Sound Level Indicators in RTP

draft-ivov-avt-slic-00
draft-lennox-avt-rtp-audio-level-exthdr-00

Emil Ivov
Enrico Marocco
Jonathan Lennox

IETF 75

Where we come from
(the new RAI process)

- DISPATCH

  - Problem statement discussion

    - draft-ivov-dispatch-slic-ps

  - Several solution approaches

    - 4575 extension

    - RTP/RTCP extension

    - CSRC hacks

  - Converged on a specific solution

 - Forwarded to AVT
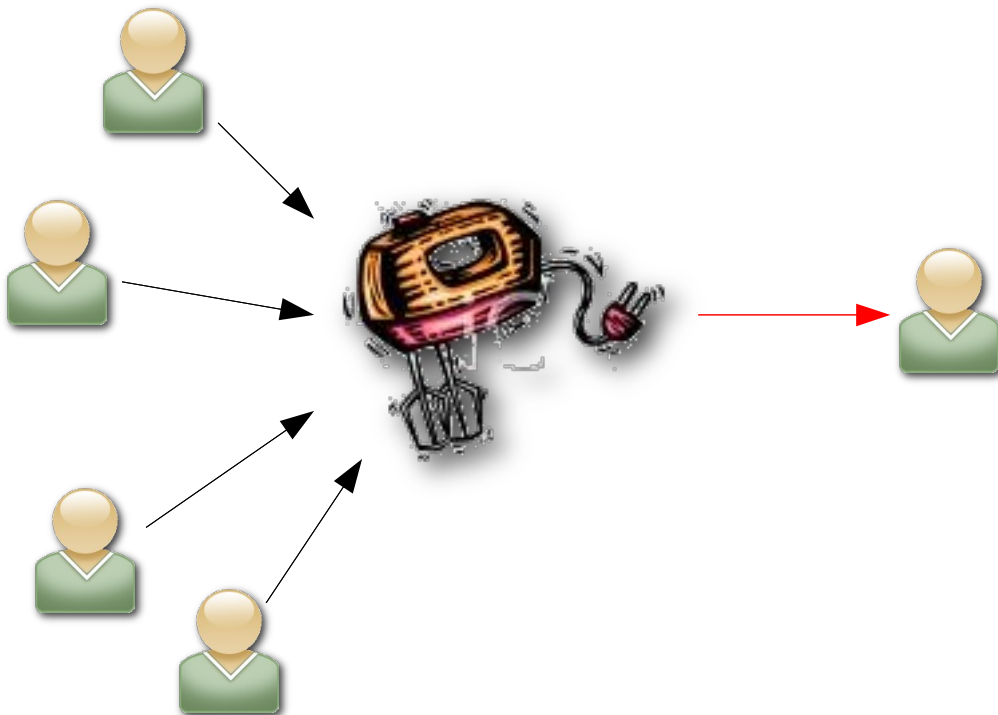
Where we come from
(the new RAI process)

- AVT

  - 5285-based solution proposed

    - draft-ivov-avt-slic

  - Similar/complementary problem and solution proposed

    - draft-lennox-avt-rtp-audio-level-exthdr

The problem (part 1)

draft-ivov-avt-slic

Enabling mixers to deliver extended information about the audio activity (sound level) of participants in a conference call
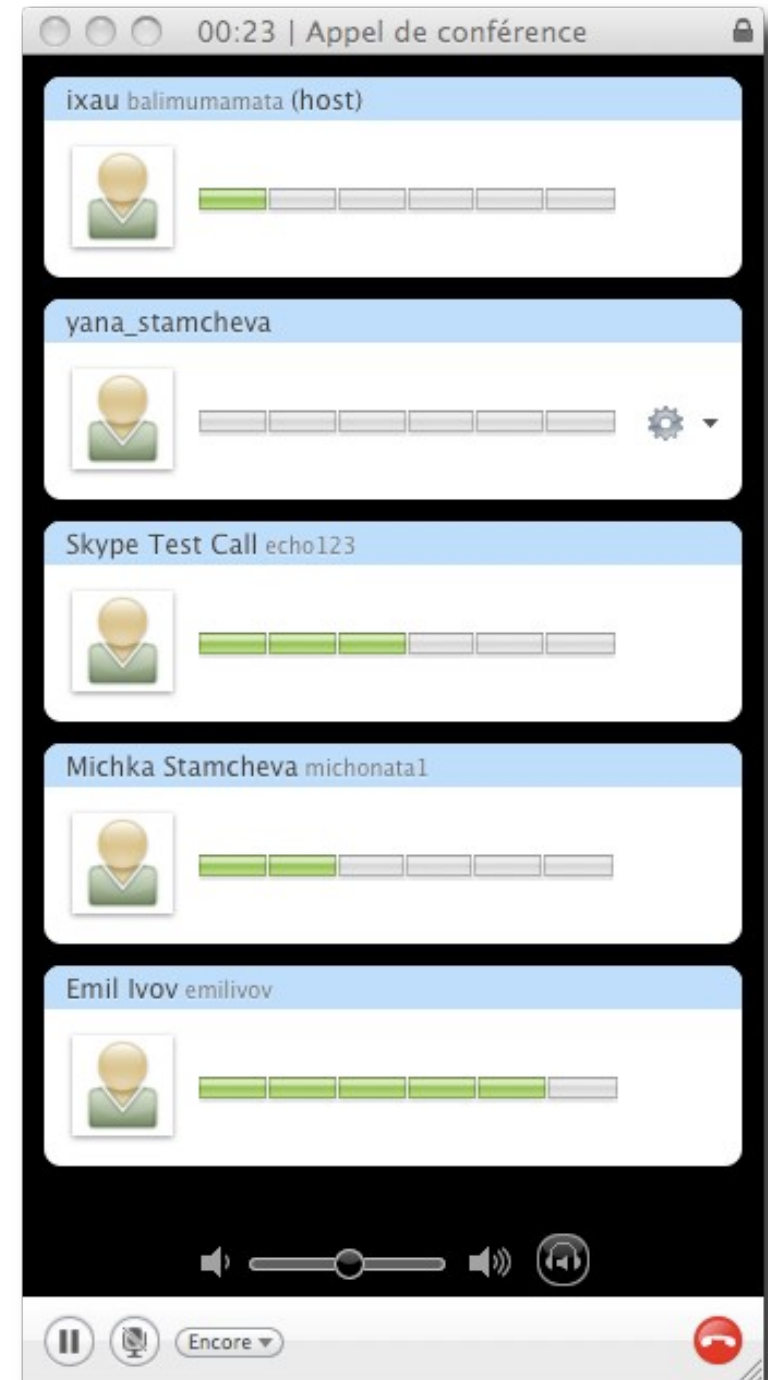
Why?

- Because it's cool!
  - Skype users are used to it

- To identify who's talking
  - When participants do not know each other
  - Esp. when following a heated debate
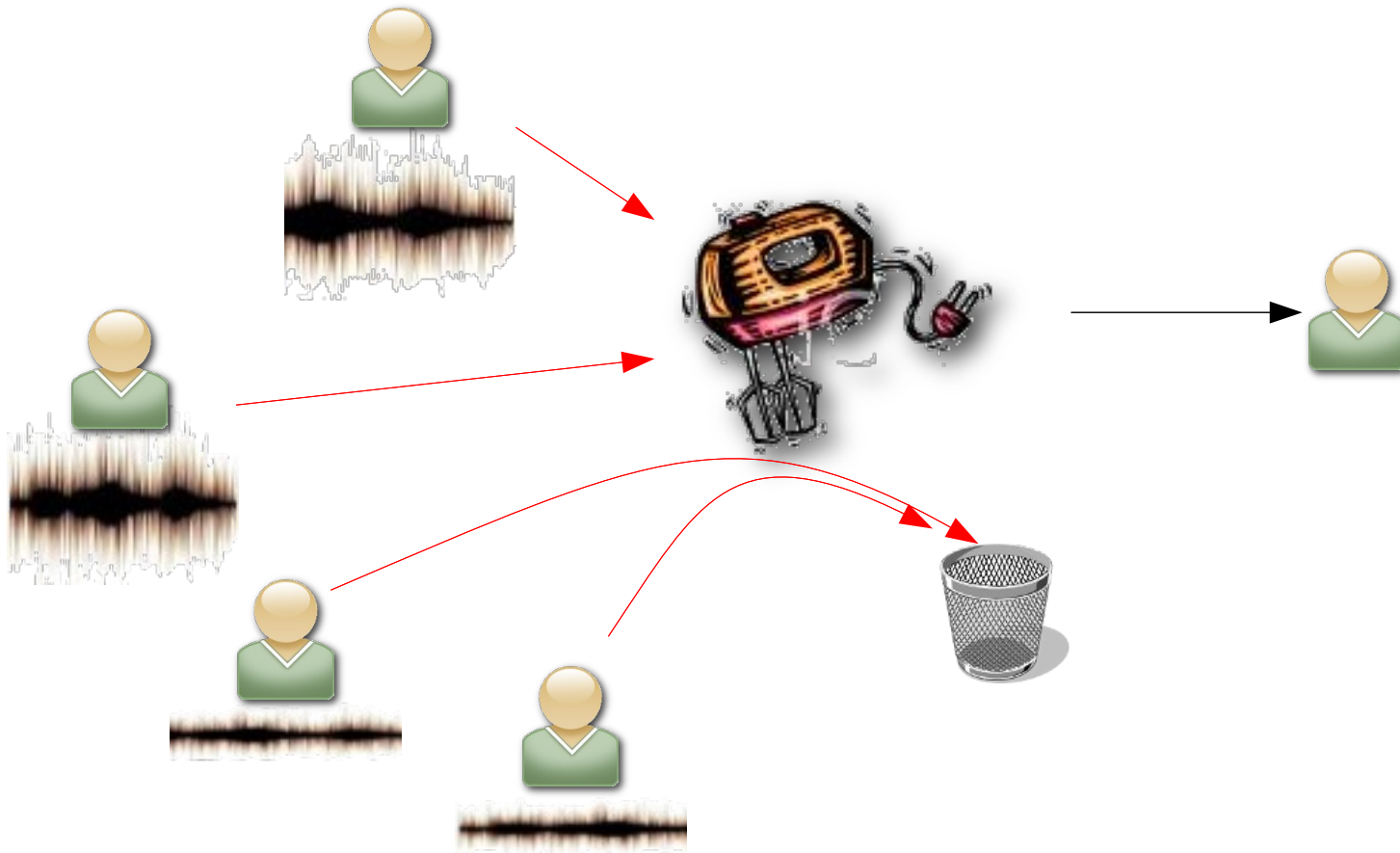
- To identify noise sources

Why?

Why?

The problem (part 2)
draft-lennox-avt-rtp-audio-level-exthdr

Enabling senders of audio packets to indicate the audio level of their packets' payload
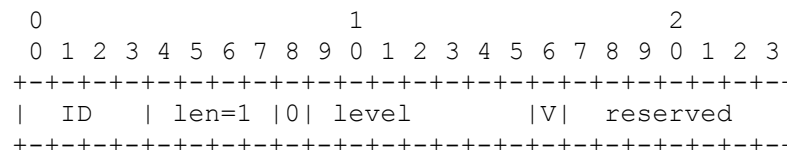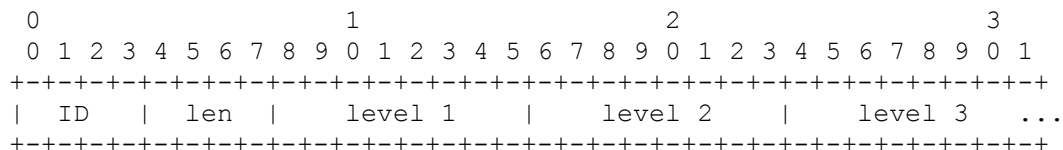
Why?

- Usually only the loudest streams get mixed
- To save mixer resources for under-the-threshold streams
  - Decoding
  - Audio level measurement
  - Voice detection (optional)

Solution(s) approach

- 5285-based RTP header extension to encode sound levels

  - For each contributing stream (i.e. associated to CSRCs)

  - For the whole stream (i.e. associated to the SSRC)

- Levels defined as for comfort noise (3389)

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  ID   |  len  |   level 1     |    level 2    |    level 3   ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

```
 0                   1                   2
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  ID   | len=1 |0| level      |V|  reserved   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Main Issue

- One problem or two problems (i.e. one extension or two extensions)?

- One-size-fits-all?

  - One extension for both SSRC <u>and</u> CSRC levels

  - Pro: Non-mixing clients won't need to send CSRC levels

  - Pro: Mixers passing CSRC levels are also likely to want to pass the SSRC level...

    … to help in the case of mixers mixing mixed streams

  - Con: Tricky to pass levels for SSRC or CSRCs <u>only</u> if both the SSRC and CSRC IDs are present in the RTP header

Minor issues

- Format

- Voice detection support
  - Complex implementation (cannot mandate it)
  - May be useful