

SIV IN MRCP

W₃C Biometrics Workshop

March 2009

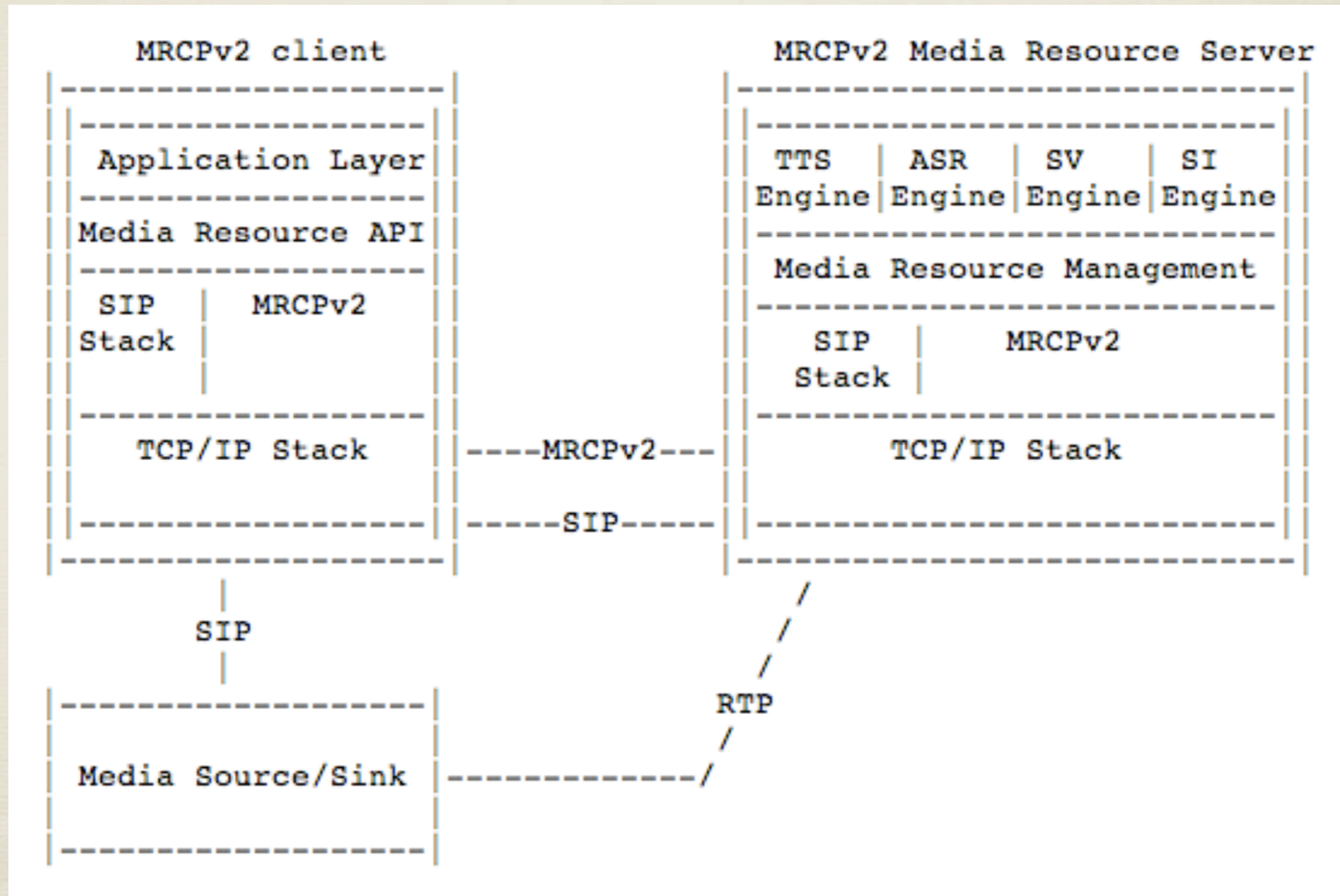
Overview

- * What is MRCP?
- * MRCP_{v1}
- * SIV in MRCP_{v1}
- * MRCP_{v2}
- * SIV Goals
- * Verification Resource
- * Security Model
- * Verification Example

What is MRCP?

- * Media Resource Control Protocol
- * Protocol-level API standard for controlling speech-related technologies (ASR, TTS, SIV)
- * Server: connects to ASR/TTS/SIV engines
- * Client: implemented inside, e.g., VoiceXML interpreter

What is MRCP?



MRCPv1

- * Developed in 2001-2002 by Cisco, Nuance, and Speechworks
- * “Tunneled” over RTSP ANNOUNCE method
- * NOT an IETF standards-track document
- * but Informational RFC 4463

SIV in MRCPv1

- * Not officially included in MRCPv1
- * Proposed at 58th IETF (http://www.standardtrack.com/ietf/speechsc/slides58/si_sv_ietf58.ppt)
- * Add-on created by Nuance and Intervoice in Dec 2003
- * Can be found at <http://www.standardtrack.com/ietf/speechsc/drafts/draft-burnett-mrcpext-01.txt>

MRCPv2

- * Standards-track document in SPEECHSC Working Group, with many participating companies
- * Its own protocol -- not tunneled over RTSP
- * Borrows structure from HTTP
- * SIV included from the beginning
- * Also adds voice enrolled grammars

SIV Goals

- * See <http://www.ietf.org/rfc/rfc4313.txt>, section 6
- * Support both separate and simultaneous ASR and SIV
- * Create simple model for both identification and verification
- * Support training on data recorded earlier in session
- * No explicit model access
- * Support text independent and text dependent SIV

Verification Resource

- * Session
- * Buffering
- * Simultaneous ASR/SIV
- * Training/Verification
- * Result in same structure as ASR
- * Speaker identification is special case of verification

Security Model

- * Audio is expected to be secured via channel-specific mechanisms -- technically out of scope for this control protocol
- * voiceprints are never handled by the client or server
- * cookies can be used for server authentication

Verification Example

C->S: MRCP/2.0 123 **START-SESSION** 314161
Channel-Identifier:32AECB23433801@speakverify
Repository-URI:<http://www.example.com/voiceprintdbase/>
Voiceprint-Mode:verify
Voiceprint-Identifier:johnsmith.voiceprint
Adapt-Model:true

S->C: MRCP/2.0 49 314161 200 **COMPLETE**
Channel-Identifier:32AECB23433801@speakverify

C->S: MRCP/2.0 123 **END-SESSION** 314174
Channel-Identifier:32AECB23433801@speakverify
Abort-Model:true

S->C: MRCP/2.0 49 314174 200 **COMPLETE**
Channel-Identifier:32AECB23433801@speakverify

C->S: MRCP/2.0 49 **VERIFY** 543260
Channel-Identifier:32AECB23433801@speakverify

S->C: MRCP/2.0 49 543260 200 **IN-PROGRESS**
Channel-Identifier:32AECB23433801@speakverify

S->C: MRCP/2.0 123 **VERIFICATION-COMplete** 543259 COMPLETE
Completion-Cause:000 success
Content-Type:application/nlsml+xml
Content-Length:739

```
<?xml version="1.0"?>
<result xmlns="http://www.ietf.org/xml/ns/mrcpv2"
  grammar="What-Grammar-URI">
  <verification-result>
    <voiceprint id="johnsmith">
      <incremental>
        <utterance-length> 500 </utterance-length>
        <device> cellular-phone </device>
        <gender> male </gender>
        <decision> accepted </decision>
        <verification-score> 0.85 </verification-score>
      </incremental>
      <cumulative>
        <utterance-length> 1500 </utterance-length>
        <device> cellular-phone </device>
        <gender> male </gender>
        <decision> accepted </decision>
        <verification-score> 0.75 </verification-score>
      </cumulative>
    </voiceprint>
  </verification-result>
</result>
```

SIV in MRCP

Dan Burnett
Director of Speech Technologies
Voxeo