IETF 90 AVTCORE PRIVACY ENSURED CLOUD CONFERENCING

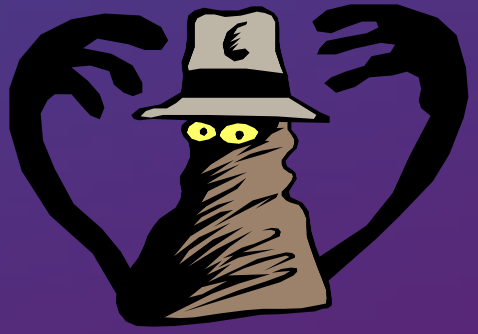
DRAFT-MATTSSON-AVTVORE-CLOUD-CONFERENCING-USE-CASE-00

JOHN MATTSSON

ERICSSON RESEARCH

MOTIVATION

- Industry transformation to cloud based, virtualized, and software based conferencing
  - One enabling factor is increased end-point capabilities, enabling them to process multiple media streams.
  - From mixing to selection, switching, and forwarding
  - This has a number of positive effects on flexibility, cost efficiency, ease of use, etc.
- But use of third-party cloud services increases the threats to privacy.
  - We know that there are many organizations actively performing large scale pervasive monitoring
- IETF should make cloud services viable and trustworthy from a pervasive monitoring perspective.



GOALS AND NON-GOALS

- Goals
  - Support use of third-party Cloud Services
  - Ensure End-To-End Confidentiality
  - Ensure End-To-End Source Authentication
  - Ensure End-To-End Replay Protection
  - More Efficient than Full-Mesh

- Non-Goals (or would be good but is difficult to accomplish)
  - Securing the Endpoints
  - Individual Media Source Authentication
  - Preventing Access before joining / after leaving.

WHICH RTP TOPOLOGY?

- Which RTP topology? (RTP Mixer, Video Switching MCU, …)

```
   RTCP  <------------> ? <------------->
    RTP  <------------> ? <------------->
Payload  <------------------------------>
       +---+      +------------+      +---+
       | A |<---->|            |<---->| B |
       +---+      |   Cloud    |      +---+
                  |  Service   |
       +---+      |            |      +---+
       | C |<---->|            |<---->| D |
       +---+      +------------+      +---+
```

RTP TOPOLOGY? RTP MIXER?

- RTP Payload needs to be sent end-to-end.
  - Receiver needs info to find context, authenticate, and decrypt.
- Duplicating and forwarding SRTP packets would prevent the mixer from doing any RTP and RTCP rewrites.
  - Switching causes gaps in RTP sequences, hiding packet loss.
    - Can cause repair attempts, buffering issues, and trigger bit-rate adaptation.
    - Significant difficulties for congestion control
  - Requires RTP stacks capable of handling multiple remote peers, including adaptation of congestion control.
  - Mixer cannot authenticate packets from end-points.
  - No confidentiality for information needed by the mixer.

PROBLEMS WITH CURRENT TECHNOLOGY

- Effective cloud based conferencing while protecting from pervasive monitoring, requires two layers of security.
  - This is not supported by SRTP. SRTP derives everything from a single master key.

- Middle boxes needs to take local switching decisions
  - Which streams: Each sender needs to include some speaker activity indication. However, this indication needs to leak as little information as possible about the actual content of the speech.
  - Where in the stream: Need to know from which points in the video streams a receiving endpoint will be able to decode. Thus markers for switching points in the media stream are needed

NEXT STEPS?

- Should IETF work on this?
- What should be standardized?
  - Minimum protocols for interoperability with third-party service or larger solution (interaction with conference, identity, and key servers)?
    - RTP topology? RTP Mixer?
    - What is needed beside two security contexts? Speaker activity indication? Switching point markers?
  - Native clients only? WebRTC? (Must standardize APIs, key management etc.)?
  - Use cases: Cloud Conferencing? Caching protected media?
- Where?
  - Avtcore?
  - Other IETF WGs?
  - W3C?

ERICSSON