

Sistema de Videocolaboração interoperável entre webconferência e *endpoints* de videoconferência

Valter Roesler
Instituto de Informática, UFRGS –
Universidade Federal do RGS
Porto Alegre, RS, Brasil
roesler@inf.ufrgs.br

Mário Gasparoni Júnior
Instituto de Informática, UFRGS –
Universidade Federal do RGS
Porto Alegre, RS, Brasil
mariogasparoni@gmail.com

Felipe Cecagno
Mconf Tecnologia Ltda
Porto Alegre, RS, Brasil
felipe@mconf.com

Rafael Valle
RNP – Rede Nacional de Ensino e
Pesquisa
Rio de Janeiro, RJ, Brasil
rafael.valle@rnp.br

Ronaldo Husemann
Engenharia Elétrica, UFRGS -
Universidade Federal do RGS
Porto Alegre, RS, Brasil
ronaldo.husemann@ufrgs.br

ABSTRACT

The purpose of this paper is to present advances in the Mconf web conferencing system to support transparent interoperability with room videoconferencing equipment (known as endpoints), Multipresence system, and Phone@RNP. The system can behave as SFU (Switching Forwarding Unit) or MCU (Multipoint Control Unit), making decisions about the best signal to send to each destination. The following main benefits can be cited: 1) Improving user interaction in the use of such systems, so that it simply enters the portal “Video Collaboration Service”, and the system make the necessary adaptations to obtain the best user experience; 2) Have a single MCU + SFU service, savings resources, as it concentrates programmers on the same code.

KEYWORDS

Web conference, Videoconference, Remote Meetings.

1 Introdução

As soluções envolvendo videocolaboração na RNP são compostas principalmente pelos serviços de *Videoconferência*, *Conferência Web*, *Multipresença*, *Video@RNP* e *Fone@RNP* [1]. Tais serviços atingem milhares de usuários, entretanto, possuem os seguintes problemas (ou espaços para melhoria):

Problema 1: os serviços citados interagem entre si até certo ponto com limitações, como descrito nos próximos exemplos.

- **Exemplo 1:** durante uma reunião do serviço de videoconferência via MCU (*Multipoint Control Unit*) com n *endpoints*, é possível integrar uma sala de webconferência com m participantes. Porém, como o

modelo *endpoint* só envia um vídeo e recebe um vídeo, não é possível fazer com que todos os participantes da webconferência sejam vistos simultaneamente. Assim, todos os participantes da sala web são vistos em uma única janela de vídeo do *endpoint*, com comutação por voz (ver Figura 1a). No lado da webconferência, os participantes irão visualizar cada pessoa da sala individualmente (via SFU), porém todos os participantes vindos pela conexão com o MCU irão aparecer numa única janela (ver Figura 1b).



Figura 1: Problemas de interoperabilidade entre plataformas no serviço de Videocolaboração RNP

- **Exemplo 2:** no caso do sistema de Multipresença [2], existe um número de painéis (TVs, projetores ou monitores) que geram uma qualidade de experiência do usuário superior a um *endpoint* simples, entretanto, o sistema Multipresença se comporta exatamente como um *endpoint* quando se conecta a um MCU, utilizando uma só dessas várias telas. Haveria uma melhor experiência do usuário caso o sistema distribuisse esses vídeos, aproveitando todas as telas do painel.
- **Exemplo 3:** Ainda no caso do Multipresença, supondo que ele receba vários vídeos em Full HD separados, cria-se um problema de desperdício de banda para muitos vídeos, pois os mesmos estarão sendo mostrados em frações de tela, mas ocupando uma banda de rede alta. Como exemplo pode-se citar o caso onde uma das telas do Multipresença estivesse recebendo um fluxo de 1Mbit/s com 9 vídeos, ao invés de 9 fluxos de 1Mbit/s cada. Isso seria possível caso alguns vídeos fossem exibidos de forma composta (como um MCU) e outros em alta definição (apresentadores e conteúdo, por exemplo). Isso sem detrimento da experiência do usuário de visualizar quem está falando numa tela exclusiva, em alta definição.

Problema 2: cada solução possui seu próprio código-fonte e sua própria equipe, entretanto, alguns serviços poderiam compartilhar a mesma arquitetura e o mesmo código-fonte, otimizando equipes em termos de tempo de desenvolvimento, manutenção e suporte, e permitindo a unificação proposta.

Problema 3: os serviços Conferência Web e Videoconferência da RNP utilizam portais de acesso separados, ou seja, os usuários OU fazem uma conferência web (acessando sua sala pelo navegador web) OU fazem uma videoconferência via MCU (discando para uma sala pré-agendada). Entretanto, todos sinais são de vídeo, só que com protocolos diferentes. É necessária uma quebra desse paradigma, permitindo aos usuários se conectarem de forma transparente entre os serviços, possuindo uma sala única com um número de acesso a *endpoint*. Dessa forma, não haveriam dois portais de acesso, dois serviços, duas salas. A mesma sala e o mesmo serviço atenderiam as duas demandas, melhorando a experiência do usuário.

Problema 4: uma consequência do problema anterior (portais separados) é o tempo dispendido para a própria RNP na gerência de dois portais. O serviço de Videoconferência necessita agendamento de horário para uso da sala. Isso gera demandas administrativas para a RNP, que deve fazer esse controle, e demandas de tempo também para o usuário, que deve ter o trabalho de agendar a sala (o recurso de sala de videoconferência é escasso). Além disso, como o usuário hoje não tem poder de moderação sobre a sala criada (não é possível alterar o layout da sala de MCU, nem colocar um usuário em mute, por exemplo), gerando outro custo para a equipe de suporte da RNP, que é de atender as pessoas que telefonam para fazer esse tipo de

moderação na sala em andamento. Além de consumir tempo e recursos, não é prático.

A proposta deste artigo é apresentar um modelo que solucione todos os problemas apresentados.

2 Análise da realidade

Um conceito muito buscado por empresas é a *Comunicação Unificada*, ou *Unified Communication*¹, que tem o objetivo de unificar e simplificar o acesso a serviços de comunicação. Ao propor uma solução de comunicação unificada no lugar de webconferência, entrega-se maior valor com a solução, o que permite ao produto atender um público antes não atingido, como também proporcionar uma maior satisfação aos usuários². Como se trata de uma área de alto valor de mercado³, sendo transversal, ou seja, impactando educação, saúde, segurança, entre outros, existem empresas de grande porte trabalhando nisso. Alguns são os seguintes: Zoom, BlueJeans, Vidyo Connect e Microsoft Skype for business. Para analisar os diferentes sistemas, contratou-se uma sala ou se utilizou a versão gratuita dos mesmos.

O produto da Zoom (<https://zoom.us/>) oferece uma boa qualidade, e integra *endpoints* com o aplicativo da empresa. Alguns pontos negativos são: a) o aplicativo da Zoom necessita ser instalado na máquina, ou seja, não funciona no navegador; b) o modelo de negócios da Zoom prevê venda de licenças de um número de salas, necessitando agendamento dessas salas; c) a sala deve obrigatoriamente ser aberta no aplicativo, para então poder integrar com *endpoints* de hardware. Não existe o conceito de sala unificada como proposto neste artigo, onde pode haver uma reunião somente com *endpoints*; d) uma vantagem da Zoom é que suporta também H.323, enquanto a proposta neste artigo se propõe a integrar somente com o protocolo SIP.

O produto da Bluejeans (<https://www.bluejeans.com/>) também oferece boa qualidade e integra com *endpoints* de hardware, suportando SIP e H.323 TCP. Não suporta SIP / H.323 UDP. Ele já possui interface HTML5, melhorando a experiência do usuário, que não precisa instalar aplicativo. Para o *endpoint* ingressar, a sala deve ser primeiramente aberta no navegador (como dito no item anterior, a proposta do artigo melhora esse ponto com a sala unificada). O modelo de negócios prevê, da mesma forma que o Zoom, um número de salas para o cliente (e agendamento), o que é melhorado na proposta com o portal unificado e acesso via federação.

¹ <https://www.cisco.com/c/en/us/products/unified-communications/index.html>

² <https://www.bluejeans.com/press-releases/bluejeans-receives-2018-unified-communications-product-year-award>

³ <https://edition.cnn.com/2019/04/18/tech/zoom-ipo-video-conference/index.html>

O Vido Connect (<https://www.vidyo.com/>) oferece uma solução instalável ou via HTML5. Foi testada a versão HTML5, com ótima qualidade. Houve dificuldade para integrar com SIP / H.323, apesar do sistema dizer que suporta. Talvez precise de um módulo adicional. O modelo de negócios prevê um mínimo de 5 usuários com direito ao máximo de 3 salas por mês (de novo a questão do agendamento). Outras soluções de mercado seguem na linha dos produtos analisados.

3 Proposta de solução

A solução proposta segue na unificação / adaptação na arquitetura e código-fonte de três serviços / produtos, que são: 1) Conferência web; 2) Multipresença; 3) MCU.

A arquitetura da proposta se caracteriza por um sistema central que recebe áudios e vídeos (não importando se a origem é de um *endpoint*, webconferência, voz sobre IP ou Multipresença) e toma a decisão da melhor forma de enviar esse vídeo para o destino (se de forma composta como MCU ou de forma individualizada tipo SFU, ou mesmo um misto de ambos para otimizar banda).

A Figura 2 apresenta a arquitetura unificada proposta. No lado esquerdo (participantes) pode-se ver a forma que o sistema se adapta a sistemas heterogêneos, ou seja, através de gateways que convertem um determinado protocolo (SIP, WebRTC, etc) no padrão definido pelo grupo (API unificada). Esse modelo permite vários tipos de expansão. Por exemplo, se for necessário integrar com RTSP, basta criar um gateway RTSP. O mesmo vale para protocolos novos que ainda não foram criados.

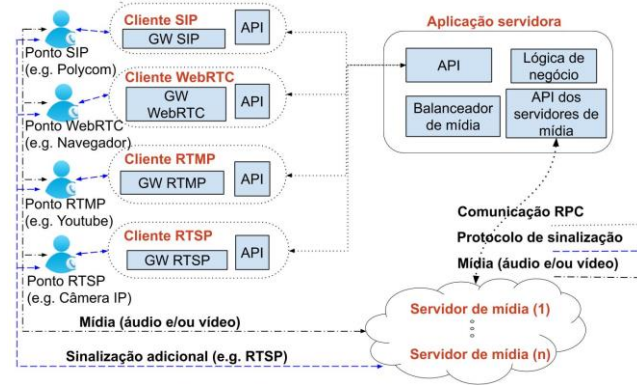


Figura 2: Arquitetura de solução proposta

O lado direito superior (*Aplicação servidora*) mostra o núcleo principal do sistema, onde as decisões são tomadas. Ali existem os blocos de *API* (interface com a API unificada), *Lógica de negócio* (inteligência do sistema), *Balancedor* (permite escalabilidade quando uma máquina virtual atinge seu limite), e *Interface com os servidores de mídia* (que lidam com o áudio e vídeo de uma determinada sala de conferência).

Resumidamente, existe um bloco de sinalização, que permite a homogeneização a protocolos heterogêneos, um bloco servidor, onde está a inteligência do sistema, e um bloco de áudio e vídeo

(codificação, decodificação, envio e recebimento), que interage com os participantes em cada sala virtual de videoconferência. Um ponto chave da unificação é a interface de comunicação (*API*) entre as diferentes soluções, permitindo o compartilhamento de mídias mesmo que provenientes de fontes distintas (*endpoint* de hardware, aplicação web ou smartphone). Com a arquitetura proposta a experiência do usuário é bastante melhorada, resolvendo os problemas apresentados na seção 1. A Figura 3 apresenta um cenário integrando os três serviços numa única sala virtual. Cada participante possui uma cor distinta, e é visto em todas as outras interfaces. A figura representa também o envio de conteúdo, que no *endpoint* tradicional e no Multipresença é visualizado em uma TV separada. No dispositivo móvel e no desktop é visto na mesma interface. Observe que os quatro participantes do *endpoint* tradicional estão recebendo os vídeos (TV da esquerda) de forma composta, e o sistema faz o papel de MCU nesse caso. Já no Multipresença a TV de cima à esquerda pode receber os vídeos no modo MCU (para economizar banda) ou separadamente (modo SFU), conforme decisão do servidor do GT-videocolaboração. Para os dois participantes Web, os vídeos são enviados separadamente (modo SFU) nesse exemplo, mas é uma decisão do servidor.

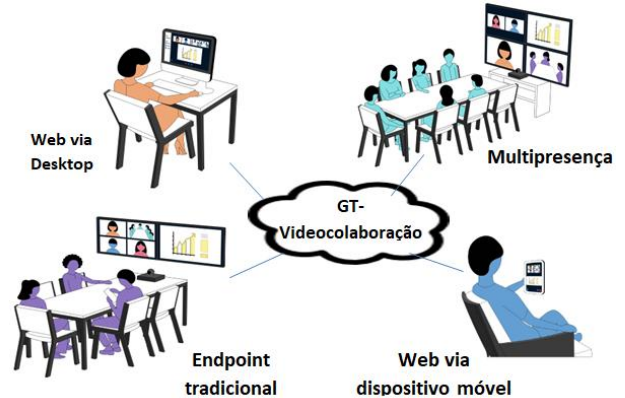


Figura 3: Cenário integrando os serviços propostos

O elemento central do projeto toma decisões para melhorar a experiência do usuário. Como esse elemento efetua funções de MCU e SFU, recebendo todos os vídeos, ele pode tomar uma decisão de enviar para os *endpoints* o único vídeo composto todos os outros vídeos (inclusive os da webconferência), pode enviar todos os vídeos separadamente para a webconferência, e pode enviar de forma mista para o Multipresença, resolvendo a questão de banda indicada na seção anterior.

Outro diferencial da proposta é o portal unificado (problemas 3 e 4 da seção 1), que permite milhares de salas (uma ou mais por pessoa na organização), eliminando o agendamento e melhorando a experiência do usuário, que utiliza o mesmo portal tanto para Conferência Web como Videoconferência. Dessa forma, os custos administrativos de agendamento e moderação de sala seriam reduzidos ou eliminados. A sala de Conferência

Web tem um número de videoconferência associado, que permite abrir a sala tanto via portal web como discando para o número via *endpoint*.

Adicionalmente às vantagens apresentadas, o sistema também se integra com serviços de VoIP SIP, e também com serviços de streaming RTMP.

4 Resultados

O ambiente de validação foi implementado em uma VM (*Virtual Machine*) no ambiente da Universidade. A Figura 4 mostra uma reunião efetuada com 4 pessoas, sendo 3 via navegador web e 1 através do software *Real Presence Desktop*, da Polycom. A transmissão de conteúdo está sendo feita a partir do Polycom através do protocolo BFCP (*Binary Floor Control Protocol*). Observa-se que a tela do computador onde está o Polycom está sendo compartilhada nos usuários da web.

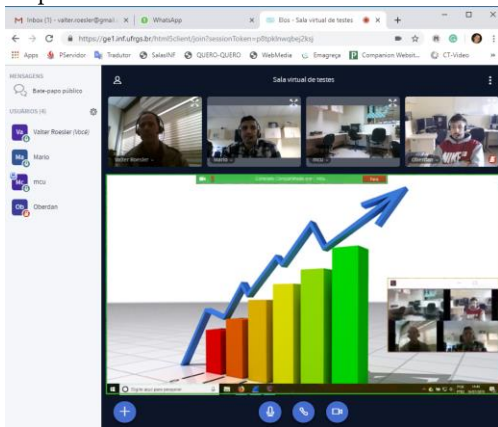


Figura 4: Exemplo de funcionamento híbrido (visão web)

O sistema desenvolvido permite a integração de forma transparente para o usuário, que percebe a mesma sensação para todos os participantes, independente do dispositivo pelo qual ingressaram na reunião. O conteúdo, que é a tela da máquina, aparece numa janela destacada no navegador web.

A Figura 5 mostra a visão do lado do sistema de videoconferência, que tem por característica enviar um vídeo e receber um vídeo. Pode-se perceber que o sistema desenvolvido compõe os quatro vídeos em um só, fazendo o papel de um MCU. Observa-se também que todos os vídeos dos participantes da web são mostrados, e não existe a comutação por voz (que foi o problema citado no início do artigo). Do ponto de vista do usuário, a experiência é similar à utilização de um MCU. O conteúdo (tela da máquina com apresentação) é mostrado ao fundo. O equipamento consegue comutar entre vários layouts, sendo um deles o vídeo das pessoas em tela cheia e a apresentação em miniatura.



Figura 5: Exemplo de funcionamento híbrido (visão Polycom)

A reunião poderia receber outros usuários de outras plataformas, como o sistema Multipresença ou sistemas de VoIP. No caso de VoIP, o usuário teria somente áudio. Na videoconferência, está se estudando uma forma de indicar sua presença, mas a mesma ainda não está implementada.

5 Visão de Futuro na área

No curto prazo, uma tendência é permitir aos usuários uma experiência única entre diferentes dispositivos, como equipamentos de hardware, navegador web, dispositivos móveis e sistemas com múltiplas telas. Além disso, essa experiência deve ser simples, ou seja, o mais transparente possível. Detalhes técnicos devem ser minimizados para o usuário, permitindo que pessoas de qualquer área do conhecimento consigam participar a partir de qualquer dispositivo.

Em médio e longo prazo estima-se uma evolução de formas alternativas de videoconferência, mais imersivas. Dentre essas formas pode-se citar a realidade aumentada e a holografia^{1, 2}.

AGRADECIMENTOS

Este trabalho está sendo executado na UFRGS e Mconf Tecnologia, sendo financiado pela RNP e pela Mconf Tecnologia Ltda.

REFERÊNCIAS

- [1] Ciuffo, L. et al. **Mapeamento de soluções de video colaboração da RNP**. Em "O futuro da video colaboração: perspectivas", capítulo 11. Editora SBC (Sociedade Brasileira de Computação). 368p. 2017. Disponível em http://www.inf.ufrgs.br/webmedia2017/wp-content/livro_worksho_ctvideo.pdf
- [2] ROESLER, Valtter; COELHO, L.; LONGONI, G.; MARINS, A.; DARONCO, L.; CIUFFO, L.; DUARTE, R. . Multipresença: towards videoconference and collaboration in multi-use environments. In: TNC16 Networking Conference, 2016, Praga. 32th TNC Networking Conference, 2016.

¹ <https://www.youtube.com/watch?v=7d5906cfaM0>

² <https://www.youtube.com/watch?v=1cQbMP3I5Sk>