

# Technical Disclosure Commons

---

Defensive Publications Series

---

September 03, 2019

## Discreet Dialogues During a Conference Using Stereo Audio Mixing Capabilities

Jeremy ZRIHEN

*ALE International*

Follow this and additional works at: [https://www.tdcommons.org/dpubs\\_series](https://www.tdcommons.org/dpubs_series)

---

### Recommended Citation

ZRIHEN, Jeremy, "Discreet Dialogues During a Conference Using Stereo Audio Mixing Capabilities", Technical Disclosure Commons, (September 03, 2019)

[https://www.tdcommons.org/dpubs\\_series/2454](https://www.tdcommons.org/dpubs_series/2454)



This work is licensed under a [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).

This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.

<b>Docket Number</b>	FR82019001
<b>Title</b>	“Discreet dialogues during a conference using stereo audio mixing capabilities”
<b>Contributor</b>	Jerremy ZRIHEN
<b>Company</b>	ALE International

## Description of the technical solution:

### A. INTRODUCTION

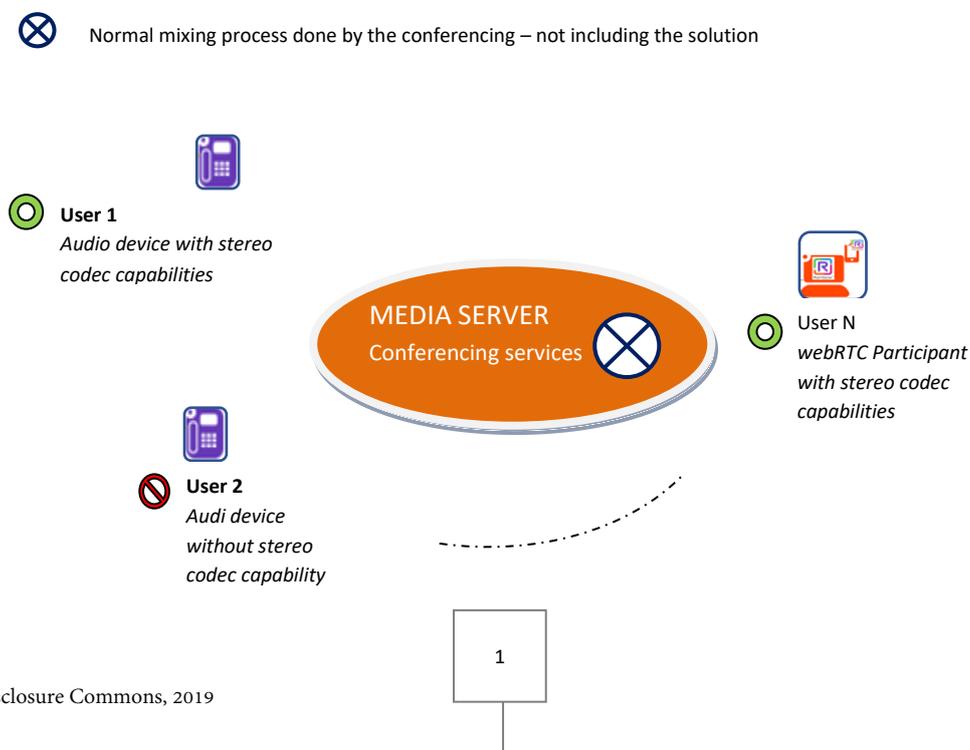
During a conference, when some events occurred, the emission of an audio dialogues can appear. For example, we can often hear “User X joins the conference” to inform all participants of a conference when a new user joins the conference ... or “User X leaves the conference” to inform all participants of a conference when a new user leaves the conference; but also, many others kind information.

⇒ ***The issue in such circumstance, is that the audio of the conference is cut during the emission of the dialog. This can be very disruptive.***

When conferencing system implements this solution, such solution is available for all participants with an audio device with stereo capabilities. They can hear the dialogue and the audio of the conference at the time and without any disturbance.

Let assume that N users participate in a conference using conferencing system with stereo capabilities and implementing this solution.

Each user using audio device with stereo codec capabilities will benefits on this solution.

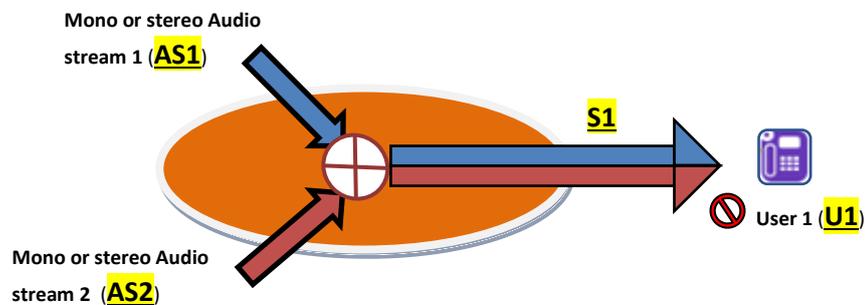


**B. THE SOLUTION**

Core of this solution is to use stereo capabilities of a conferencing system to give to end users the ability to receive several audio streams at the same time, and therefore to be able to distinguish them by using spatialization effects.

For example, in the scheme below, we can see that a mixing stereo output **S1** for user **U1** can contain a spatial rendering of the two audio inputs **AS1** and **AS2**. Independently with input streams mode (mono or stereo), it is always possible for mixing processes to render the two audio streams information in a single stereo output **S1** using spatialization effects.

⊕ Supplementary mixing process and audio rendering proposed by the solution



This stereo mixing process, that consists on the mix of two input audio streams to generate a single stereo stream with a spatial separation of the two inputs, needs to be distinguished with the classical mixing process done by a conferencing server and that acts in the way to render an audio containing the addition of the main actors of the conference (speakers with the higher level).

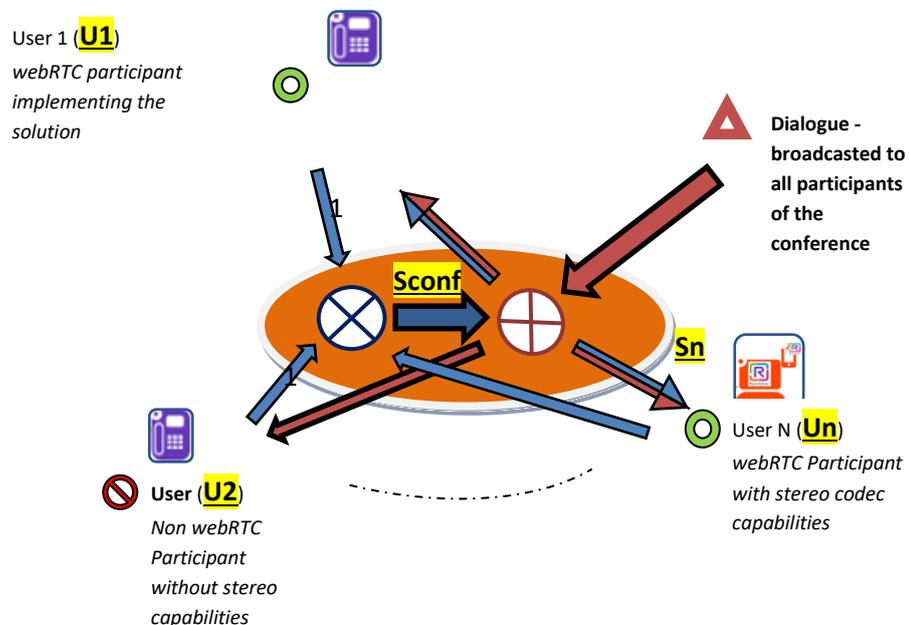
The idea to add such mixing process in the way to remove perturbation of dialogues in a conferencing server is the main purpose of this solution.

For that, the solution proposes to add a stereo mixing capability in the conferencing server.

During the emission of a dialogue, the mixing processes done by the conferencing server for a user (**Un**) with stereo codec capabilities will produce one stereo audio stream. This stereo audio stream (**Sn**) is the spatialized restitution of the audio mixing process already done by the conferencing server in a normal case (**Sconf**) plus the audio of the dialogue.

- ⇒ In example below, User (**U2**) does not have stereo capabilities and so, only ear the dialogue.
- ⇒ Other participants having stereo capabilities can listen, the dialog without any cut of the audio of the conference.

-  Normal mixing process done by the conferencing – not including the solution
-  Supplementary mixing process and audio rendering proposed by the solution



It is to consider that output of the main conference (**Sconf**) is composed of several independent audio streams. Depending on the mixing process of the main conference, several participants can be declared active talkers, and be part of distinct audio; each participant declared as active receives a specific audio stream without his own voice, and all other participant declared as non-active received mixing result without any removal.

- ⇒ The spatialized audio restitution (**Sn**) done for each participant having stereo capabilities such as (**Un**), includes the audio of the dialogue plus the (**Sconf**) stream part that is dedicated to this participant, depending on his active / non-active talker status in the main conference.

### C. **IMPLEMENTATION**

This section describes in details the implementation that can be proposed for a conferencing system (conferencing server + end user audio devices) to fully benefit from the solution.

First, conferencing server must implement mixing processes, exposed in the previous section, in the way to remove disruption when dialogues need to be transmitted to participant of the conference.

Optionally, conferencing server and audio communication devices implementing the solution can have capabilities to exchange some specific messages and interact with. In many systems, a protocol is already existing to initiate media communication between end user audio devices and conferencing system, and allows to transport some information to interact with. SIP is for example the most popular one and offer mandatory facilities for the solution.

Those facilities will be used to allow some new interactions between each audio communication devices implementing the solution and the conferencing system.

A Man Machine Interface (MMI) on audio communication devices implementing the solution allows end users to benefit from the solution giving the possibility to trigger following additional use cases and scenarios;

- **UC3: MODIFY AUDIO LEVEL MAIN CONF:**
  - *Modify level of the main conference audio stream*
- **UC4: MODIFY AUDIO LEVEL DIALOGUE:**
  - *Modify level of a private communication audio stream*
- **UC5: MODIFY POSITION MAIN CONF:**
  - *Modify spatial position of the main conference audio stream*
- **UC6: MODIFY POSITION DIALOGUE:**
  - *Modify spatial position of a private communication audio stream*

An implementation of this solution can consist on following scenarios / sequence diagram:

### **MANAGE AUDIO RENDERING**

Following use cases are commands send from end user audio communication device to conferencing server.

Those commands allow to manage audio rendering (level and spatial position) of the main conference and of dialogues when appeared.

#### ***MODIFY\_AUDIO\_LEVEL\_MAIN\_CONF:***

Modify level of the main conference audio stream.

Parameters of the command are following:

Required audio level

#### ***MODIFY\_AUDIO\_LEVEL\_DIALOGUE:***

Modify level of dialogues audio stream.

Parameters of the command are following:

Required audio level

#### ***MODIFY\_POSITION\_MAIN\_CONF:***

Modify spatial position of the main conference audio stream

Parameters of the command are following:

Required position – xyz spatial adaptation to do from current position

#### ***MODIFY\_POSITION\_PRIVATE\_COM:***

Modify spatial position of dialogues audio stream

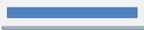
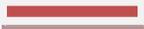
Parameters of the command are following:

Required position – xyz spatial adaptation to do from current position

They are implemented only for end users with stereo audio capabilities during the conference.

When received by conferencing server, mixing processes is adapted to consider new required values.

Sequence diagram below illustrate commands where acknowledgement of commands is optional:

-  End user interacts only with main conference
-  End user participates in private discussion. He can also ear main conference discussion.

