

WWW 管理インターフェースを備えた IPv6 と SIP による集中制御型会議システムの開発

吉内 英也[†] 武田 幸子[†]
松木 譲介^{††} 星 徹^{†††}

ブロードバンドアクセス環境が一般家庭に普及し、ユーザが利用できる通信帯域は増加の一途をたどっている。このような背景を受け、IP ネットワーク上での音声・映像による双方向コミュニケーションシステムの開発、及びサービスの検討が進められている。この種のシステムの代表例としてテレビ会議システムがあるが、従来のテレビ会議システムでは会議室管理用インターフェースが専用ツールで提供されていたり、一般ユーザに公開されていない等の問題点があった。また、ユーザは指定の日に会議室にログインしなければならなかった。本研究では会議室管理インターフェースを WWW (World Wide Web) アプリケーションとして実装して一般ユーザ向けに公開し、サーバがユーザを会議室に召集する集中制御型の会議システムを提案する。集中制御型会議システムを実現するために、SIP(Session Initiation Protocol) の応用技術として標準化が進められている 3PCC (3rd Party Call Control) を用いる。提案手法のシステムを IPv6 ネットワーク上で構築し、接続時間の評価により有用性を示した。

Development of Center Controlled Conferencing System based on IPv6 and SIP with WWW Based Management Interface

HIDEYA YOSHIUCHI,[†] YUKIKO TAKEDA,[†] JOUSUKE MATSUKI^{††}
and TOHRU HOSHI^{†††}

Broadband access environment spreads over office and home, and people can get much faster network than ever used. According to this technological trends, bidirectional communication systems and services, especially video conferencing systems over IP network are rapidly improved. But video conferencing system has some problems. First, conference room management tools are provided as private ones or provided for only conference managers. Second, users must enter the conference rooms on the appointed time. In this paper, we propose center controlled conferencing system based on IPv6 and SIP with WWW (World Wide Web) based conference room management interface opened to end users. We use 3PCC (3rd Party Call Control) technology to invite users which is standardized as SIP applied technology. We implemented this system over IPv6 network and showed the usefulness by evaluated call control time.

1. 緒 論

一般家庭における IP (Internet Protocol) ネットワークへのブロードバンドアクセス環境が急速に普及し、ユーザの利用可能な通信帯域が増加したことを受け、VoIP による音声通話サービス、さらには映像を併用した双方向コミュニケーションシステムの検討が行われている。VoIP システムにおいては現在 IETF (Internet Engineering Task Force) で標準化が行われているアプリケーション層の呼制御プロトコルである SIP (Session Initiation Protocol)¹⁾²⁾ が急

速に広まりつつあり、ISP (Internet Service Provider) やサーバのベンダは、双方向コミュニケーションシステムの一つとして 3 人以上の多人数によるテレビ会議システムを付加価値サービスとして検討している。

テレビ会議システムのような多者通話における会話の場を”会議室”と呼ぶことにする。多者通話においては会議室の管理方法が問題となるが、従来の多者通話では管理者が専用のクライアント、又は WWW ブラウザ等から会議室を作成する方式が主流を占めていた。これを解決すべく、一般ユーザが会議室を作成できる方式も提案されている³⁾⁴⁾⁵⁾が、この方式では会議に参加するためにユーザが指定の日時に会議室に入室する必要があった。このため、ユーザは事前に会議室の識別子を URL (Uniform Resource Locators) 等の形で知らなければならず、また定例ミーティングのように日時が固定でタイムによる会議開催が可能な場合でも、ユーザからの能動的な参加を必要とするという問題がある。

[†] 日立製作所 中央研究所
Central Research Laboratory, Hitachi, Ltd.

^{††} 日立コミュニケーションテクノロジー
Hitachi Communication Technologies, Ltd.

^{†††} 東京工科大学 コンピュータサイエンス学部
School of Computer Science, Tokyo University of Technology

これらの問題を解決するため、会議室管理を WWW アプリケーションとして実装し、サーバ駆動で会議を開催できる多者通話システムを提案する。本研究では多者通話システムにおいてサーバ駆動でユーザを会議室に召集するために、3PCC(3rd Party Call Control)⁶⁾ という SIP の応用技術を適用した。また、システムの実装に際してはアドレス枯渇、NAT(Network Address Translation) による P2P(Peer-to-Peer) 型通信の阻害等の問題解決を視野に入れ、IPv6 を採用した。

以下、第 2 章では多者通話のモデルについて比較検討し、集中制御型多者通話モデルを採用した根拠を示す。第 3 章では集中制御型多者通話について詳説し、第 4 章で評価結果を示す。最後に第 5 章で結論と今後の課題を述べる。

2. SIP による多者通話モデル

多者通話モデルは交換機網におけるシステム以来さまざまな形式が提案されている。VoIP による多者通話においては、従来 ITU-T (International Telecommunication Union Telecommunication Standardization Section) で策定された H.323 が呼制御プロトコルとして採用されており、これに基づく会議モデルの提案、及び比較が行われていた⁷⁾ が、近年は IETF で標準化が進められている SIP が急速に広がりつつあり、SIP による多者通話モデルの議論が活発に行われている^{8) 9)}。本章では SIP について簡単に説明した後、多者通話モデルの比較を行う。

2.1 SIP

SIP はアプリケーション層のプロトコルであり、主に IP 上での音声、映像コミュニケーションにおいて通話セッションを確立し、制御することを目的としている。SIP は Request、Response の 2 種類のメッセージを備える。Request を受信した端末は Response を返す。SIP Response は HTTP の Response に準拠した 3 桁のステータスコードを持つ。SIP におけるセッション制御は、端末が Request を送信して要求を出し、相手端末、またはサーバがこれに対する Response を返すという形で進行する。

図 1 に 2 つの端末と SIP プロキシによる SIP の正常シーケンスを示す。通話を開始する端末 (端末 1) は通話開始用のリクエストである INVITE を中継サーバ (SIP プロキシ) 経由で相手端末 (端末 2) に送信する。INVITE には以下の情報が含まれている。

- 端末 1 が以降の通信で用いることを望むアドレス (Contact アドレス)
- 通話に用いるメディアの種類 (音声、映像、あるいは両方)
- 端末が処理可能なメディアの圧縮方式 (CODEC)
- 通話用の IP アドレス
- 通話用の通信ポート番号

これらの情報を「通話メディア情報」と呼ぶことにする。これらの情報のうち Contact アドレスは SIP ヘッダに、そ

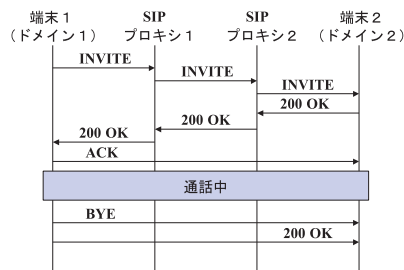


図 1 SIP 基本シーケンス

れ以外の情報は SDP (Session Description Protocol)¹⁰⁾ によりメッセージボディにそれぞれ格納される。

INVITE を受信した端末 2 は 200 応答 (OK) を返し、端末 1 からの通話要求に応じる。この 200 応答は端末 2 に関する通話メディア情報を含む。SIP の主目的は通話メディア情報を交換することにある。

2.2 多者通話モデル

ユーザの召集方法に着眼すると、多者通話は以下の 3 つのモデルに分類される。

- 端末ミキシングモデル
- ダイヤルインモデル
- ダイヤルアウトモデル

2.2.1 端末ミキシングモデル

会議サーバを用いずに端末間で全ての呼制御を行い、メディアのミキシングも端末側で行うモデル。端末ミキシングモデルではユーザが他のユーザを会議に召集する形で参加者が増加する。呼制御信号、メディアストリームは共に全ての端末間で送受信される。端末ミキシングモデルの図を 2(a) に示す。

2.2.2 ダイヤルインモデル

会議サーバにより呼制御を、メディアサーバによりメディア配信の集中制御を行うモデル。ダイヤルインモデルにおいては会議サーバがユーザからの要求を受けて呼制御用の会議室を準備し、各ユーザが会議室にログインすることで会議が形成される。ダイヤルインモデルの図を 2(b) に示す。

2.2.3 ダイヤルアウトモデル

会議サーバにより呼制御を、メディアサーバによりメディア配信の集中制御を行うモデル。ダイヤルアウトモデルに対してダイヤルアウトモデルでは会議サーバが会議の初期参加者リストを所有しており、会議サーバがユーザを会議室に召集する。各ユーザは会議サーバからの召集を受けてメディアサーバとの間にメディア通信用セッションを確立する。ダイヤルアウトモデルの図を 2(c) に示す。

2.3 会議モデルの比較

会議モデルの比較結果を表 1 に示す。

端末ミキシングモデルは会議室を作成する端末が他端末を会議に召集する手法であり、会議室の作成は可能だがそれと同時にユーザの召集も行うため、会議室の管理ができない。また、他のユーザが会議に途中参加する場合に、開

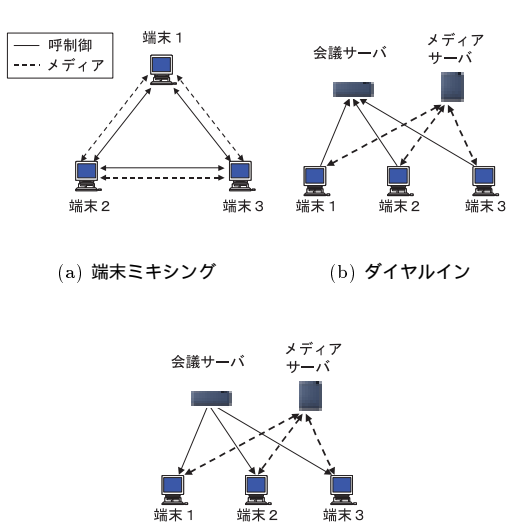


図 2 多者通話モデル

	端末 Mix	Dial-In	Dial-Out
会議室作成			
会議室管理	x		
ユーザ召集		x	
タイマ駆動	x	x	

催中の会議に関する情報をユーザに別途提供する必要がある。ダイヤルインモデルは会議サーバで会議室を管理するため、途中参加に関する問題は軽減されるが、端末が会議室に能動的にログインすることを必要とするため、タイマ駆動による会議召集や、会議に参加するべきユーザが不在の場合にユーザに参加を促したりすることができない。以上の結果を踏まえて、ダイヤルアウトモデルを採用した。

ユーザが会議室を作成する端末ミキシングモデルとは異なり、ダイヤルイン/アウトモデルでは会議室の作成・管理方法が問題となるが、会議室の予約、参加申込等を行う管理インターフェースを WWW アプリケーションとして実装し、これをユーザに公開することで利便性を確保した。

3. 集中制御型多者通話システム

前章で述べたダイヤルアウトモデルでは、会議サーバによる集中型の呼制御を行う。本章ではシステム構成について検討した結果を示す。

3.1 システム構成

集中制御型多者通話システムの構成を図 3 に示す。

システムは呼制御機能、会議室管理機能と、これらを管理する WWW インターフェースを備える会議サーバ、会議ソフトと音声、ビデオの入出力機構を有する PC からなる端末 (1~N)、端末からメディアを受信し、他端末へのメ

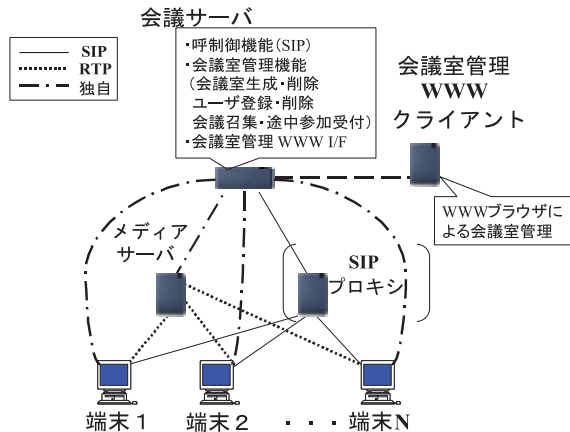


図 3 システム構成

ディア配信を行うメディアサーバと会議室管理 WWW クライアントからなる。会議サーバが提供する会議室管理機能を以下に列挙する。ユーザは会議室管理 WWW クライアントを通して会議室管理を行う。

- 会議室生成
- 会議室削除
- 会議室へのユーザ登録
- 会議室からのユーザ削除
- 会議召集
- 会議への途中参加

本システムにおける会議開催処理は以下の手順で進行する。

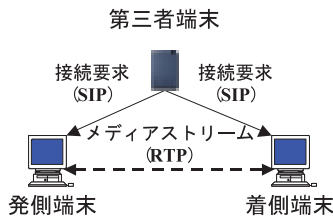
- (1) 会議を開催したいユーザ (主催者) が会議室を生成する
- (2) 主催者は会議に召集したいユーザを会議室に登録する、または会議に参加したいユーザは生成された会議室に自身を登録する
- (3) 主催者は WWW インターフェースから会議を召集して多者通話を開始させる。
- (4) 会議に途中から参加したいユーザは途中参加機能により会議に途中から入る。

会議の召集は主催者による召集の他にもタイマによる自動召集等が考えられるが、会議サーバからユーザ (主催者含む) に対する呼制御を行い、各端末とメディアサーバのストリームを確立するには、交換機網における網内発呼の機能が必要になる。この要求を満たす SIP の応用技術として、IETF で標準化が進められている 3PCC を適用する。

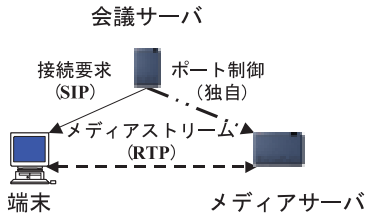
3.2 3PCC (3rd Party Call Control)

3PCC は第三者が二つの端末間にメディアストリームを確立する技術である。3PCC による呼制御では、INVITE をはじめとするメッセージは端末間ではなく、各端末と第三者端末との間で交換される。3PCC のレファレンスモデルの構成を図 4(a) に示す。

3PCC では第三者端末が 2 つの端末に対して呼制御サーバの役割を担うため、厳密には端末には発側・着側の区別



(a) レファレンスモデル



(b) 多者通話モデル

図 4 3PCC 構成要素

はないが、ここでは 第三者端末が最初に INVITE を送信する端末を発側、2 番目に INVITE を送信する端末を着側と呼ぶことにする。

3PCC においては呼制御メッセージの通信相手とメディアストリームの通信相手が異なる。SIP による接続要求は各端末と第三者端末の間で送信されるのに対し、音声、映像などのメディアストリームは RTP(Real-time Transport Protocol)¹¹⁾ により端末間で直接送受信が行われる。端末間で呼制御メッセージを交換する場合は通信相手に通話メディア情報を直接送信することができるが、3PCC では第三者端末が発側端末に着側端末の通話メディア情報を送信する必要がある。着側端末に関しても同様である。

このため、第三者端末は何らかの手段で発側・着側端末から通話メディア情報を収集する必要がある。通話メディア情報を収集するには (1) 事前にメディアの種類、CODEC、ポート番号を設定する (2) 呼制御セッション中に収集する、の 2 つの方法がある。

3.3 集中制御型多者通話への 3PCC の適用

集中制御型多者通話では会議サーバが第三者端末の役割を担い、端末とメディアサーバ間にメディアストリームを確立する。図 4(b) に集中制御型多者通話における 3PCC の構成を示す。

メディアサーバは通常の端末とは異なり、自身でメディアを送信することはなく、端末から受信したメディアに合成・複製などの処理を施して別の端末へ配信する。メディアを処理するには端末毎にポートを割り当てる必要がある。そこで、会議サーバはメディアストリームを確立する前にメディアサーバに端末を登録する。端末を登録するとメディア

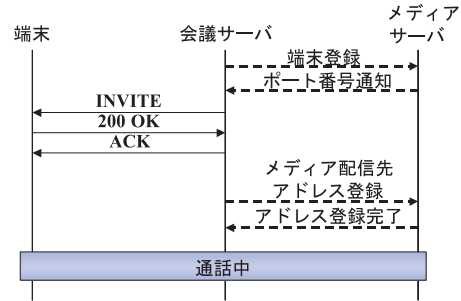


図 5 3PCC シーケンス

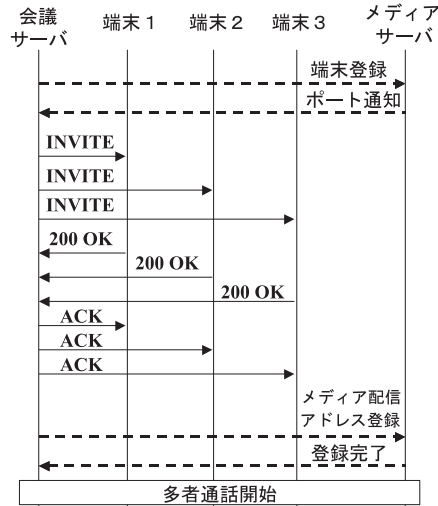


図 6 多者通話シーケンス (通話開始)

サーバはその端末用のポートを確保し、登録要求に対する応答でポートを要求元に通知する。ポート番号以外の通話メディア情報については、メディアの種類、CODEC はメディアサーバの仕様から、メディアサーバの IP アドレスについてはシステム構築時に設定することが可能である。

ポートの割り当ての他に、メディアサーバにはメディアの配信先アドレスを指定する必要がある。これらの処理を SIP で行うことは困難なため、会議サーバによるメディアサーバの制御は独自のプロトコル (以下、メディアサーバ制御プロトコル) で実装した。

端末へは前述のメディアサーバ制御プロトコルで収集したメディアサーバの通話メディア情報を SIP の INVITE メッセージを用いて送信する。端末からの 200 応答を受信した時点で会議サーバは端末の通話メディア情報を獲得する。最後に会議サーバは端末の通話メディア情報をメディアサーバに通知する。これらの処理シーケンスを図 5 に示す。また、多者通話開始のシーケンスを図 6 に示す。

会議サーバはメディアサーバに多者通話に参加する端末 1, 2, 3 をメディアサーバ制御プロトコルを用いて登録する。メディアサーバは登録要求を受け、各端末に対応するポート 1, 2, 3 をそれぞれ割り当て、要求に対する応答でこれを

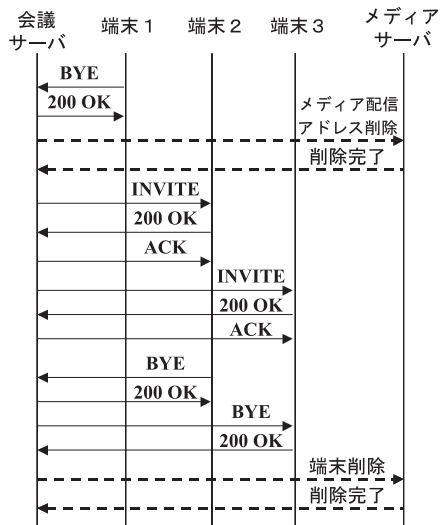


図 7 多者通話シーケンス (終了)

会議サーバに通知する。これは SIP における INVITE と 200 応答の交換に相当する。ポート番号以外の通話メディア情報は事前に会議サーバに設定することができるため、この時点で会議サーバは各端末に対するメディアサーバの通話メディア情報を獲得する。

次に会議サーバはメディアサーバの通話メディア情報を INVITE を用いて端末に送信する。通話メディア情報の記述は SDP によるオファー/アンサーモデル¹²⁾に従う。

端末は INVITE を受け、自身の通話メディア情報を 200 応答を含めて会議サーバに送信する。これらのシーケンスは端末毎に送受信が行われる。200 応答を受信した会議サーバは、SIP の規定に従い ACK を端末に送信して通話セッションが確立したことを端末に通知する。

通話セッション確立後に会議サーバはメディアサーバにメディアを配信するアドレスを登録する。各端末は自身に割り当てられたポートに対してメディアを送信し、メディアサーバは受信したメディアを送信元以外の端末に対して配信する。例えば、端末 1 はポート 1 に対して自身の音声・映像を送信し、メディアサーバはポート 1 で受信したメディアを端末 2、端末 3 に対して配信する。この処理を行うために、ポート 1 に端末 2、端末 3 のアドレスを登録する。ポート 2、ポート 3 についても同様である。以上の処理を経て多者通話が始まる。

次に終了シーケンスを図 7 に示す。終了処理は端末からの退出通知を受けて始まる。端末 1 が最初に退出するものとする。端末 1 は BYE を会議サーバに送信する。会議サーバは 200 応答で端末 1 に退出完了を通知すると同時に、メディアサーバに対して端末 1 へ配信していた全てのメディア配信を停止させるためにメディア配信アドレス削除を要求する。メディアサーバは端末 1 のアドレスをポート 2、ポート 3 から削除し、同時にポート 1 から端末 2、端末 3 へのメディア配信も停止させる。端末 1 が再度会議に

表 2 評価システム構成機器

		仕様
会議サーバ	Sun WS	Solaris8 SIP(RFC3261), メディアサーバ制御プロトコル
メディアサーバ	PC	会議室収容人数 4 名 Intel Pentium IV 1.7GHz Linux RTP(RFC1889) メディアサーバ制御プロトコル
端末	ノート PC	Intel Mobile Pentium IV 1.8GHz WindowsXP SIP, RTP

参加する可能性があるため、この時点ではメディアサーバから端末の削除は行わない。

不要なメディアの配信を停止させた後、会議サーバは端末 2、端末 3 に INVITE を送信して、各端末が端末 1 との通信用に解放していたポートを閉じるように要求する。これは、端末 1 との通信用のポートに対するメディア配信を停止させたことに伴う処置である。各端末は 200 応答を返し、メディアを受信しなくなったポートを閉じる。

次に、端末 2 が退出を要求し、会議サーバは 200 応答でこれを確認する。この時点で会議の参加者は端末 3 ののみとなり、二者通話も維持できない状態になる。そこで会議サーバは端末 3 に BYE を送信し、会議の終了を通知する。端末 3 からの 200 応答を受信した会議サーバは、メディアサーバに端末の削除を要求し、会議の終了をメディアサーバに対して通知する。以上で多者通話が終了する。

4. 実装・評価

本研究の有用性を示すために、提案手法の実装を行い、呼制御に必要な時間を測定した。システム構成は図 3 において SIP プロキシを要素に含めず、会議サーバが端末を直収する形とした。ネットワークは 100Base-T の LAN で構築し、システムの各要素には表 2 に示す機器を用いた。会議室制御 WWW インターフェースについては、会議サーバ上で WWW サーバを動作させ、Java Servlet で管理アプリケーションを実装した。

図 8(a) に会議室管理 WWW クライアントの画面を、図 8(b) に多者通話クライアントの画面をそれぞれ示す。

このシステム構成において 3 台の端末に対して WWW ブラウザから会議召集を行った。表 3 に接続時間を測定した結果を示す。表は WWW ブラウザにおいて召集ボタンを押下した時間を 0 として、各メッセージを会議サーバでキャプチャした相対時間を示す。

WWW ブラウザからの召集要求を受けて召集を開始するまで、即ち端末 1 に INVITE を送信するまでに 0.119 秒を要した。端末 1、2、3 のセッション確立に必要な時間はそれぞれ 0.667 秒、0.400 秒、0.228 秒であり、平均値は 0.431 秒であった。

