlichting, R.D.: Cholla: A Framework for Composing and Coordinating Adaptations in Networked Systems, *IEEE Trans. Comput.*, Vol.58, No.11, pp.1456–1469 (2009).

[10] Nishida, H. and Nguyen, T.: Optimal Client-Server Assignment for Internet Distributed Systems, *IEEE Trans. Parallel and Distributed Systems*, Vol.24, No.3, pp.565–575 (2013).

[11] Cao, J., Hwang, K., Li, K. and Zomaya, A.Y.: Optimal Multiserver Configuration for Profit Maximization in Cloud Computing, *IEEE Trans. Parallel and Distributed Systems*, Vol.24, No.6, pp.1087–1096 (2013).

[12] Nallur, V. and Bahsoon, R.: A Decentralized Self-Adaptation Mechanism for Service-Based Applications in the Cloud, *IEEE Trans. Softw. Eng.*, Vol.39, No.5, pp.591–612 (2013).

[13] Zhiyuan, F., Wei, L., Zhang, F., Fang, Z., Huang, D. and Dai, X.: A Cloud-Based Pan-Terminal Video Conferencing System, *Proc. IEEE 10th International Conference on e-Business Engineering*, pp.445–449 (2013).

[14] Feng, Y., Li, B. and Li, B.: Airlift: Video conferencing as a cloud service using inter-datacenter networks, *Proc. IEEE 20th International Conference on Network Protocols*, pp.1–11 (2012).

[15] Feng, H. and Wu, W.: Framework and user migration strategy of cloud-based video conference multigateway system, *Proc. 19th International Conference on High Performance Computing*, pp.1–8 (2012).

[16] エヌ・ティ・ティ・ビズリンク株式会社: ニュースリリース | NTT ビズリンク, 入手先 (<http://www.vcd.nttbiz.com/search/newsrelease.php?id=58>) (参照 2015-01-21).

[17] Zhao, H., Smilkov, D., Dettori, P., Nogima, J., Schaffa, F.A., Westerink, P. and Wah, C.: A Feasibility Study of Collaborative Stream Routing in Peer-to-Peer Multiparty Video Conferencing, *Proc. IEEE International Symposium on Multimedia*, pp.233–240 (2011).

[18] Wang, Z., Zhao, J., Xi, W. and Jiang, Z.: A Scalable P2P Video Conferencing System Based on VCStream Model, *Proc. IEEE/ACIS 11th International Conference on Computer and Information Science*, pp.77–82 (2012).

[19] Guo, H., Lo, K.-T., Qian, Y. and Li, J.: Peer-to-Peer Live Video Distribution under Heterogeneous Bandwidth Constraints, *IEEE Trans. Parallel and Distributed Systems*, Vol.20, No.2, pp.233–245 (2009).

[20] Hashimoto, K. and Shibata, Y.: MidField: An Adaptive Middleware System for Multipoint Digital Video Communication, Digital Video, De Rango, F. (Ed.), ISBN: 978-953-7619-70-1, InTech, DOI: 10.5772/8032 (2010). available from <http://www.intechopen.com/books/digital-video/midfield-an-adaptive-middleware-system-for-multipoint-digital-video-communication> (accessed 2015-01-21).

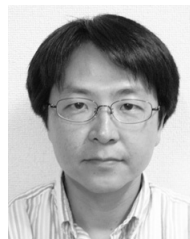
[21] MSDN Library: DirectShow, available from [https://msdn.microsoft.com/en-us/library/windows/desktop/dd375454\(v=vs.85\).aspx](https://msdn.microsoft.com/en-us/library/windows/desktop/dd375454(v=vs.85).aspx) (accessed 2015-01-21).

- [22] 玉木秀和, 東野 豪, 小林 稔, 井原雅行: 音声遅延が遠隔会議中の発話衝突と精神的ストレスに与える影響, 電子情報通信学会論文誌 D, Vol.J96-D, No.1, pp.35-45 (2013).
- [23] OpenLDAP Foundation: OpenLDAP, Main Page, available from (<http://www.openldap.org>) (accessed 2015-01-21).
- [24] RELATIVE TIMING OF SOUND AND VISION FOR BROADCASTING: RECOMMENDATION ITU-R BT.1359-1 (1998).
- [25] 四国情報通信懇談会:【部会】ICT 研究交流フォーラム 共催行事「第 5 回地域防災情報シンポジウム~災害時に有効な ICT~」の開催 (26.3.7 高知市)-四国情報通信懇談会, 入手先 (<http://shikoku-ict.jp/?p=4283>) (参照 2015-01-21).
- [26] 岩手日報 WebNews: いわたの防災, 入手先 (<http://www.iwate-np.co.jp/bousai/y2014/bousai1403081.html>) (参照 2015-01-21).
- [27] 静岡県立大学: 第 5 回地域防災情報シンポジウム | ICT イノベーション研究センター, 入手先 (<http://icti.u-shizuoka-ken.ac.jp/event/bjs2014>) (参照 2015-01-21).
- [28] JGN-X: JGN-X ウェブサイト, 入手先 (<http://www.jgn.nict.go.jp>) (参照 2015-01-21).
- [29] SINET4: 学術情報ネットワーク (SINET4, サイネット・フォー), 入手先 (<http://www.sinet.ad.jp>) (参照 2015-01-21).
- [30] 総務省: 総務省 | 東北総合通信局 | 「ICT 推進フェア 2014 in 東北」を開催, 入手先 (<http://www.soumu.go.jp/soutsu/tohoku/kohoshi/ICTsuisin260528.html>) (参照 2015-01-21).
- [31] Polycom: Video Conferencing, Voice Conferencing, Telepresence, available from (<http://www.polycom.com>) (accessed 2015-01-21).
- [32] Business Communications Solutions from Avaya - Avaya USA, available from (<http://www.avaya.com>) (accessed 2015-01-21).
- [33] 会議ソリューション - Cisco Systems, 入手先 (<http://www.cisco.com/web/JP/product/hs/webex/index.html>) (参照 2015-01-21).
- [34] ビデオ会議システム | 法人のお客様 | ソニー, 入手先 (<http://www.sony.jp/pes>) (参照 2015-01-21).
- [35] 製品ラインナップ・仕様 | ビデオ会議システム Visual Nexus | OKI, 入手先 (<https://www.oki.com/jp/visualnexus/product/spec.html>) (参照 2015-01-21).
- [36] 小規模から大規模までのビジネスに対応する HD ビデオ会議ソリューション | Lifesize, 入手先 (<http://www.lifesize.com/ja>) (参照 2015-01-21).
- [37] テレビ会議・ビデオ会議システム「HD コム」| Panasonic, 入手先 (<http://panasonic.biz/com/visual>) (参照 2015-01-21).



橋本 浩二 (正会員)

1996 年東洋大学大学院工学研究科電気工学専攻博士前期課程修了. 同年 (株) CSK 総合研究所入社. 1998 年岩手県立大学ソフトウェア情報学部助手. 2001 年東北大学大学院情報科学研究科博士後期課程修了. 2002 年岩手県立大学ソフトウェア情報学部講師. 2005 年同大学助教授 (2007 年より准教授), 現在に至る. 博士 (情報科学). 映像通信および分散システムの研究に従事. 電子情報通信学会, IEEE 各会員.



白崎 博生

1996 年大阪大学大学院基礎工学研究科物理系専攻博士前期課程修了. 同年 (株) インターネットイニシアティブ入社. 1996~1999 年 JPCERT コーディネーションセンター出向, 設立に従事. IJ サービスのバックエンドシステムの開発を経て, 2012 年より (株) ストラトスフィアに出向, SDN 技術の研究と開発に従事.



柴田 義孝 (正会員)

1985 年 UCLA 大学院コンピュータサイエンス学専攻博士課程修了. Ph.D. in Computer Science. 1985 年から 1988 年まで Bellcore (旧 AT&T ベル通信研究所) にて専任研究員としてマルチメディア情報ネットワークの研究に従事. 1989 年より東洋大学工学部情報工学科助教授. 1997 年同大学教授. 1998 年より岩手県立大学ソフトウェア情報学部教授, 同大学メディアセンター長. 2005 年より同大学大学院ソフトウェア情報学専攻主任, 2012 年より同大学理事副学長, 現在に至る. 電子情報通信学会, 感性工学会, IEEE, ACM 各会員.