RFC 3960
Early Media and Ringing Tone Generation in the Session Initiation Protocol (SIP)

Presented by Heng-Te Chu
Instructed by Assistant Professor Quincy Wu
Outline

- Introduction to RFC 3960
- Session Establishment in SIP
- The Gateway Model
- The Application Server Model
- Alert-Info Header Field
- Security Considerations
- Conclusion
Introduction to RFC 3960

- Early Media
  - unidirectional
  - or bidirectional
  - Generated by
    - Caller
    - Callee
    - Or both
Session Establishment in SIP

- **Offer/ Answer model**
  - SDP messages.
  - Offer in INVITE, answer in 200 OK,
  - Offer in 200 OK, answer in ACK
  - In PRACK, UPDATE, and etc

- **Media Clipping**
  - Caller is prepared to listen, no media clipping
Media Clipping

UAS cannot send any media until the answer in the ACK arrives.

Media Clipping – some first words lost

UAC cannot send media until the 200 (OK) arrives

Media Clipping – some first words lost
The Gateway Model

- Manage early media sessions
  - using offer/answer exchanges
    - in reliable provisional responses, PRACKs, and UPDATEs.
- If INVITE forks, media clipping may occur.
Forking

UAC  UAS  UAS  UAS

INVITE  offer
183 Session Progress  answer
PRACK
200 OK (PRACK)

Early Session

Mute randomly

200 OK (INVITE)
ACK
Re-INVITE or UPDATE new offer/answer

Regular Session
In the PSTN, local exchange of the callee generates a standardized ringing tone while the callee is being alerted.

But, SIP UAs have different capabilities, different user interfaces.
- UAC is supposed to generate ringing tones locally as long as no early media is received from the UAS.
- If the UAS generates early media, UAC is supposed to play it instead.
- Not easy to decide
Ringing Tone Generation

- **Common local policy**
  - No local ringing without 180 response received,
  - If 180 (Ringing) received but no incoming media packets, generate local ringing.
  - If 180 (Ringing) received and incoming media packets, play them and do not generate local ringing.
  - i.e. play incoming media packets and stop local ringing to avoid media clipping, even if the 200 (OK) response has not arrived.
Applicability of the Gateway Model

- The gateway model produces media clipping in forking scenarios and requires media detection to generate local ringing properly.
- Only acceptable in situations where the UA cannot distinguish between early media and regular media.
  - E.g., PSTN Gateway
The Application Server Model

- Having the UAS behave as an application server to establish early media sessions with the UAC.
- UAC indicates support
  - early-session option tag
- Different offer/answer exchange
  - avoid media clipping in cases of forking.
- UA still have to choose which to mute and which to render to the user.
The Application Server Model

UAC

UAS

1. INVITE (offer) Contect-Disposition: session
2. 183 Session Progress (early-offer) Contect-Disposition: early-session
3. PRACK (early-answer) Contect-Disposition: early-session
4. 200 OK (PRACK)
5. 200 OK (INVITE)
6. ACK

Early Session

Regular Session
Alert-Info Header Field

- For both gateway and application server models
- Allows to tell the UAC which local tone to play.
  - In case that local ringing is generated.
- does not tell the UAC when to generate locally.
Security Considerations

- Session description contains the IP address and port number where to receive media,
  - Attackers may intrude it.
- UAs should encrypt their session descriptions
  - e.g., using S/MIME (RFC 2633)
  - Still, may be guessed. Because many UAs always pick up the same initial media port.
  - media-level authentication mechanisms
    - Secure Realtime Transport Protocol (SRTP) (RFC 3711)
Security Considerations

- Attackers may attempt to make a UA send media to a victim as part of a DoS attack.
  - UA should handshake with the owner to verify the willingness to receive media
    - by using a connection oriented transport protocol,
    - STUN (RFC3489) in an end-to-end fashion,
    - or by the key exchange in SRTP.
Security Considerations

- Early media-specific risk
  - Like "toll fraud" in the PSTN
    - may try to establish a bidirectional early media session and never send a 200 (OK) response for the INVITE.
- Servers may use bidirectional early media to obtain information from the callers
  - e.g., the PIN code of a calling card.
- Remedy – to charge early media that last too long or stop them at media level.
Conclusion

- The gateway model is acceptable in situations where the UA cannot distinguish between early media and regular media.
  - e.g., PSTN
- The application server model resolves the issues present in the gateway model.
  - is strongly recommended to use this model.
Thank You!

Q & A