VoIP Analysis Fundamentals with Wireshark…

Phill Shade (Forensic Engineer – Merlion's Keep Consulting)

Phillip D. Shade (Phill)
phill.shade@gmail.com

- Phillip D. Shade is the founder of Merlion's Keep Consulting, a professional services company specializing in Network and Forensics Analysis

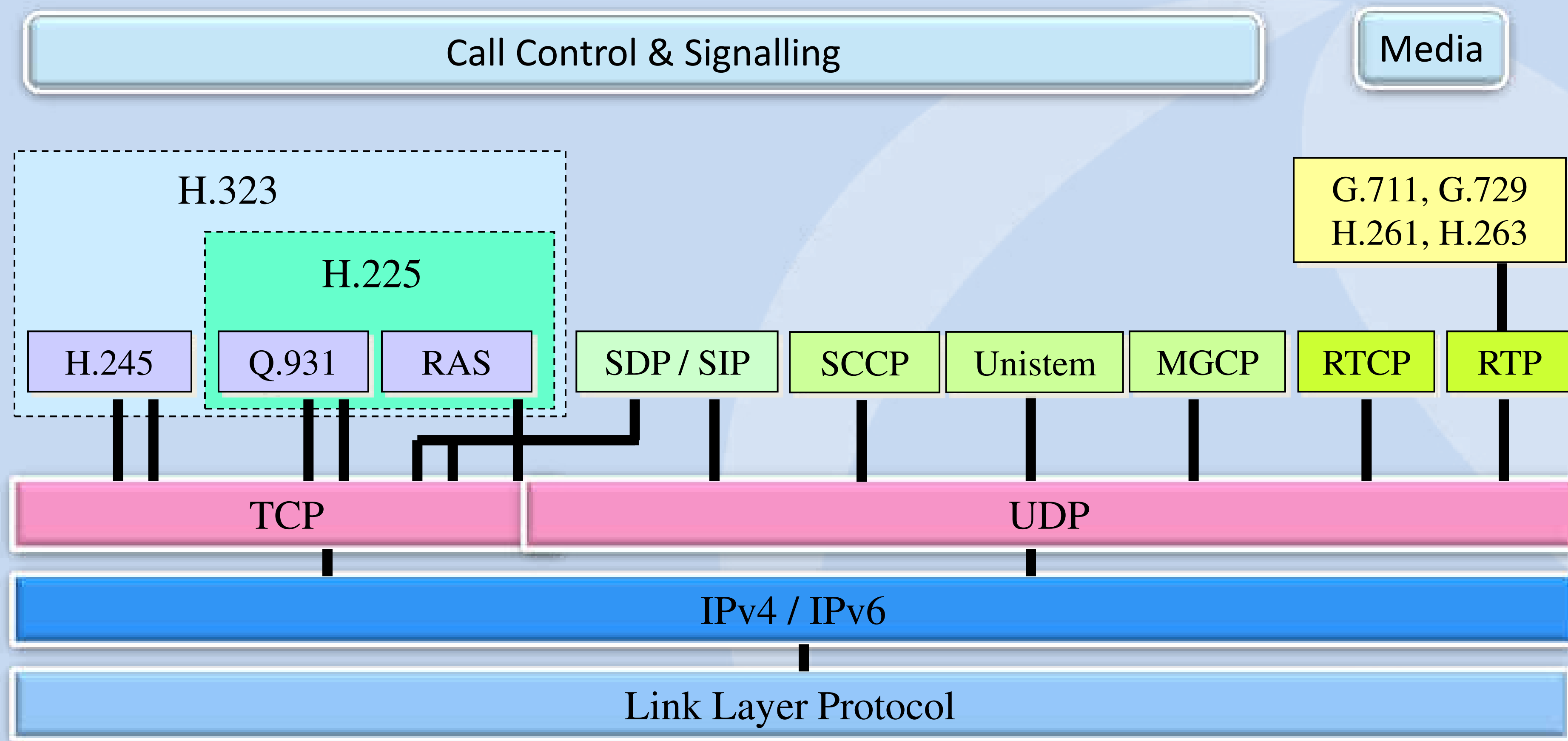
- Internationally recognized Network Security and Forensics expert, with over 30 years of experience

- Member of FBI InfraGard, Computer Security Institute, the IEEE and Volunteer at the Cyber Warfare Forum Initiative

- Numerous certifications including CNX-Ethernet (Certified Network Expert), Cisco CCNA, CWNA (Certified Wireless Network Administrator), WildPackets PasTech and WNAX (WildPackets Certified Network Forensics Analysis Expert)

- Certified instructor for a number of advanced Network Training academies including Wireshark University, Global Knowledge, Sniffer University, and Planet-3 Wireless Academy.

VoIP / Video Protocol Stack



VoIP Protocols Overview (Signaling)

- **MGCP - Media Gateway Control Protocol**
  - Defined by the IETF and ITU
  - Used to control signaling and session management (also known as H.248 or Megaco)
- **SCCP - Skinny Client Control Protocol**
  - CISCO proprietary protocol used to communicate between a H.323 Proxy (performing H.225 & H.245 signaling) and a Skinny Client (VoIP phone)
- **SIP - Session Initiation Protocol**
  - Defined by the IETF / RFC 2543 / RFC 3261
- **H.323 – Defines a Suite of ITU designed protocols**
  - H.225, H.245, Q.931, RAS, etc…

VoIP Protocols Overview (Data)

- **RTP** - Real Time Protocol
  - Defined by the IETF / RFC 1889
  - Provides end-to-end transport functions for applications transmitting real-time data over Multicast or Unicast network services
    - Audio, video or simulation data

- **RTCP** - Real Time Control Protocol
  - Defined by the IETF
  - Supplements RTP's data transport to allow monitoring of the data delivery in a manner scalable to large Multicast networks
  - Provides minimal control and identification functionality

- **RTSP** - Real Time Streaming Protocol
  - Defined by the IETF / RFC 2326
  - Enables the controlled delivery of real-time data, such as audio and video
  - Designed to work with established protocols, such as RTP and HTTP

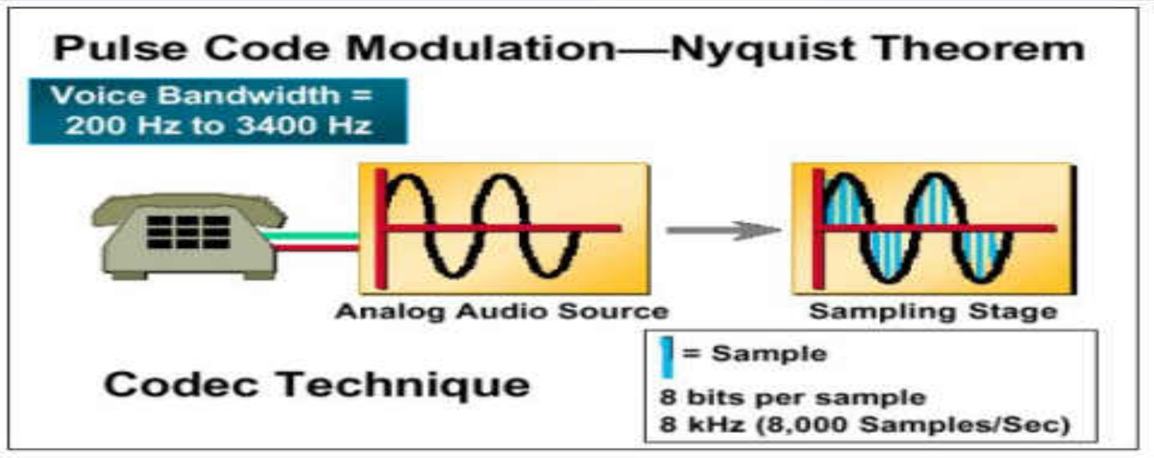
VoIP Codecs (Audio Conversion)

- CODEC = Compressor / Decompressor or Coder / Decoder or Reader
  - Provides conversion between Audio/Video signals and data streams at various rates and delays

- Designations conform to the relevant ITU standard
  - Audio Codecs (G.7xx)
    - G.711a / u - PCM Audio 56 and 64 Kbps (Most common business use)
    - G.722 - 7 Khz Audio at 48, 56 and 64 Kbps
    - G.723.1 / 2- ACELP Speech at 5.3 Kbps / MPMLQ at 6.3 Kbps
    - G.726 - ADPCM Speech at 16, 24, 32 and 40 Kbps
    - G.727 - E-ADPCM Speech at 16, 24, 32 and 40 Kbps
    - G.728 - LD-CELP Speech at 16 Kbps
    - G.729 - CS-ACELP Speech at 8 and 13 Kbps (Very common for home use)
  - Video Codecs (H.2xx)
    - H.261 - Video >= 64 Kbps
    - H.263 - Video <= 64 Kbps

VoIP Codecs

- CODEC = Compressor / Decompressor or Coder / Decoder or Reader
  - Provides conversion between Audio/Video signals and data streams at various rates and delays



Sample VoIP Codec Comparison

Codec	Data Rate	Typical Datagram Size	Packeti -zation Delay	Combined Bandwidth for 2 Flows	Typical Jitter Buffer Delay	Theoretical Maximum MOS
G.711u	64.0 kbps	20 ms	1.0 ms	174.40 kbps	2 datagrams (40 ms)	**4.40**
G.711a	64.0 kbps	20 ms	1.0 ms	174.40 kbps	2 datagrams (40 ms)	**4.40**
G.726-32	32.0 kbps	20 ms	1.0 ms	110.40 kbps	2 datagrams (40 ms)	**4.22**
G.729	8.0 kbps	20 ms	25.0 ms	62.40 kbps	2 datagrams (40 ms)	**4.07**
G.723.1 MPMLQ	6.3 kbps	30 ms	67.5 ms	43.73 kbps	2 datagrams (60 ms)	**3.87**
G.723.1 ACELP	5.3 kbps	30 ms	67.5 ms	41.60 kbps	2 datagrams (60 ms)	**3.69**

- MOS and R value include Packetiaztion delay + Jitter buffer delay
- Common bandwidth – real bandwidth consumption:

  # Payload = 20 bytes/p (40 bytes/s)

  # Overhead includes 40 bytes of RTP header (20 IP + 8 UDP + 12 RTP)

Competing Signaling Standards

- **Several different standards are currently competing for dominance in the VoIP field:**
  - **H.323 -** Developed by the International Telecommunications Union (ITU) and the Internet Engineering Task Force (IETF)
  - **MGCP / Megaco/ H.248 -** Developed by CISCO as an alternative to H.323
  - **SIP -** Developed by 3Com as an alternative to H.323
  - **SCCP –** Cisco Skinny Client Control Protocol – used to communicate between a H.323 Proxy (performing H.225 & H.245 signaling) and a Skinny Client (VoIP phone)
  - **UNISTEM –** Proprietary Nortel protocol, developed by as an alternative to H.323

H.323 - Packet-based Multimedia Communications Systems

- An umbrella standard defined by the International Telecommunications Union (ITU) and the Internet Engineering Task Force (IETF)
- Defines a set of call controls, channel set up and Codec's for multimedia, packet-based communications systems using IP-based networks

H.450.1	Supplemental, generic protocol for use under H.323
H.225	Call Signaling / RAS
H.245	Control messages for the H.323 Terminal (RTP / RTCP)
H.235	Security Enhancements
Q.931	Call setup and termination
G.711, G.723.1 G.728	Audio Codec's
H.261, H.263, H.264	Video Codec's

SIP VoIP Standard (SIP)

- Defined in RFC 2543 and RFC 3261 and by the ITU
  - Pioneered by 3Com to address weaknesses in H.323
- Application layer signaling protocol supporting real time calls and conferences (often involving multiple users) over IP networks
  - Can replace or complement MGCP
    - SIP provides Session Control and the ability to discover remote users
    - SDP provides information about the call
    - MGCP/SGCP Provides Device Control
    - ASCII text based
    - Provides a simplified set of response codes
- Integrated into many Internet-based technologies such as web, email, and directory services such as LDAP and DNS
  - Extensively used across WANs

MGCP / Megaco VoIP Standards

- Defined by RFC 2705 / 3015 and the ITU in conjunction with the H.248 standard
  - Pioneered by CISCO to address weaknesses in H.323
- Used between elements of distributed Gateways (defined later) as opposed to the older, single all-inclusive Gateway device
  - Extensively used in the LAN environment
- Utilizes Media Gateway Control Protocol (MGCP) to control these distributed elements
  - Often considered a “Master/Slave” protocol

Quality Of Service (QoS) - Overview

- Provides a guarantee of bandwidth and availability for requesting applications
  - Used to overcome the hostile IP network environment and provide an acceptable Quality of Service
    - Delay, Jitter, Echo, Congestion, Packet loss and Out of Sequence packets
  - Mean Opinion Score (MoS) / R-Factor is sometimes used to determine the requirements for QoS.
  - Utilized in the VoIP environment in one of several methods:
    - Resource Reservation Protocol (RSVP) defined by IETF
    - IP Differentiated Services
    - IEEE 802.1p and IEEE 802.1q

Assessing Voice Quality

- Voice Quality can be measured using several criteria

  1. Delay: As delay increases, callers begin talking over each other, eventually the call will sound like talking on a “walkie-talkie”.   (Over…)

  2. Jitter: As jitter increases, the gateway becomes unable to correctly order the packets and the conversation will begin to sound choppy

  - Some devices utilize jitter buffer technology to compensate

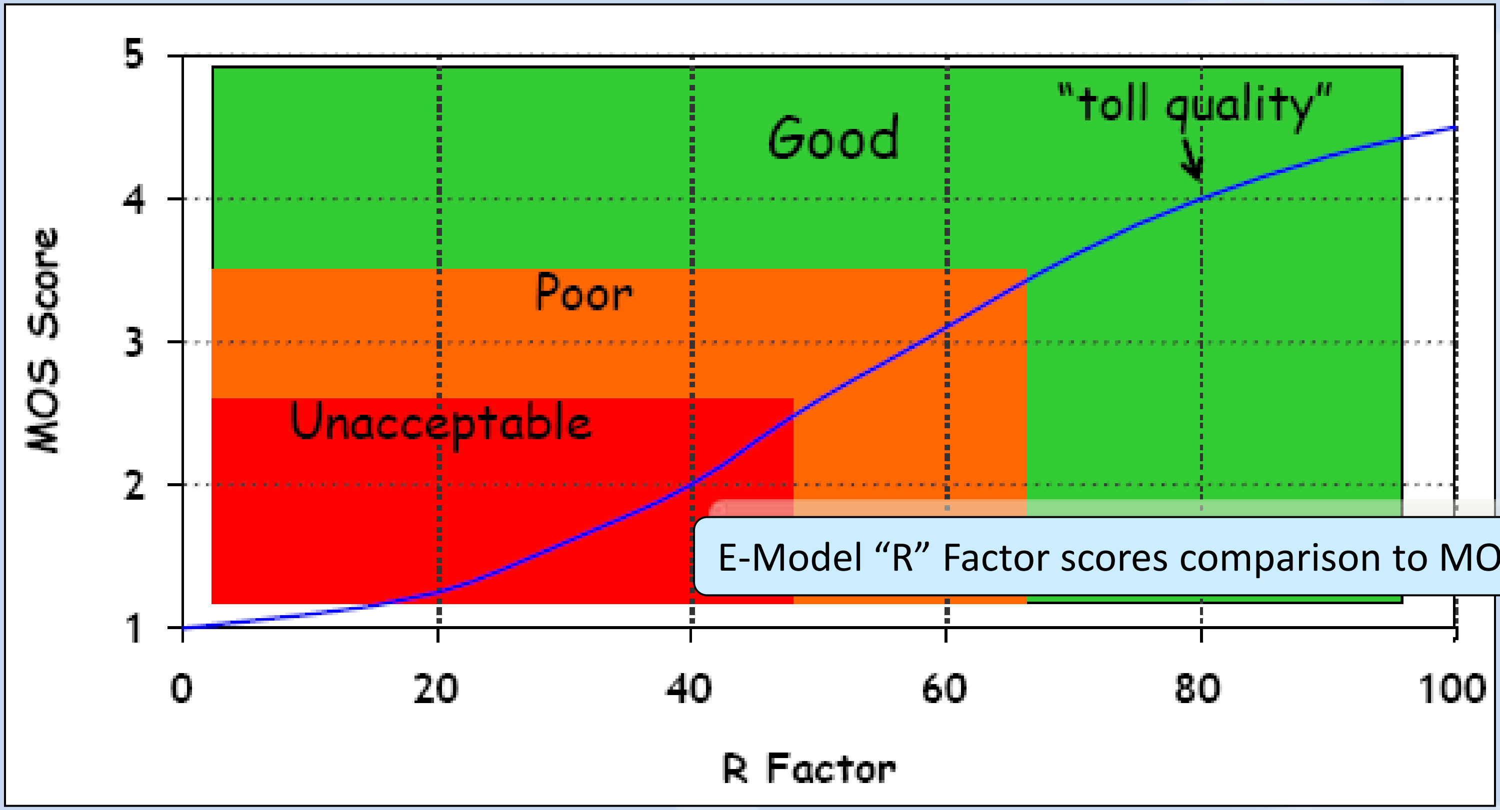
  3. Packet Loss: If packet loss is greater than the jitter buffer, the caller will hear dead air space and the call will sound choppy

  - Gateways are designed to conceal minor packet loss

Different VoIP Quality Measurement Terms

- MoS – Mean Opinion Score
  - Numerical measure of the quality of human speech at the destination end of the circuit
- PSQM (ITU P.861)/PSQM+ - Perceptual Speech Quality Measure
- PESQ (ITU P.862) – Perceptual Evaluation of Speech Quality
- PAMS (British Telecom) Perceptual Analysis Measurement System
- The E-Model (ITU G.107) – (R-Factor)
  - Send a signal through the network, and measure the other end!

Measures of Voice Quality



E-Model "R" Factor scores comparison to MOS score

- MOS can only be measured by humans
- R-value can be calculated in software
- PMOS values can be determined from R-value

MOS (Mean Opinion Score)

MOS	Quality Rating
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

1. Quality Goal is the same as PSTN and is widely accepted criterion for call quality

2. Call quality testing has always been subjective (Humans) - International Telecommunications Union (ITU) P.800

MOS - *Mean Opinion Score*

- Numerical measure of the quality of human speech at the destination end of the circuit (affected extensively by Jitter)
- Uses subjective tests (opinionated scores) that are mathematically averaged to obtain a quantitative indicator of the system performance
- Rating of 5.0 is considered perfect

E-Model (R-Factor)

- The E-Model - Recommendation ITU G.107
  - The "E-Model" is a parameter based algorithm based on subjective test results of auditory tests done in the past compared with current “system parameters”
  - Provides a prediction of the expected quality, as perceived by the user
  - The result of the E-Model calculation is “**E-Model Rating R**” (0 - 100) which can be transformed to “**Predicted MOS (PMOS)**” (1 – 5; 5 is non-extended, non-compressed)
    - Typical range for R factors is 50-94 for narrowband telephony and 50-100 for wideband telephony

Cascade Pilot Computes the R-Factor and MOS scores

“R” Factor vs. MOS in Cascade Pilot

Caller Number ▲ Receiver Number ▲ Call-ID ▲

Hierarchy (Caller Number/Receiver Number/Call-ID)	RTP Src IP	RTP Src Port	RTP Dst IP	RTP Dst Port	SSRC	Payload Type	Avg R-Factor	Max R-Factor
- Caller Number: 3290	[3]	[4]	[3]	[4]	[3]	[1]	79.62	93.34
- Receiver Number: 4672	[2]	[2]	[2]	[2]	[2]	[1]	68.90	93.34
- Call-ID: 003094c3-438b0085-4ef3a663	[2]	[2]	[2]	[2]	[2]	[1]	68.90	93.34
	45.210.3.90	19716	45.210.9.72	2238	0x8b43c394	PCMU	68.98	93.34
	45.210.9.72	2238	45.210.3.90	19716	0x13c443d3	PCMU	68.83	93.34
- Receiver Number: 4697	[2]	[2]	[2]	[2]	[2]	[1]	90.33	93.34
- Call-ID: 003094c3-438b0083-6f807304	[2]	[2]	[2]	[2]	[2]	[1]	90.33	93.34
	45.210.9.97	5004	45.210.3.90	19712	0x7ef3a938	PCMU	90.33	93.34
	45.210.3.90	19712	45.210.9.97	5004	0x8b43c394	PCMU	90.33	93.34
Summary	[3]	[4]	[3]	[4]	[3]	[1]	79.62	93.34

Cascade Pilot computes both “R” Factor and MOS in multiple formats:
1. Average - R Factor / MOS
2. Maximum - R Factor / MOS

Caller Number ▲ Receiver Number ▲ Call-ID ▲

Hierarchy (Caller Number/Receiver Number/Call-ID)	RTP Src IP	RTP Src Port	RTP Dst IP	RTP Dst Port	SSRC	Payload Type	Avg MOS	Max MOS
- Caller Number: 3290	[3]	[4]	[3]	[4]	[3]	[1]	3.83	4.41
- Receiver Number: 4672	[2]	[2]	[2]	[2]	[2]	[1]	3.35	4.41
- Call-ID: 003094c3-438b0085-4ef	[2]	[2]	[2]	[2]	[2]	[1]	3.35	4.41
	45.210.3.90	19716	45.210.9.72	2238	0x8b43c394	PCMU	3.35	4.41
	45.210.9.72	2238	45.210.3.90	19716	0x13c443d3	PCMU	3.34	4.41
- Receiver Number: 4697	[2]	[2]	[2]	[2]	[2]	[1]	4.30	4.41
- Call-ID: 003094c3-438b0083-6f8	[2]	[2]	[2]	[2]	[2]	[1]	4.30	4.41
	45.210.9.97	5004	45.210.3.90	19712	0x7ef3a938	PCMU	4.30	4.41
	45.210.3.90	19712	45.210.9.97	5004	0x8b43c394	PCMU	4.30	4.41
Summary	[3]	[4]	[3]	[4]	[3]	[1]	3.82	4.41

Cascade Pilot – Quality Details

Caller Number | Receiver Number | Call-ID

Hierarchy (Caller Number/Receiver Number/Call-ID)	P Src Port	RTP Dst IP	RTP Dst Port	SSRC	Payload Type	Avg Jitter	Max Jitter	Avg Delta	Max Delta
- Caller Number: 3290	[4]	[3]	[4]	[3]	[1]	7.151ms	507.953ms	24.340ms	-296318us
- Receiver Number: 4672	[2]	[2]	[2]	[2]	[1]	8.330ms	507.953ms	23.070ms	-332398us
- Call-ID: 003094c3-43800081-4ef5a663	[2]	[2]	[2]	[2]	[1]	8.330ms	507.953ms	23.070ms	-332398us
	16	45.210.9.72	2238	0x8b43c394	PCMU	8.379ms	488.079ms	23.070ms	-333296us
	8	45.210.3.90	19716	0x13c443d3	PCMU	8.280ms	507.953ms	23.071ms	-332398us
- Receiver Number: 4697	[2]	[2]	[2]	[2]	[1]	5.973ms	396.187ms	25.610ms	-296318us
- Call-ID: 003094c3-43800083-6f807304	[2]	[2]	[2]	[2]	[1]	5.973ms	396.187ms	25.610ms	-296318us
	4	45.210.3.90	19712	0x7ef3a938	PCMU	6.200ms	395.187ms	25.605ms	-296788us
	12	45.210.9.97	5004	0x8b43c394	PCMU	5.745ms	394.989ms	25.616ms	-296318us
Summary	[4]	[3]	[4]	[3]	[1]	7.151ms	507.953ms	24.340ms	-296318us

Cascade Pilot computes both Jitter and Delta in multiple formats:

1. Average / Maximum Jitter
2. Average / Maximum Delta

Making the Call - SIP...

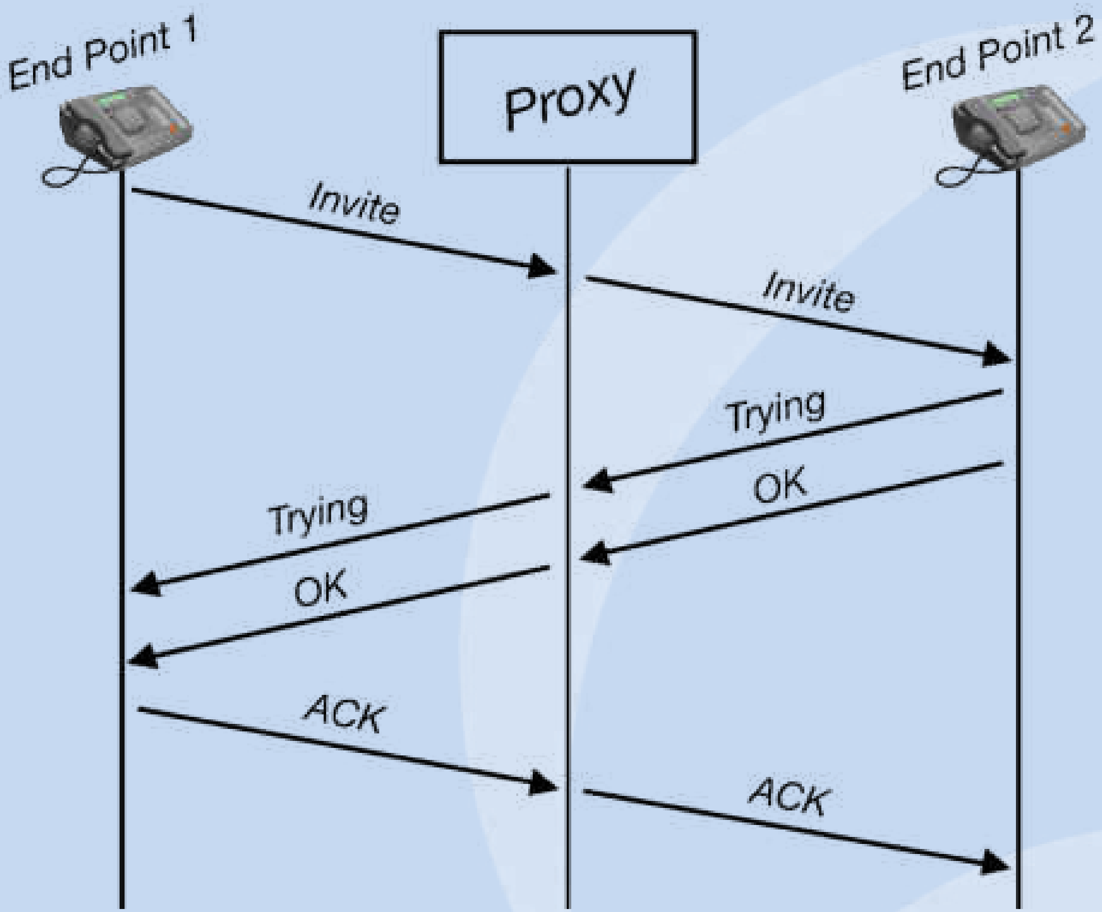
So what happens when we engage SIP VoIP?

Expected SIP Operation

- To initiate a session
  - Caller sends a request to a callee's address in the form of a ASCII text command
    - “Invite”
  - Gatekeeper/Gateway attempts phnoe number -> IP mapping/resolution
    - Trying / Response code = 100
    - Ringing / response code = 180
  - Callee responds with an acceptance or rejection of the invitation
    - “Accept” / response code=200 “OK”
  - Call process is often mediated by a proxy server or a redirect server for routing purposes

- To terminate a session
  - Either side issues a quit command in ASCII text form
    - “Bye”

SIP Call Setup



Session Initiation Protocol (SIP - Invite)

```
⊟ Session Initiation Protocol
  ⊟ Request-Line: INVITE sip:4697@cisco.sip.ilabs.interop.net;user=phone SIP/2.0
      Method: INVITE
    ⊞ Request-URI: sip:4697@cisco.sip.ilabs.interop.net;user=phone
      [Resent Packet: False]
  ⊟ Message Header
    ⊞ Via: SIP/2.0/UDP 45.210.3.90:5060;branch=z9hG4bK6137b728
    ⊟ From: "Cisco 3290" <sip:3290@cisco.sip.ilabs.interop.net>;tag=003094c3438b00cd52bdf1e8-0d2f4d4b
        SIP Display info: "Cisco 3290"
      ⊟ SIP from address: sip:3290@cisco.sip.ilabs.interop.net
          SIP from address User Part: 3290
          SIP from address Host Part: cisco.sip.ilabs.interop.net
        SIP tag: 003094c3438b00cd52bdf1e8-0d2f4d4b
    ⊟ To: <sip:4697@cisco.sip.ilabs.interop.net;user=phone>
      ⊟ SIP to address: sip:4697@cisco.sip.ilabs.interop.net;user=phone
          SIP to address User Part: 4697
          SIP to address Host Part: cisco.sip.ilabs.interop.net
      Call-ID: 003094c3-438b0083-6f807304-47943c3c@45.210.3.90
      Date: Thu, 13 May 2004 18:11:17 GMT
    ⊞ CSeq: 101 INVITE
      User-Agent: CSCO/6
    ⊞ Contact: <sip:3290@45.210.3.90:5060>
      Expires: 180
      Content-Type: application/sdp
      Content-Length: 244
      Accept: application/sdp
  ⊞ Message Body
```

SIP “Invite”

SIP is data is carried in text format

Session Initiation Protocol (SIP - Bye)

```
⊟ Session Initiation Protocol
  ⊟ Request-Line: BYE sip:3290@45.210.3.90:5060 SIP/2.0
      Method: BYE
    ⊞ Request-URI: sip:3290@45.210.3.90:5060
      [Resent Packet: False]
  ⊟ Message Header
    ⊞ Via: SIP/2.0/UDP 45.210.3.36:5060;branch=a84121e1-2d6f00ce-2bb702b0-fd00f62c-1
    ⊞ Via: SIP/2.0/UDP 45.210.3.36:5060;received=45.210.3.36;branch=cb89efff-be63b1bc-83f907fe-69cf5fcc-1, SIP/2.0/UDP
    ⊞ To: "Cisco 3290" <sip:3290@cisco.sip.ilabs.interop.net>;tag=003094c3438b00cf087acf0f-1340dfed
    ⊞ From: <sip:4672@cisco.sip.ilabs.interop.net;user=phone>;tag=614790957
      Call-ID: 003094c3-438b0085-4ef5a663-56f32b68@45.210.3.90
      Content-Length: 0
      Allow: INVITE,ACK,BYE,CANCEL,OPTIONS,INFO,MESSAGE,SUBSCRIBE,NOTIFY,PRACK,UPDATE,REFER
      User-Agent: PolycomSoundPointIP-UA/1.0.9
      Max-Forwards: 67
      k: com.nortelnetworks.firewall,100rel,p-3rdpartycontrol
    ⊟ CSeq: 36515 BYE
        Sequence Number: 36515
        Method: BYE
```

SIP - “Bye”

Challenges of VoIP

- Minimize Delay, Jitter and data loss
  – Excessive Delay variations can lead to unacceptable data lost or distortion
- Implementing QoS
  – RSVP designed to reserve required resources for VoIP traffic
- Interoperability of equipment beyond the Intranet
  – Different vendors Gateways utilize different Codec's
- Compatibility with the PSTN
  – Seamless integration required to support services such as smart card and 800 service

Factors Affecting Delay and VoIP Quality - 1

- Latency
  - Round trip latency is the key factor in a call having an "interactive feel"
  - <100 msec is considered idle

- Jitter
  - Occurs when packets do not arrive at a constant rate that exceeds the buffering ability of the receiving device to compensate for
  - If excessive Jitter occurs, larger Jitter buffers will be required which cause longer latency

Latency → Jitter Buffer → **Latency**

- Packet Loss
  - Loss of > 10% (non-consecutive packets) will be perceived as a bad connection

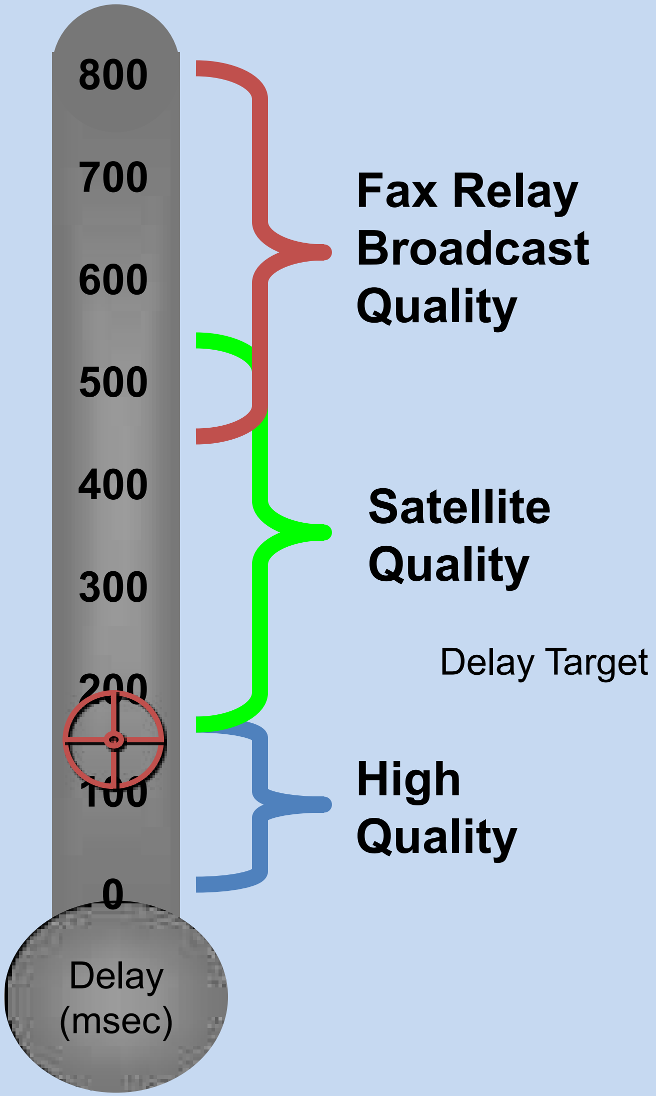
Factors Affecting Delay and VoIP Quality - 2

- Codec Choice
  - Add delay
    - Processing
    - Encoding / Decoding
  - Greater the compression factors result in lowered quality

- Bandwidth Utilization
  - Less utilization = lower latency, jitter and loss due to collisions

- Priority
  - Voice is extremely sensitive to delay
  - QoS is used to allow network devices to handle VoIP ahead of other traffic

Voice Quality & Delay

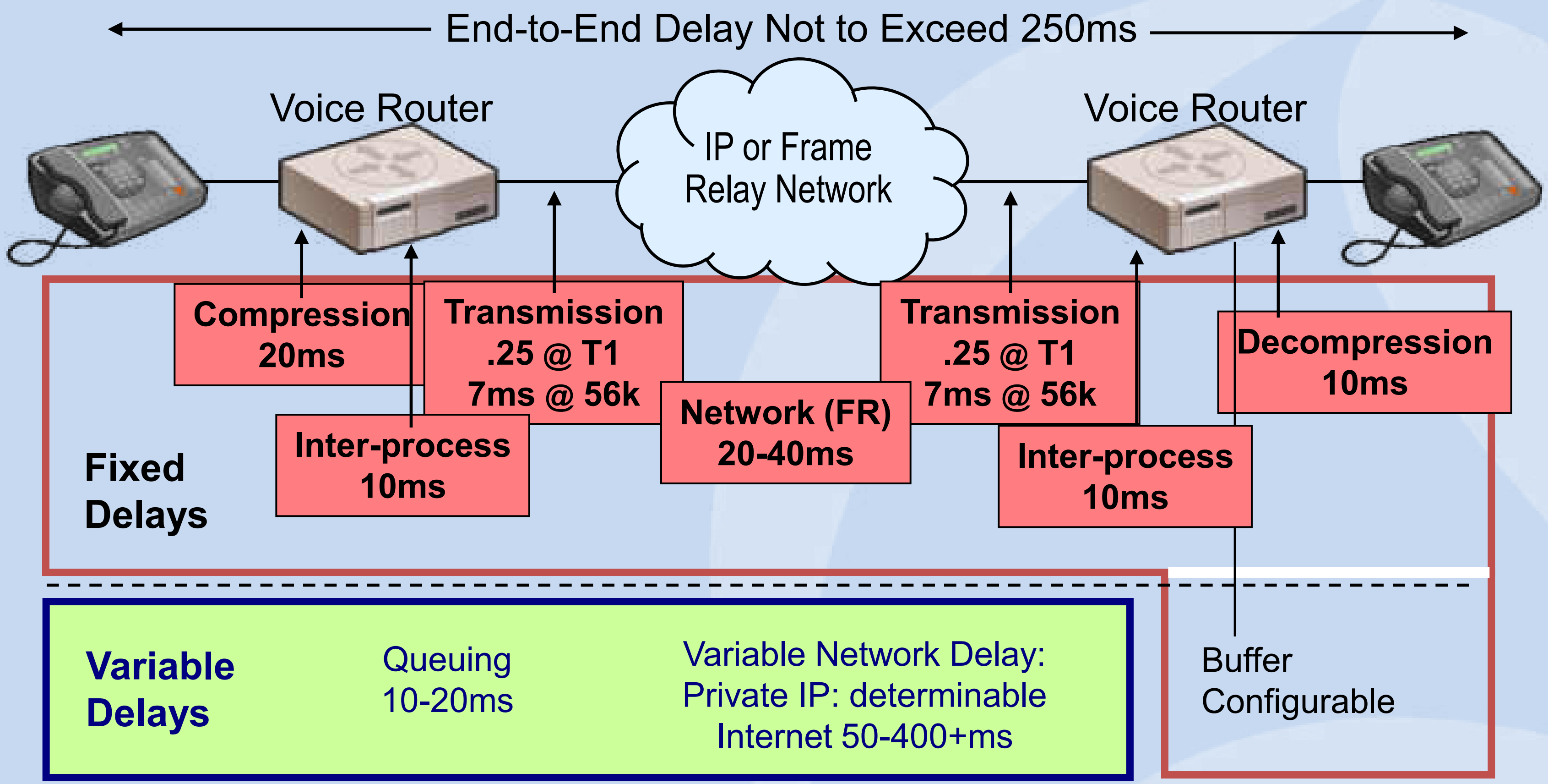


Many factors that contribute to the overall delay are fixed:
-Codec delay
-Hardware delay
-Processing delay
-Network physical delay

However, several delay factors are variable:
-Queuing delay
-Network propagation delay

It is the sum of all of these factors that determines overall delay as shown in the chart to the left

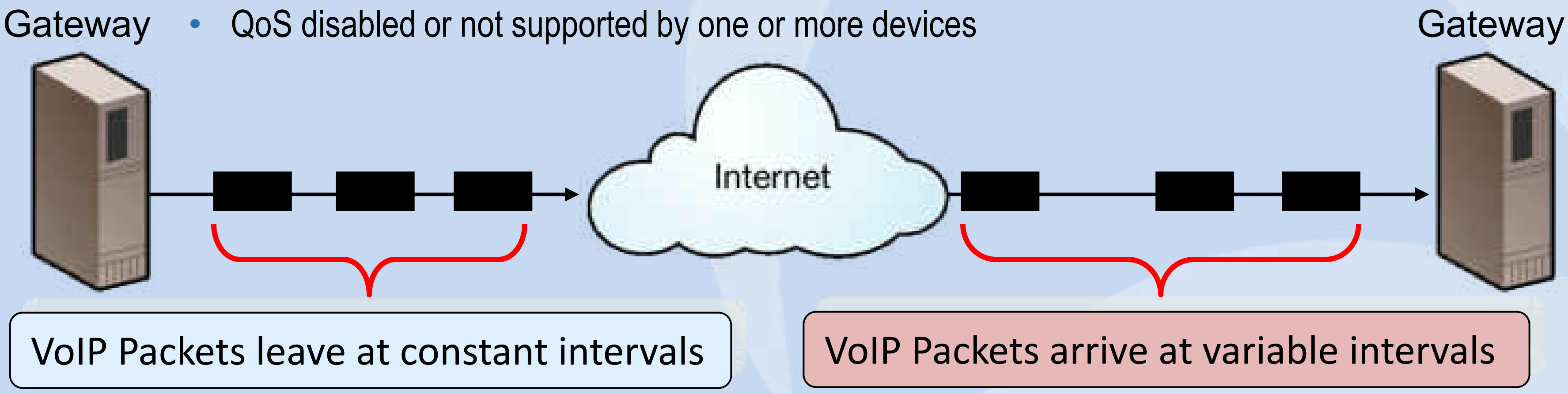
VoIP Delay Example



Total Fixed Delays (w/o buffer) 71-129ms

The #1 Result of Excessive Delay - Jitter

- Occurs when packets do not arrive at a constant rate that exceeds the buffering ability of the receiving device to compensate for
  - Symptoms
    - Often noticed as garbles or a annoying screech during a conversation
  - Typical Causes
    - Insufficient bandwidth for the conversation
    - Excessive number of Hops in the signal path
    - QoS disabled or not supported by one or more devices

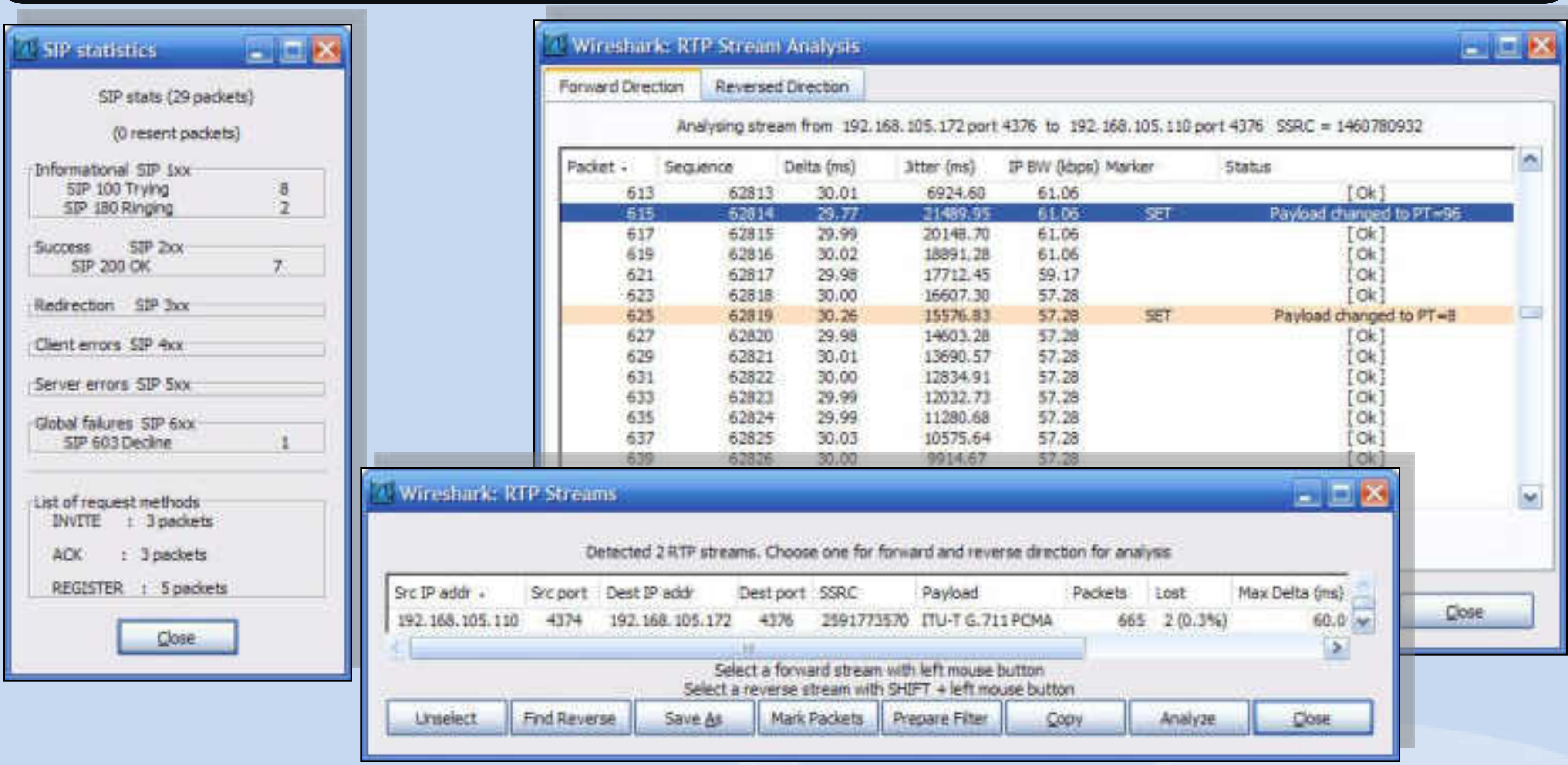


Customer Symptoms

- Customer Reported Symptoms
  - Cannot place or receive calls
  - Hear foreign voices not supposed to be on call
    - Cross-Talk
  - Volume noticeably low or high
  - Choppy Audio
  - Features do not work properly

- Equipment Alarm Indications
  - Ring Pre-trip Test Fails
  - Internal indications (card, power, etc)
  - Loss of Signal
  - High Error Rate
  - Connectivity failures

Analysis of Telephony Protocols



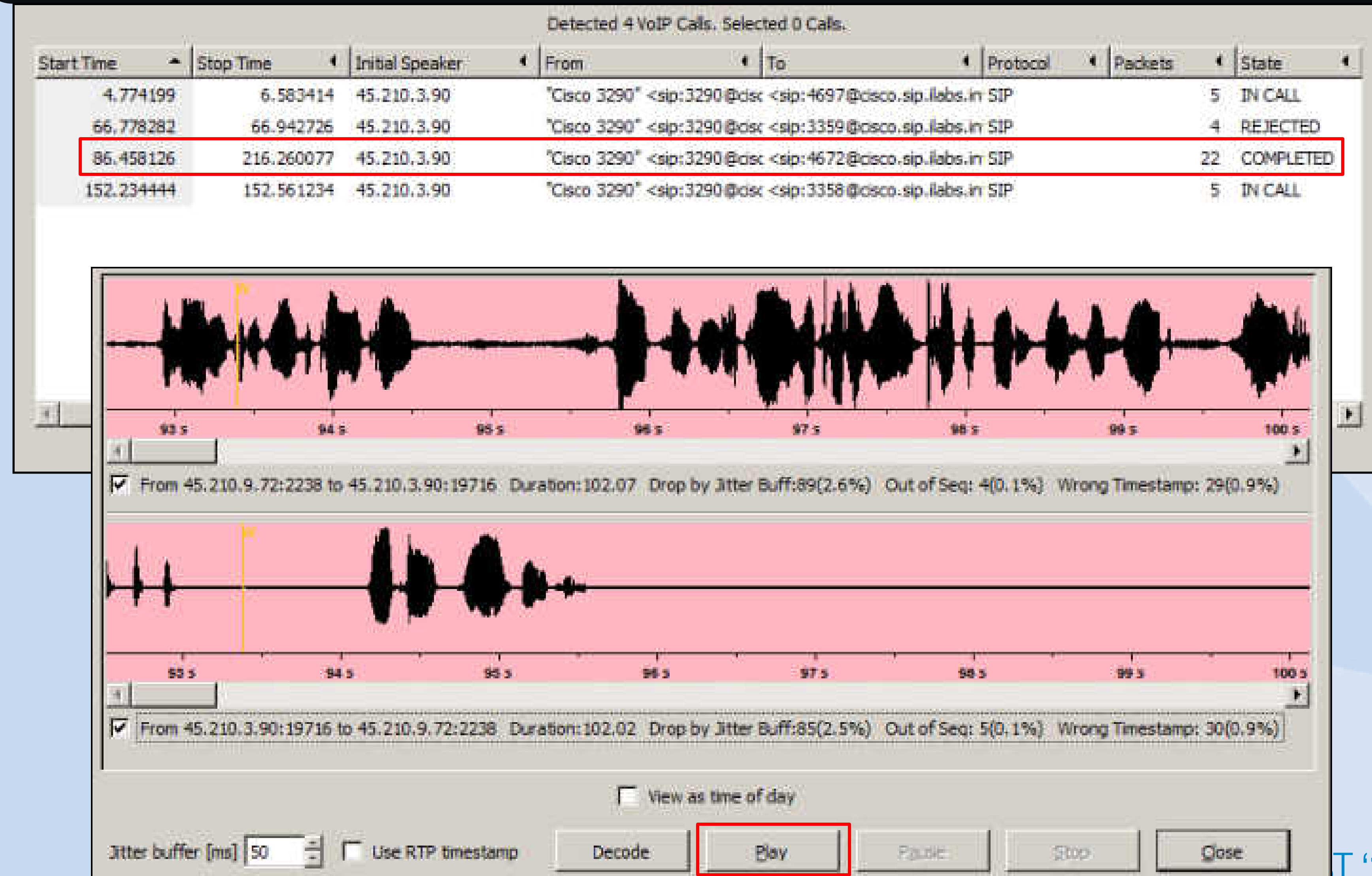
<u>VoIP Analysis Tip:</u> Wireshark has the ability to reconstruct not only VoIP conversations, but also other media streams for later analysis.

Packet Capture File

No.	IP - Src	IP - Dest	Time	Protocol	Length	Info
4	45.210.3.90	45.210.3.36	4.774198532	SIP/SDP	824	Request: INVITE sip:4697@c
5	45.210.3.36	45.210.3.90	4.774234772	SIP	390	Status: 100 Trying
6	45.210.3.36	45.210.3.90	4.855833054	SIP	556	Status: 180 Ringing
10	45.210.3.36	45.210.3.90	6.430492401	SIP/SDP	1078	Status: 200 OK , with ses
11	45.210.3.90	45.210.3.36	6.583414078	SIP	603	Request: ACK sip:3290.a75
12	45.210.9.97	45.210.3.90	6.616043091	RTP	214	PT=ITU-T G.711 PCMU, SSRC=
13	45.210.9.97	45.210.3.90	6.634405136	RTP	214	PT=ITU-T G.711 PCMU, SSRC=
14	45.210.3.90	45.210.9.97	6.648046493	RTP	214	PT=ITU-T G.711 PCMU, SSRC=
15	45.210.9.97	45.210.3.90	6.655860901	RTP	214	PT=ITU-T G.711 PCMU, SSRC=
16	45.210.3.90	45.210.9.97	6.675859451	RTP	214	PT=ITU-T G.711 PCMU, SSRC=
17	45.210.9.97	45.210.3.90	6.675891876	RTP	214	PT=ITU-T G.711 PCMU, SSRC=
18	45.210.3.90	45.210.9.97	6.687984466	RTP	214	PT=ITU-T G.711 PCMU, SSRC=
19	45.210.9.97	45.210.3.90	6.695211410	RTP	214	PT=ITU-T G.711 PCMU, SSRC=
20	45.210.3.90	45.210.9.97	6.707969665	RTP	214	PT=ITU-T G.711 PCMU, SSRC=
21	45.210.9.97	45.210.3.90	6.714948654	RTP	214	PT=ITU-T G.711 PCMU, SSRC=
22	45.210.3.90	45.210.9.97	6.728021622	RTP	214	PT=ITU-T G.711 PCMU, SSRC=
23	45.210.9.97	45.210.3.90	6.734687805	RTP	214	PT=ITU-T G.711 PCMU, SSRC=
24	45.210.3.90	45.210.9.97	6.748052597	RTP	214	PT=ITU-T G.711 PCMU, SSRC=
25	45.210.9.97	45.210.3.90	6.754869461	RTP	214	PT=ITU-T G.711 PCMU, SSRC=

This example contains four (4) calls and is from a VoIP network using Cisco phones and SIP signaling with G.711 audio codec

VoIP Call Detection, Analysis and Playback



Thank You !