



高度なリアルタイムコミュニケーションを実現する 通信制御基盤

IP (Internet Protocol) ネットワーク上でのコミュニケーションサービス提供のためには、多様な参加端末と接続形態に応じた音声・映像のセッション制御の仕組みをサービスごとに開発する必要があり、コミュニケーションサービス提供者が本来尽力すべきコミュニケーションサービスのコンテンツ拡充にリソースを集中できない問題が起きています。本稿では、この問題を解決するため、通信キャリアが提供するさまざまなコミュニケーションサービスに適用可能な通信制御基盤を提供する研究開発の取り組みを紹介します。

キーワード：#リアルタイムコミュニケーション、#WebRTC、#標準化

コミュニケーションサービスの 高度化

IP (Internet Protocol) ネットワーク上でのコミュニケーションサービスでは、PCやスマートフォンなど多様なUE (User Equipment：ユーザ端末) が固定網や移動網等の異なる環境から接続し、音声や映像メディアをやり取りします。近年ではAI (人工知能) 解析エンジンやNPC (非プレイヤーキャラクタ) がメディア接続を介して感情などの非言語情報を収集・解析して、メタコミュニケーションを実現するユースケースが登場しています。このため、CP (Content Provider) は提供するコミュニケーションサービスごとにUEやネットワーク環境、メディアの違いを考慮したセッション制御方法を開発する必要があり、本来注力すべきコンテンツ拡充等にリソースを集中できない問題が起きています。

私たちはこの問題を解決するために、キャリア品質のコミュニケーションサービスに組込可能な柔軟性と品質を両立した通信制御基盤の研究開発に取り組んでいます。本稿では、メディア制御信号の標準化活動と、その制御信号に対応した通信制御基盤であるImmersive RTC (Real-Time Communication) 基盤を実現する技術を紹介し

キャリアWebRTC基盤の 実現仕様にかかわる標準化活動

新しいXR (Extended Reality) デバイスの登場や産業・教育・エンタテインメント分野におけるXRアプリケーションの利用拡大を受けて、通信キャリア・ベンダ各社は没入型インタラクティブな次世代リアルタイムコミュニケーションの重要な要素と位置付けています。これを受けてモバイル通信技術の国際標準仕様を定める3GPP (3rd Generation Partnership Project) においても、モバイル通信事業者が提供するメタバース・XRサービス向けの技術仕様の検討が行われています。

没入空間におけるリアルタイムコミュニケーションサービスを実現する方式としては大別して、既存の公衆IP電話のサーバ網であるIMS (IP Multimedia Subsystem) を拡張して提供する方式と、新たにWebRTCのサーバ網を導入して提供する方式が検討されており、NTTはWebRTC方式に注力し検討を主導してきました。これは仕様が厳格であり機能拡張による複雑化・安定運用が進むIMSに拡張を加えるよりも、Web開発との親和性や開発・導入のスピードを上げ、ユーザのニーズをいち早く取り込むために適した方式と考えられるためです。

WebRTC方式の標準化上重要な課題の1つとして、異ベンダ・キャリアのサーバ網間および端末デバイスとサーバ網間の接続性が保証されないことが挙げられます。

はら ゆうすけ^{†1} いじゅういん あきら^{†1}
原 佑輔 / 伊集院 明
やました やすはる^{†1} まえだ けんた^{†1}
山下 康治 / 前田 健太
すずき りひと^{†2}
鈴木 璃人

NTTネットワークイノベーションセンタ^{†1}

NTTネットワークサービスシステム研究所^{†2}

これは、インターネット系のプロトコルを策定するIETF (Internet Engineering Task Force) ではWebRTCのセッション制御のプロトコル仕様を規定していないためです。

このWebRTC方式の課題を解決するため、NTTは図1に示すアーキテクチャ (3GPP TS 26.506⁽¹⁾にて規定) を前提に、WebRTC向けのセッション制御信号プロトコルとしてRESPECT (REaltime&REality media Setup Protocol, Extensible and Compact), その他サービス上層機能の制御に必要なAPI (Application Programming Interface) 仕様 (サービス制御API) の検討を主導してきました。RESPECTはWebRTC標準に準拠するメディアセッション制御のプロトコルとして、高い信頼性とWeb開発との親和性を特徴とします。トランザクション管理やタイムアウトの機能、信頼性を検証できるID体系は通信キャリアでの利用に堪える信頼性の高いセッション制御を実現します。Web開発との親和性の観点では、WebSocketによるシグナリングパスを用いJSON形式のメッセージとすることで、既存のライブラリ・フレームワークを組み合わせた開発ならびにブラウザ機能を利用したデバッグ・トラブルシューティングの容易さと相互運用性の高さを実現しています。また、サービス制御APIは通信キャリアのWebRTC基盤を利用するCPがリソース管理を行うためのAPIで、メディア・データの転送制御を詳細に設定できることを特徴とします。

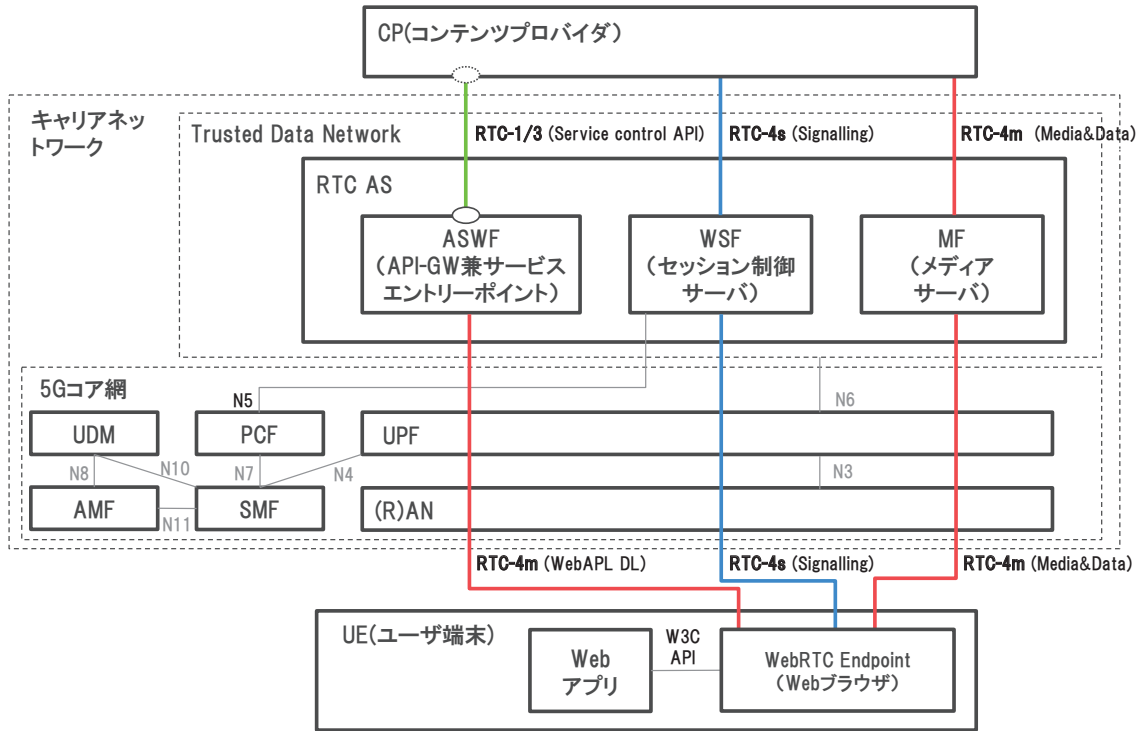


図1 RTCアーキテクチャ (コンテンツプロバイダ連携)

一般的なWebRTC通信サービスではサービスごとに固定的につくり込まれるさまざまな接続類型 (VR会議室, Webinar等) をAPI経由で設定することを可能とし、サービス提供の迅速化・柔軟化に貢献します。これらの3GPPリリース18における検討結果は技術文書3GPP TR 26.930⁽²⁾として合意・制定されています。

NTTの提供するImmersive RTC基盤は上記の3GPP標準をベースとして、アーキテクチャならびにインターフェースの設計・実装が行われています。

Immersive RTC基盤

私たちの取り組むImmersive RTC基盤の全体像を図2に示します。Immersive RTC基盤では、メディア通信のセッション制御を行うシグナリング制御部と、エンドユーザやCPのUEからのメディア通信を集約・配信するメディア処理部がUE間の音声・映像・データ等のメディア送受信を実現します。本稿では、これらのコンポーネントと、CPがシグナリング制御に必要な情報を操作するためのサービス制御API、そしてImmersive RTC基盤の安定稼働を支える監視と運用制御、高品質を

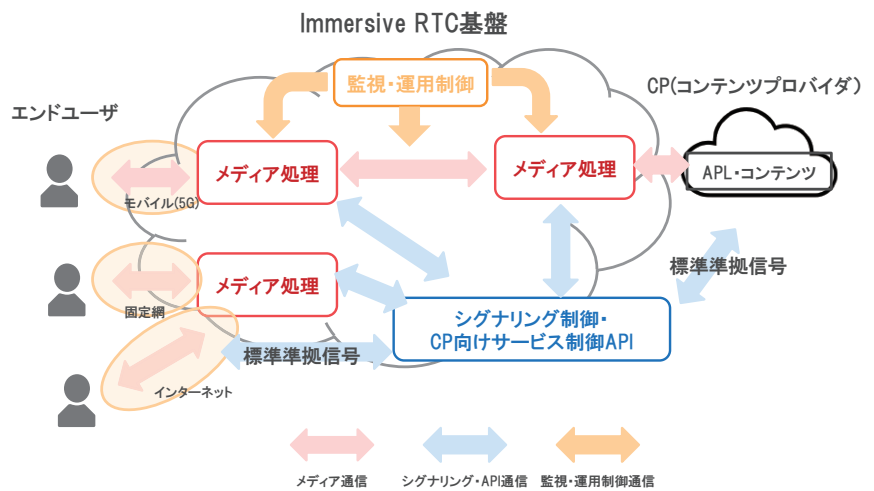


図2 Immersive RTC基盤全体像

現・維持するためのテスト・デプロイ戦略を紹介します。

Immersive RTC基盤：シグナリング制御

UEがメディア通信を行うためには、自身が送信するメディア情報を通信相手に伝えたくて、通信相手から受信するメディア情報を受け取る必要があります。Immersive RTC基盤ではUEとやり取りす

るRESPECT準拠の制御信号 (RESPECT信号) によりこれを実現します。

具体的には、まずUEからRESPECT信号を通してUEが送信するデータ・音声・映像のメディア情報を受け取り、その情報に基づいて、各メディアをどのメディア処理部のインスタンスを経由して、他のどのUEに届けるかを決定します。この決定に従い、メディア処理部に対してはUEとのメディア送受信を指示する信号を送信し、UEに対してはメディア処理部への接続情

報と他のUEから受信するメディア情報を含むRESPECT信号を送信します。

さらに、RESPECT信号によりUEのメディア送受信開始・停止の制御も行います。コミュニケーションサービスによってはUE間の接続順序関係が存在します。例えば、エンドユーザのUEには、そのエンドユーザの情報を扱うCP提供のUEを最初に接続させ、メディア送信を開始してから、次にエンドユーザのUE間でのメディア送受信を開始させるような順序関係が考えられます。Immersive RTC基盤ではUEの新規接続等のイベントをトリガーとしたCPへの通知機構との組合せにより、UEが適切にメディア送受信できるように制御します。

これらの制御を接続タイプごとに事前に定義しており、CPはImmersive RTC基盤に接続タイプを事前に登録するだけで提供するコミュニケーションサービスに必要な制御を行えます。現在は複数の主要な接続タイプに対応していますが、より多様なサービスに適用できるように、汎用的なモデル化に取り組んでいます。

Immersive RTC基盤：メディア処理

Immersive RTC基盤では、音声・映像データを配信する機能を持つSFU (Selective Forwarding Unit) や、メディアの変換や合成まで行う機能を持つMCU (Multipoint Control Unit) というメディア処理コンポーネントが中心となり、UEとの間でスター型の通信路を構成します。スター型は通信経路への負荷が低く、ユーザが増加した場合にもスケールしやすいメリットがあります。

SFUやMCUは、シグナリング制御部からの制御信号に従ってメディアを制御することにより、CPがコミュニケーションサービスを提供することを可能とします。提供するコミュニケーションサービスのユースケースとしては、下記の3種類の接続タイプを想定しています(図3)。

- ・エンドユーザが自分自身のアバターの映像を操作するような1:1のサービス
- ・エンドユーザへライブ映像を配信するような1:Nのサービス

- ・エンドユーザどうしや生成AIが会話する、エンドユーザから受け取った音声・映像を感情分析するようなN:Nのサービス

インターネット上でNAT (Network Address Translation) を介したメディア双方向通信を提供するにあたり、Immersive RTC基盤では、ICE (Interactive Connectivity Establishment) を利用します。ICEでは、STUN (Session Traversal Utilities for NAT) サーバを利用して、UEとSFU/MCUとの間で直接メディアセッションを確立します。

また、UEによるメディアの送受信制御がより容易にハンドリング可能となるようなクライアントライブラリも開発しています。CPはクライアントライブラリを利用することで、端末ブラウザのAPIに従ったクライアントサイドプログラムをフルスクラッチで開発するよりも開発負担を軽減す

ることができ、コンテンツ拡充等に注力することが可能です(図4)。

さらに、UEの受信能力に応じたメディア種別を選択して配信するサイマルキャストについても開発を進めています。サイマルキャストにより、PCやスマートフォン等の多様な端末に向けてのメディア配信が可能となるため、幅広いユーザ・利用シーンをターゲットとしたコミュニケーションサービスへの組込が可能となります。

Immersive RTC基盤：コンテンツプロバイダ向けAPI提供

シグナリング制御部やメディア処理部が動作するにあたり、コミュニケーションサービスの接続タイプのようなCPが提供するコミュニケーションサービス固有の情報が必要になります。Immersive RTC基盤では、CPがこれらの情報を登録・削除できる

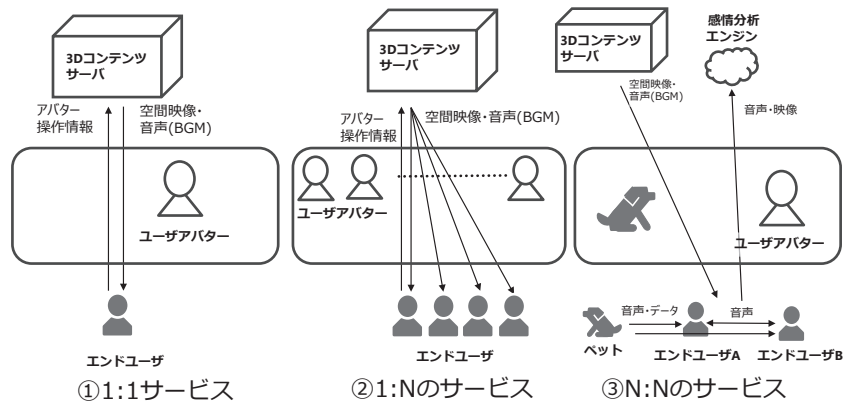


図3 コミュニケーションサービスのユースケース

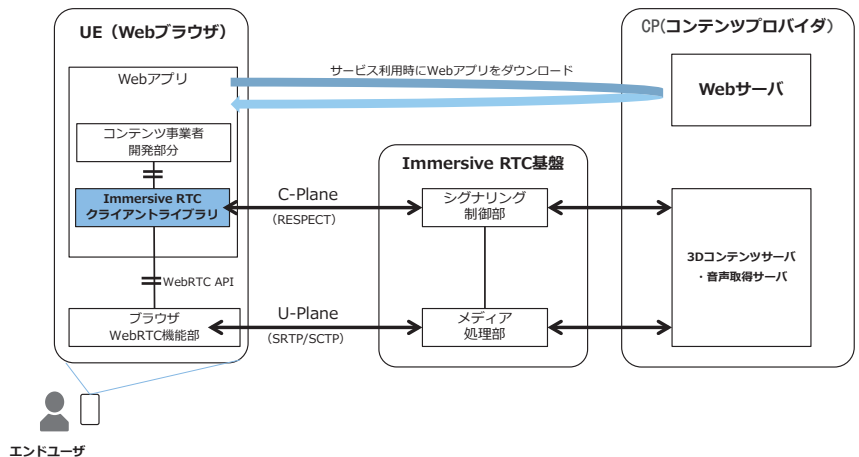


図4 クライアントライブラリ

REST (Representational State Transfer) 形式のサービス制御APIを提供します。

このAPIでは、外部認証基盤と連携したAPIキーなどのクレデンシャル情報による認証・認可機能を持っており、正しい権限を持ったCPのコミュニケーションサービス運用者による操作であることを検証しています。また、同時リクエスト数制限などの安定稼働を実現するためのセキュリティ機能を持っています。

Immersive RTC基盤：コンポーネント監視・運用制御

シグナリング制御部やメディア処理部が機能停止してしまうと、Immersive RTC基盤を組み込むコミュニケーションサービスの稼働に影響を与えてしまい、ひいてはエンドユーザのサービス体験にも影響を与えてしまいます。ここでは、Immersive RTC基盤を安定稼働させるための監視・運用制御について紹介します。

Immersive RTC基盤はパブリッククラウド上で動作しており、シグナリング制御・メディア処理部の各インスタンスの基本的な状態監視にはパブリッククラウドが提供する監視機能を使用しています。しかし、パブリッククラウドの機能ではパブリッククラウドサービスから見た動作状態しか監視・制御できないため、Immersive RTC基盤が正常に機能提供できる状態であることを保証できません。そのため、パブリッククラウドの機能では対応できない監視・運用制御を行うコンポーネントをImmersive RTC基盤では具備しています。具体的な監視方法として、シグナリング制御部とメディア処理部が連携するのと同じ機構を使用して、監視部から監視用信号を定期的送信し、監視対象のインスタンスから正常応答が返ってこなかった場合にエラー通知するアプリケーションレベルの監視により、機能を提供できない状態を早期に検知可能としています。

また、運用者がImmersive RTC基盤の各コンポーネントを制御可能とする機能の一例として、運用者がシグナリング制御・メディア処理部のインスタンスへ新規制御セッション確立を抑制、もしくは抑制解除を制御可能としています。本機能により、

当該インスタンスの制御セッションが0となった後にインスタンスを停止することで、コミュニケーションサービスへの影響なく各コンポーネントの減設が実施できます。

Immersive RTC基盤：テスト・デプロイ戦略

Immersive RTC基盤では、接続する端末や環境の変化に素早く対応するために、アジャイル開発を採用しています。現在のチームでは2週間のスプリントサイクルを採用しているため、各スプリントの成果物を速やかにデプロイし、品質評価を行うためにさまざまな工夫を取り入れています。ここでは、品質管理の観点で取り入れている3点の工夫を紹介します。

まず、既存機能に対するリグレッションテストについて説明します。リグレッションテストは、WebブラウザからのE2E (End-to-End) の自動試験によって行います。これにより、新しい機能を追加した際やコードの変更が行われた際に、既存の機能が正しく動作し続けることを保証します。具体的には、継続的インテグレーションツールと連携し、デプロイごとに自動テストを実行します。このプロセスにより、バグの早期発見と修正が可能となり、開発速度を維持しながら品質を保つことができます。

次に、新規リリースされる機能の評価についてです。新しい機能の評価計画は、開発スプリント中から着手することが重要です。スプリント期間中には、開発要件や試験計画を詳細に策定します。この段階で、ユーザーストーリーや受入れ基準を明確に定義し、テストケースを作成します。コンポーネント単位の単体テスト範囲を見定めたくうえで、統合テスト、E2Eテストのレベルで試験設計を行います。これにより、開発が進む中で段階的にテストを実施し、大きな手戻りがないよう品質確認を進めています。

さらに、継続的デプロイメントを実現するために、デプロイパイプラインを整備しています。各環境（開発、ステージング、本番）へのデプロイメントは、自動化されたパイプラインを通じて行われます。これにより、デプロイメントプロセスの一貫性が保たれ、手動の介入が減少し、デプロイメント速度が向上します。さらに、環境間

の差分もリポジトリで容易に管理ができます。これにより、環境差分による不具合発生の原因解明が早期に行えます。

以上の取り組みにより、2週間のスプリントサイクルで開発されたファイルを効率的にデプロイし、速やかな品質評価を進めています。これらのプロセスは継続的に改善を図り、常に最適な方法を模索し続けています。

多様なコミュニケーションサービス実現に向けて

Immersive RTC基盤を組み込めることとなるコミュニケーションサービスを広げていき、現在CPごとにクローズドになっているコミュニケーションサービスの領域において、オープンで新しい発想による多様なサービスの創出を促し、エンドユーザが豊富なCPによる多様かつ複合的なサービス享受できるように取り組んでいきます。

■参考文献

- (1) 3GPP TS 26.506: "5G Real-time Media Communication Architecture (Stage 2)."
- (2) 3GPP TR 26.930: "Study on the enhancement for Immersive Real-Time communication for WebRTC."



(上段左から) 原 佑輔/ 伊集院 明/
山下 康治

(下段左から) 前田 健太/ 鈴木 璃人

次世代のリアルタイムコミュニケーションは産業・教育・エンタテインメント等のさまざまな分野で革新をもたらす鍵になるととらえており、より良い世の中を実現するため取り組んでいきます。

◆問い合わせ先

NTTネットワークイノベーションセンター
ネットワーク制御ソフトウェアプロジェクト
E-mail irtc-dev@ntt.com