VoIP - Implementing the New Phone System

A presentation for the IEEE SF ComSoc

http://www.sonoma.edu/users/k/kujoory/ in "Technical Documents"

Ali Kujoory, Ph.D.

Department of Engineering Science

Sonoma State University

ali.kujoory@sonoma.edu

Outline

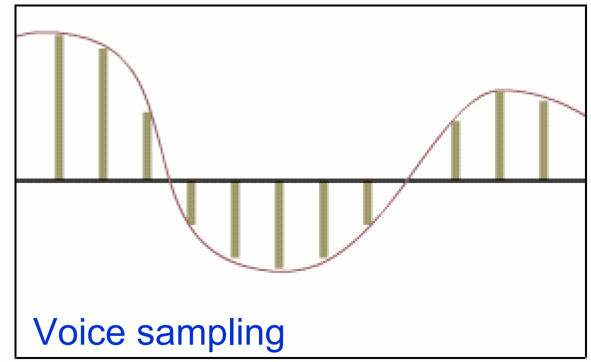
- How Does PSTN Work?
- What Is VOIP And Why?
- How Does Voice Differ From Data?
- How Is Voice Encapsulated in IP?
- How Is Voice Compressed / Encoded?
- What Are The Issues With VOIP?
- What Are The VOIP Signaling Protocols?
- What Are VOIP Security & Vulnerability Issues?
- What Are The Areas For Further Research?
- Appendix

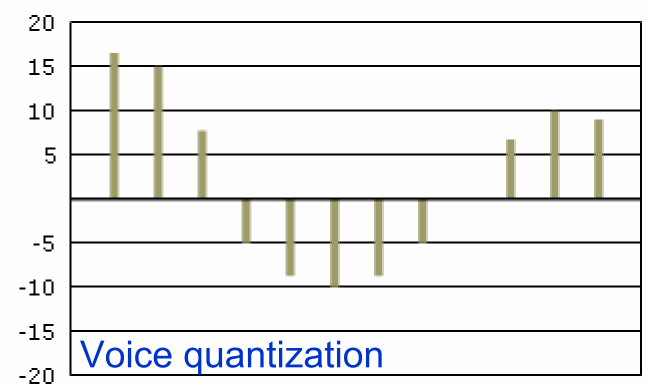
IP = Internet Protocol

PSTN = Public Switched Telephone Network

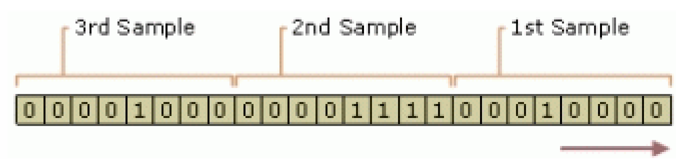
How Does The Traditional Telephony Work?

- Pick up the phone
  - Wait for a dial-tone
- Dial the destination tel. #
- Remote phone starts ringing
  - Caller is alerted of the ringing at the other side
- Destination picks up and
  - A point-to-point circuit is established
- The circuit carries a digitized version of the voice samples
- E.g., for PSTN, we use
  - One sample at a time
  - 8 bits/voice sample
  - 8000 samples/second
  - I.e., Bandwidth = 64 kbps


Voice sampling



Voice quantization



Digitized voice transmission

Are The Circuit and Bandwidth Used Efficiently?

- The circuit and bandwidth (64 kbps for PSTN) allocated to the call are devoted to
  - Only one conversation
- During the time between the digitized samples and the silent periods (when we pause)
  - The circuit is idle
  - Carries no information
- All this time is wasted
  - 50-70% of the bandwidth

PSTN = Public Switched Telephone Network

Saving the Bandwidth

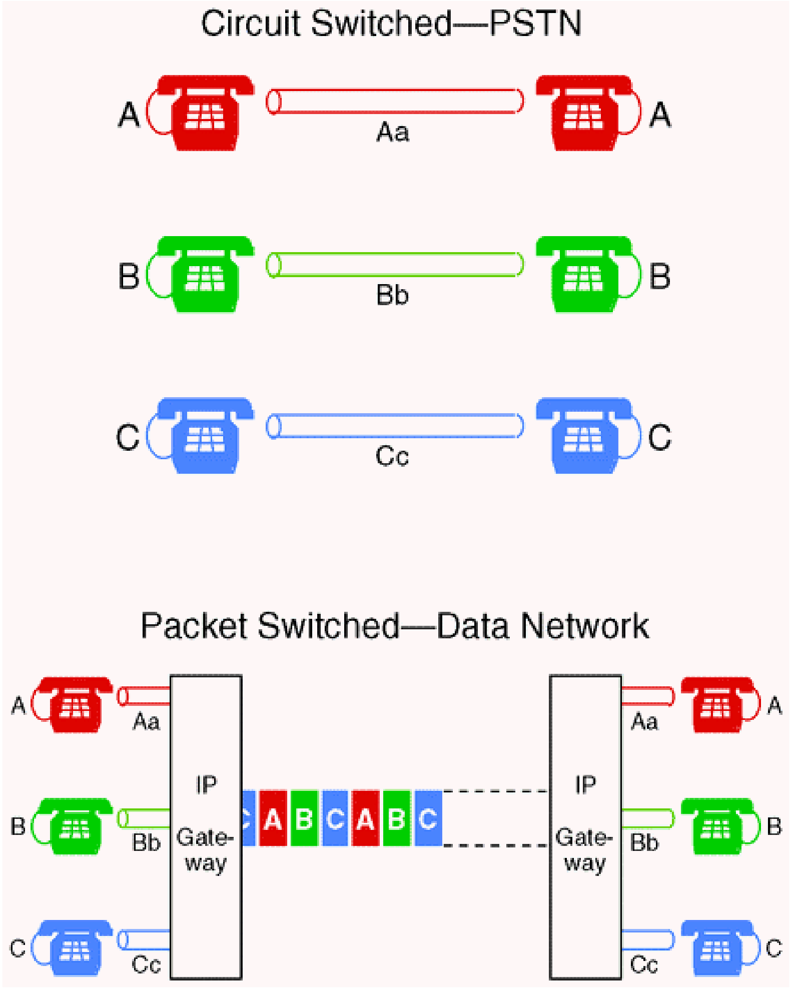
- How can we reduce the BW requirement?
  - Two ways
    - Accumulate several digitized voice samples together and send them in a packet
    - Also, signal the silence periods instead of sending them
  - Additionally, one can
    - Compress the digitized voice samples to save more BW
- What do you do with the BW that was saved over the network?
  - Use the BW for other calls or applications
    - A big advantage
- But do we lose anything?
  - What are the issues?

BW = bandwidth

What Is VOIP?

- A technology for transmitting ordinary telephone calls over a packet-switched network, Internet
  - Also called IP telephony
- VOIP works through sending digitized voice samples in packets

- Advantages to user
  - Cheaper - avoids the tolls charged by ordinary telephone service
- Advantages to service provider
  - Shares the BW among many users
  - Shares the network among voice, data, and video applications
  - Makes telephony cheaper
- Any issues / impairments?



PSTN = Public Switched Telephone Network

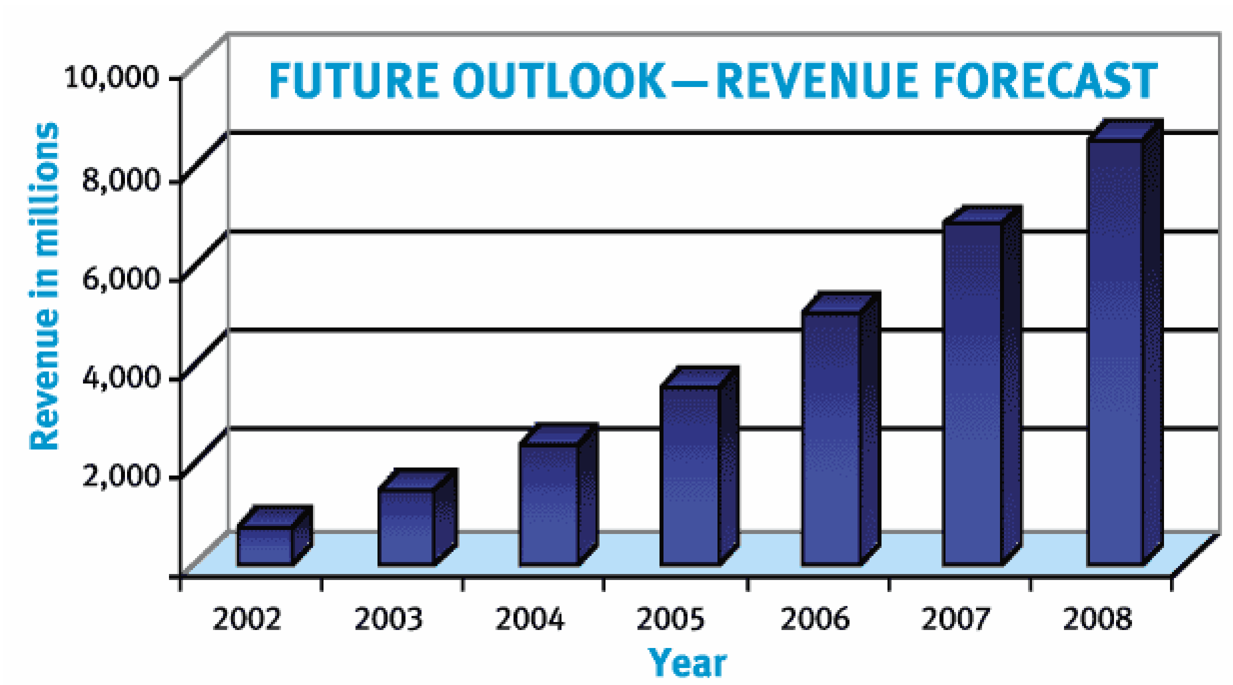
Why Network Operators Are Interested In VOIP?

- PSTN is based on TDM
  - Used primarily for voice traffic
  - Carrying little data directly
- Data and voice traffic have used separate networks
- Data traffic volume has passed voice volume since early 2000
- Makes sense to carry the voice over the data network
  - Amount of additional BW required for voice is minute with respect to data

- Using same network for voice, video, and data
  - Allows sharing resources
  - Reduces the CAPEX and OPEX
  - Allows service providers to quickly rollout new services
    - E.g., Instant Messaging
  - May allow mobility - convergence of wired and wireless services
- VOIP is believed to be much cheaper for subscribers
- Issues
  - VOIP should offer the same quality, reliability, and security as PSTN
    - Including requirements for lifeline, emergency 911, legal call tapping (CALEA)

CALEA = Communications Assistance for Law Enforcement Agencies
CAPEX = Capital Expense
OPEX = Operation Expense
TDM = Time Division Multiplexing

VOIP Future Outlook: Revenue Forecast *



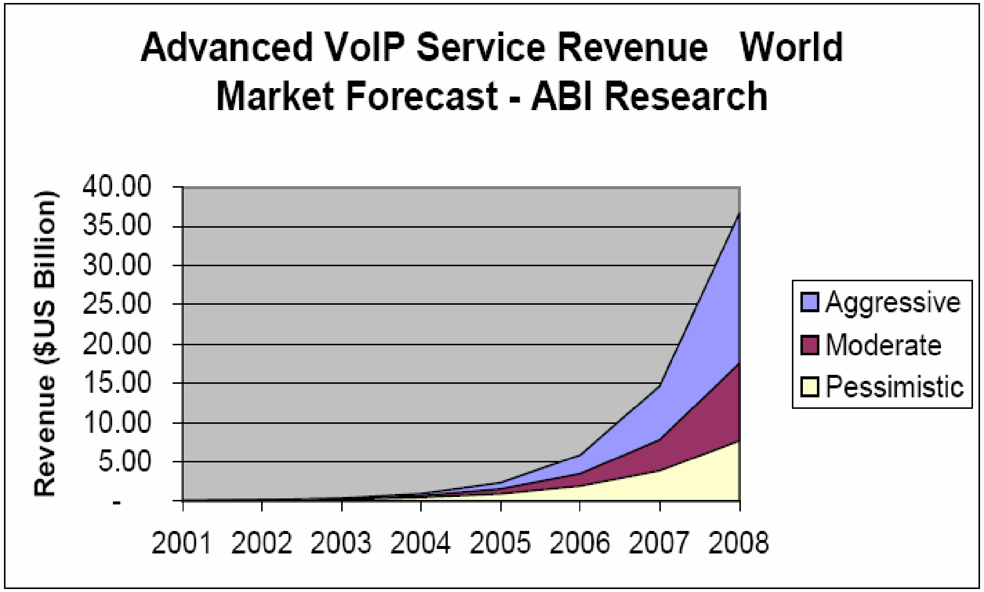
"By 2008, the world voice-over-IP equipment market is forecast to reach $8.5 billion. This represents a CAGR of 41.7%."*

* Telephony Online Intelligence for the Broadband Economy, February, 17, 2005.

CAGR = Compound Annual Growth Rate

VOIP Service Revenue

- VOIP has been picking up in the past few years
  - All service providers offer it
  - Commonly used for international calls
- Market forecasts for VOIP through 2008 vary sharply depending upon confidence in consumer uptake*
- The 'Moderate' case estimates, which exclude consumer uptake
  - Estimate the worldwide market at roughly $US 8 Billion
- If assumptions of consumer uptake are included
  - Then the market roughly doubles in the "Aggressive" case to $US 15 Billion*



Source: ABI Research

* "VOIP Market Forecast", Publisher: ABI Research, Pub Time: 2003/09, https://www.abiresearch.com/reports/IPS.html

How Does Voice Differ From Data?

Characteristics	Voice	Data
Bit rate	Continuous • Constant Bit Rate (CBR) • Queues must be small	Bursty • Variable Bit Rate (VBR) • Queues can be large
Tolerance to packet loss/errors	High • < 3 % acceptable packet loss • Uses UDP	Low • Value can change • Usually uses TCP
Tolerance to delay	Low (ITU-T Recomm. G.114) • < 100 ms desirable • 150-200 ms acceptable • 400-500 ms unacceptable by many • Meaning can change	High • May be acceptable • Unless for real-time data
Tolerance to delay variation	Low • Otherwise not acceptable • Cause distortion	High • Some delay variation acceptable

- Packetized voice over IP is VBR
  - **Challenge is how to smoothen VBR to CBR behavior**

Green = Desirable characteristics with respect to IP   Red = Undesirable characteristics with respect to IP
TCP = Transmission Control Protocol   UDP = User Datagram Protocol

How Is Voice Encapsulated In IP?

L2 Header (3-22)	IP Header (20 bytes)	UDP Header (8 bytes)	RTP Header (12 bytes)	Voice payload (Digitized voice samples) (20-160 bytes)	L2 Trailer (3-4)

- Assume that a call is established between source and destination
- How is the voice carried over the data network?
  - Voice is sampled and digitized
  - Digitized voice samples are encapsulated in RTP packets
    - RTP provides timing and sequencing
  - Each RTP packet is encapsulated in a UDP packet
    - UDP provides S/D port addresses
    - TCP is not used due to its latency effect
  - Each UDP packet is encapsulated in an IP packet for end-to-end transfer
  - An IP packet goes from one node to another in a Link Layer (L2) frame
    - E.g., Ethernet, PPP, Frame relay

Note

- Voice payload comprises one or more frames of encoded voice samples
- Voice payload and some headers may be compressed
- Voice packets are generally sent with priority against other packets
- Can end-to-end priority be guaranteed over the Internet?
  - This is an issue
  - Need a managed network
    - Control access to network resources

PPP = Point-to-Point Protocol
RTP = Real-Time Protocol
S/D = Source and Destination
TCP = Transmission Control Protocol
UDP = User Data Protocol

How Is Voice Compressed / Encoded?

- Legacy telephone channel for voice is
  - 64 kbps as specified by Nyquist's Theorem
- Can we compress the voice, to reduce the BW? YES
- However compressing/decompressing introduce additional delay
- A number of standard algorithms offer
  - Acceptable voice quality and delay
- Trade-off between these standards is
  - Quality versus BW
  - The higher the voice quality, the more BW is required
- How good voice sounds is
  - A subjective opinion
  - Difficult for people to describe what sounds good to them, specially
    - When delay and echo are present

Common ITU-T Recommendations For Voice Encoding	**BW (Kbps)**	**Codec Delay (msec)**
G.711 PCM (Pulse Code Modulation)	64	0.125
G.726 ADPCM (Adaptive Differential PCM)	32	0.125
G.728 LD-CELP (Low-Delay Code Exited Linear Prediction)	16	0.625
G.729-CS-ACELP (Conjugate Structure Algebraic CELP)	8	15
G.723.1 ACELP	5.3	37.5

Codec = Voice Coder-Decoder
ITU-T = International Telecommunications Union - Telecommunication Standardization Sector

What Are The Sources Of Delay?

- Voice encoding/decoding
  - E.g., 0.125 msec for G.711 encoding
  - Voice compression/decompression
- Accumulation
  - To collect several voice samples in a voice frame before encoding
  - In turn several speech frames may be collected in a voice packet
  - Size depends upon format
- Processing
  - Time to process the packets, RTP/UDP/IP/L2 encapsulations
  - Negligible for a fast processor
- Queuing
  - Buffering time in the gateways and routers along the path
  - Can vary
- Transmission
  - Time to transmit the packets over the links along the path
  - The lower the link capacity, the larger the transmission delay, e.g.
    - 100 bytes at 10 Mbps => 0.08 ms
    - 100 bytes at 2 Mbps => 0.4 ms
    - 100 bytes at 64 Kbps => 12.5 ms
  - The larger the packet size, the larger the transmission delay
- Propagation
  - Propagation of voice packets across the media - air, fiber, wire
  - 1000 Km fiber at 2/3 c => 5 ms
  - The longer the distance, the larger the propagation delay

c = Speed of light RTP = Real Time Protocol

How Can We Deal With Delay Variation?

Problems

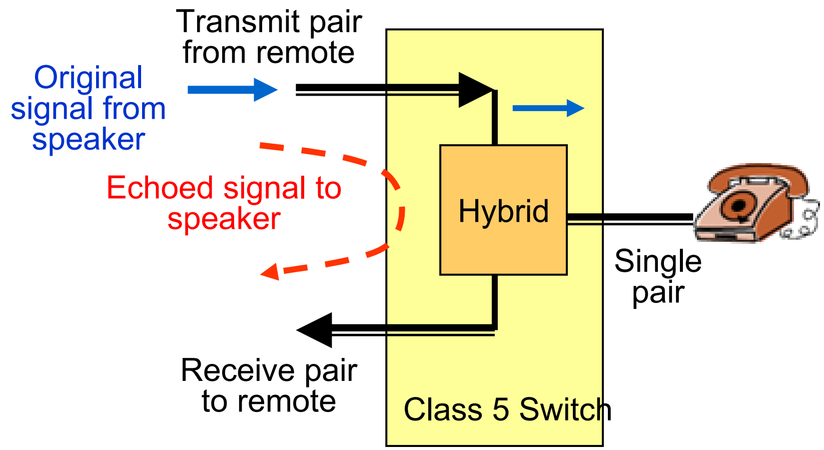
- Some delay components are variable
  - E.g., queuing delay
- IP network offers a best-effort service
  - I.e., packet loss
- Voice quality may suffer from the delay variation and packet loss in IP networks
  - Quality of Service (QoS) issues
- How can we handle these impairments?

Solutions

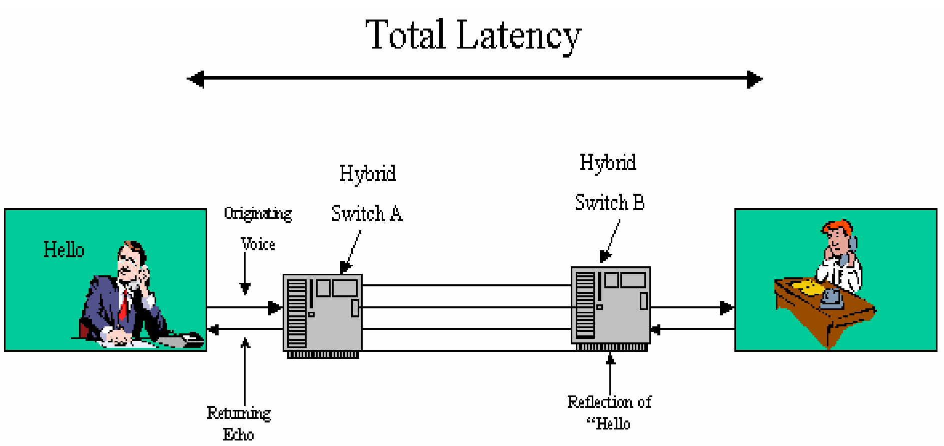
- For delay variation, the receiver buffers incoming voice packets for a period of T
  - T should be > expected delay variation
  - Then, voice samples are played out at a fixed rate
  - Note: T adds to the total delay that may not be acceptable
- For voice packet loss, the receiver may
  - Insert comfort noise, or
  - Play previous packet, or
  - Interpolate

What Is Echo?

- When voice signal passes from the 4-wire to the 2-wire
  - Some of the energy in 4-wire circuit is reflected back towards the speaker
    - Due to impedance mis-match
  - 4-wire is used since long-distance calls require amplification/direction
- The echo is acceptable
  - As long as round-trip delay is < 50 msec and is not too loud
    - As in majority circuit switched calls
  - Some echo is desired for talker
- Round-trip delays > 50 ms
  - Require echo cancellation
- In IP networks end-to-end delay > 50 msec due to additional delays
  - Echo cancellation is generally needed
  - An echo canceller is used in the gateway at each end



Speaker voice is echoed back toward speaker in the hybrid



Hello is echoed back from the hybrid in switch B

What Are The VOIP Signaling Protocols?

- Establish, modify, or terminate a session between the users
  - To participate in a call
- Three major VOIP Signaling Protocols:
  - ITU-T H.323 – a collection of protocols
    - Specifies “Packet Based Multimedia Communications System”
    - Currently most mature and is popular in enterprise networks
    - Supported by some vendors
      - Microsoft Netmeeting (MM conferencing) based on H.323
  - ITU-T H.248 = IETF MEGACO
    - Telephony signaling protocol based on existing PSTN
    - Upgrade of earlier MGCP
    - Currently supported by some VOIP vendors and service providers (Sonus Networks, Net2Phone)
  - IETF SIP
    - Text-based, Client-Server protocol for telephony applications over IP networks
    - Moves application control to the endpoints
    - Supported by some vendors and ISPs
      - Microsoft Windows Messenger, AT&T CallVantage, Vonage
- Multiplicity of protocols can result in
  - <u>Interoperability</u> issue

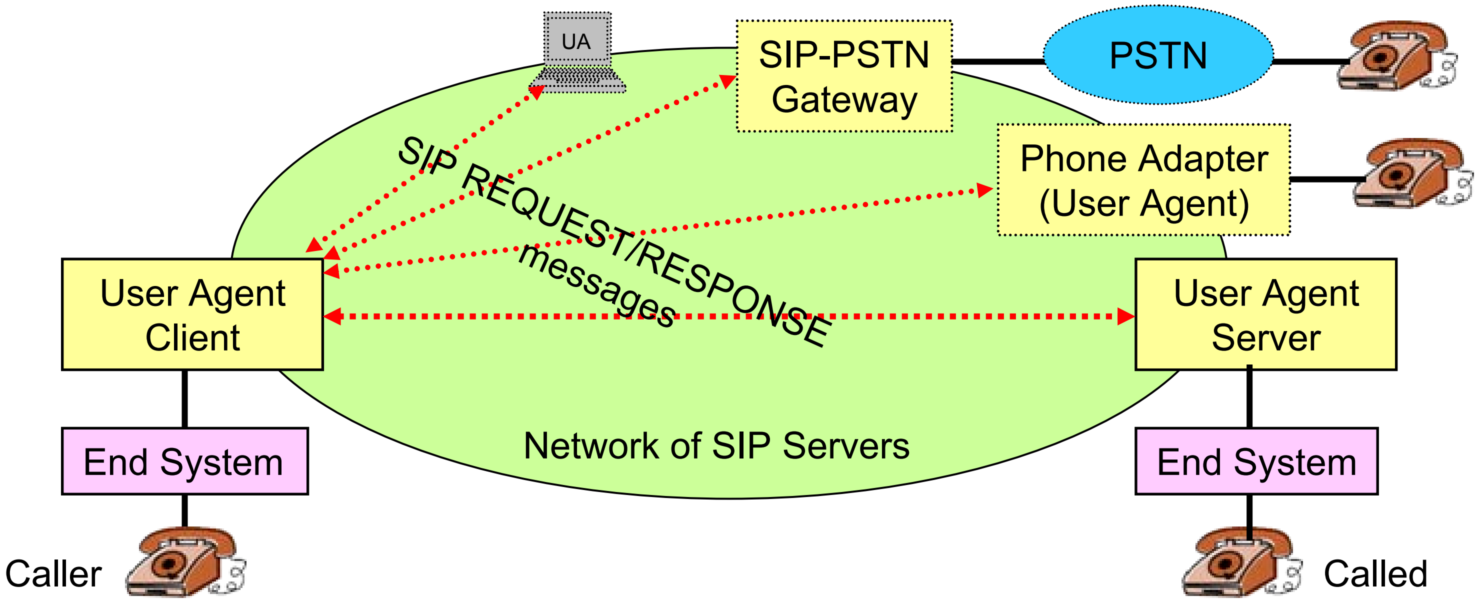
IETF = Internet Engineering Task Force
ISP = Internet Service Provider
MEGACO = MEdia GAteway Control Protocol
MGCP = Media Gateway Controller Protocol
SIP = Session Initiation Protocol

A Comparison Of VOIP Signaling Protocols

Item	**H.323**	**H.248/MEGACO (RFC 3015)**	**SIP(RFC 3265)**
Designed by	ITU-T	ITU-T and IETF	IETF
Compatibility with PSTN	Yes	Yes	Largely
Compatibility with Internet	No	Yes	Yes
Call signaling	Q.931 over TCP	Signaling over TCP or UDP	SIP over TCP or UDP
Media Transport	RTP & RTCP	RTP & RTCP	RTP & RTCP
Multiparty call	Yes	Yes	Yes
Multimedia conference	Yes	In future?	Yes w/ other protocols
Addressing	Host or telephone #	Telephone # or IP address	URL
Message encoding	Binary (ASN.1)	Text	Text
Implementation	Large and complex	Moderate	Moderate-Easy
Used by	Enterprise networks	Some telephony providers	Some ISPs and Telcos

ASN.1 = Abstract Syntax Notation
RTCP = Real Time Control Protocol is used to monitor the quality of RTP session
URL = Uniform Resource Locator

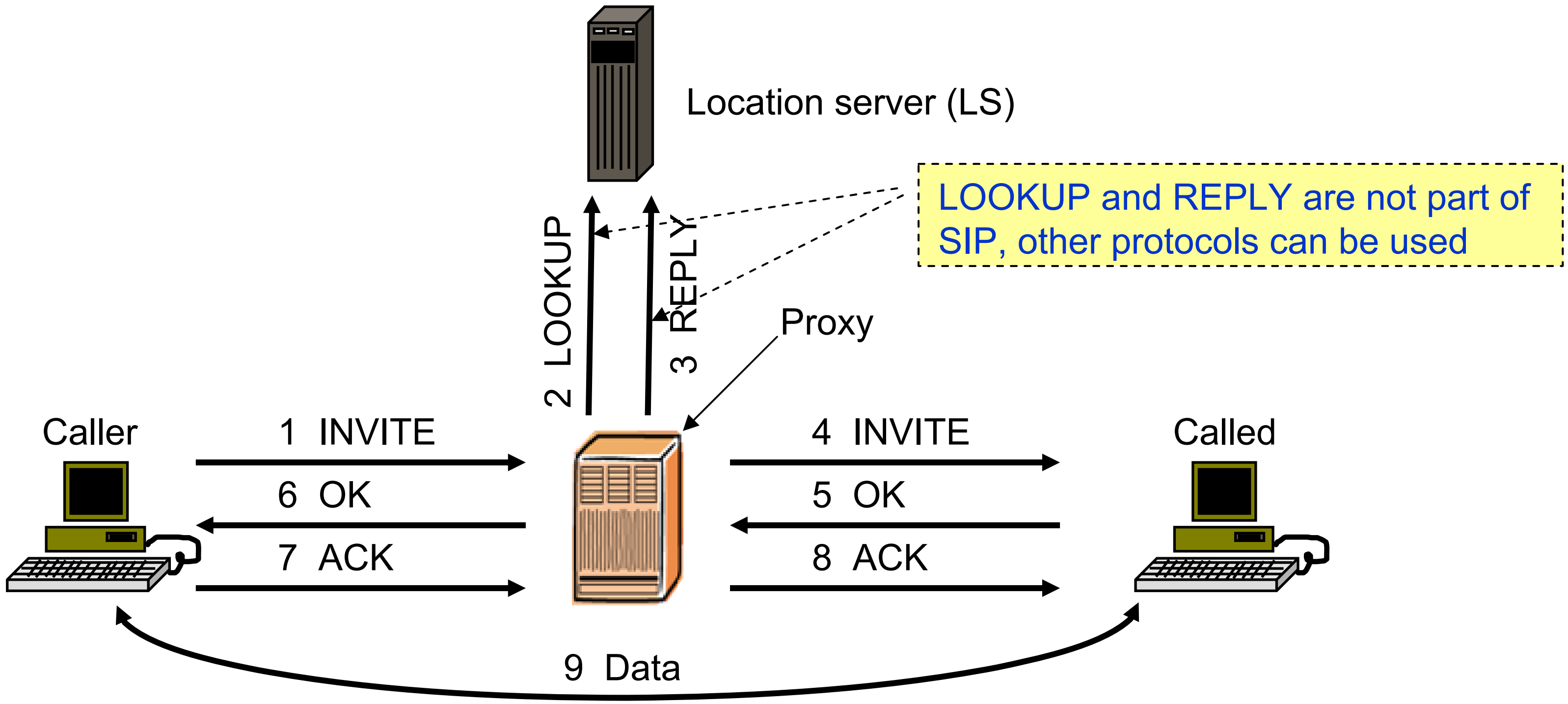
SIP Architecture Simplified



- User Agent Client (UAC)
  - Application which originates SIP requests
- User Agent Server (UAS)
  - Application which contacts user upon receiving SIP request, and
  - Returns user's response on his behalf
    - Accepts, rejects or redirects
- User Agent (UA)
  - Application which contains both UAC & UAS and exchange request/response messages

- UA is a piece of software that can be placed in a computer or a laptop
- Therefore, SIP can offer
  - Various telephony services, e.g.,
    - Internet phones-to-Internet phones
    - Internet phones-to-PSTN phones
    - PC phones-to-PC phones
  - Mobility option

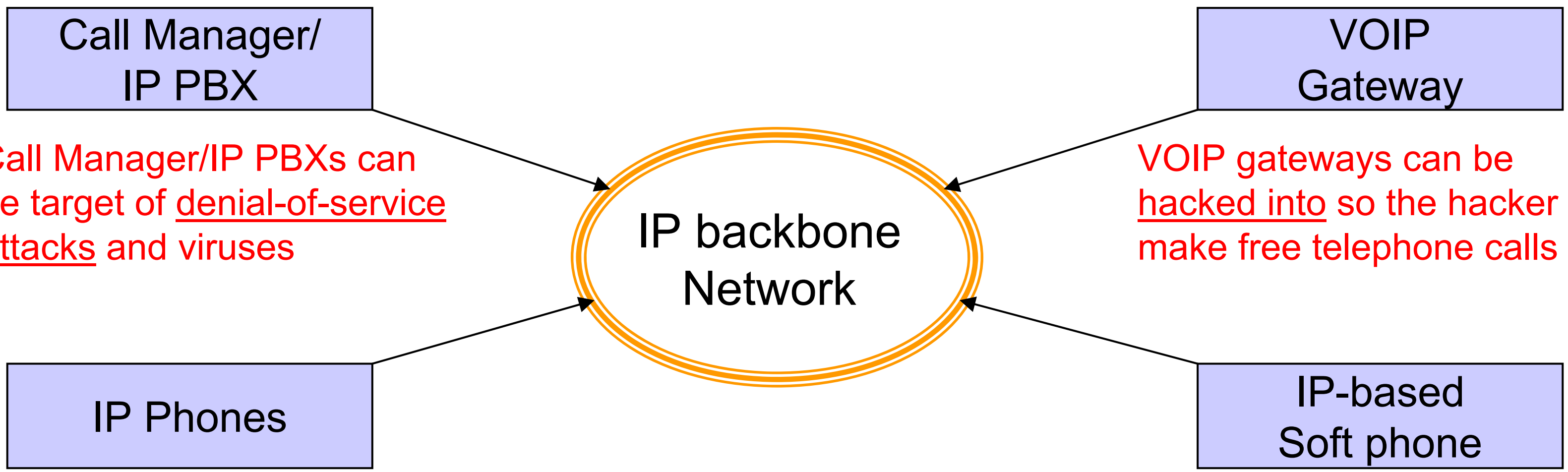
SIP With Proxy And Redirection Server



- The REGISTER method relates to SIP's ability to track down and connect to a user who is away (mobility)
  - This message is sent to a LS
    - LS keeps track of who is where
- Caller can send INVITE message to a Proxy server
  - To hide possible redirect operation
- Proxy consults LS to locate Called
- Proxy sends INVITE message to Called
- Proxy relays and performs 3-way handshake for the messages
- Two sides can start talking using RTP
- One side can terminate by using BYE

What Are VOIP Security & Vulnerability Issues?

- In addition to QoS, latency, interoperability, and redundancy issues
  - Security is a fundamental issue when we deal with the Internet
- VOIP is treated as data in the packet-switched network
  - Same security precautions are necessary for
    - Privacy, authentication, integrity, service denial
- What are the sources of vulnerabilities?



Call Manager/IP PBXs can be target of denial-of-service attacks and viruses

VOIP gateways can be hacked into so the hacker can make free telephone calls

Anyone connected to the IP network could, in theory, eavesdrop on voice calls if he/she had the right hacking and data-sniffing tools

The use of desktop-based soft phones could result in dangerous holes being punched in the firewall

PBX = Private Branch Exchange

Tips For Securing VOIP Traffic

- Encrypt VOIP traffic and run it over a VPN
- Make sure the firewalls are properly configured
- Check to see if your network and security vendors have
  - Support for SIP and H.323
- Consider segmenting voice and data traffic virtually by using a VLAN
  - To limit the threat posed by packet-sniffing tools, also
  - To minimize disruption in the event of an attack
- Think about using proxy servers in front of corporate firewalls
  - To process incoming and outgoing voice data
- Make sure that server-based IP PBXs are locked down and
  - Protected against viruses and denial-of-service attacks
- Make extra provisions for desktop soft phones
  - That initiate/receive VOIP-based calls
  - Do not allow punch holes in corporate firewall that hackers could exploit

VLAN = Virtual LAN
VPN = Virtual Private Network

Conclusion

- Legacy voice networks are converting to VOIP
- Issues under discussion
  - QoS, reliability, security, legal tapping, emergency 911, security
- Voice quality affected by delay, delay jitter, packet loss, echo
  - Mechanisms are needed to control these impairments
- Delay requirements for voice
  - < 100 msec is desirable
  - 150-200 msec is acceptable, but lower quality is noticeable
  - > 400 msec is unacceptable
- Echo > 50 msec requires echo cancellation
- Several signaling protocols
  - H.323, H.248/MEGACO, SIP
  - These will coexist for sometimes
  - The market decides the winner
- Areas for further research
  - Quality of service
  - Traffic and capacity engineering for triple-play
    - Integrated voice, video, data
  - Multicast
  - Security
  - Internetworking

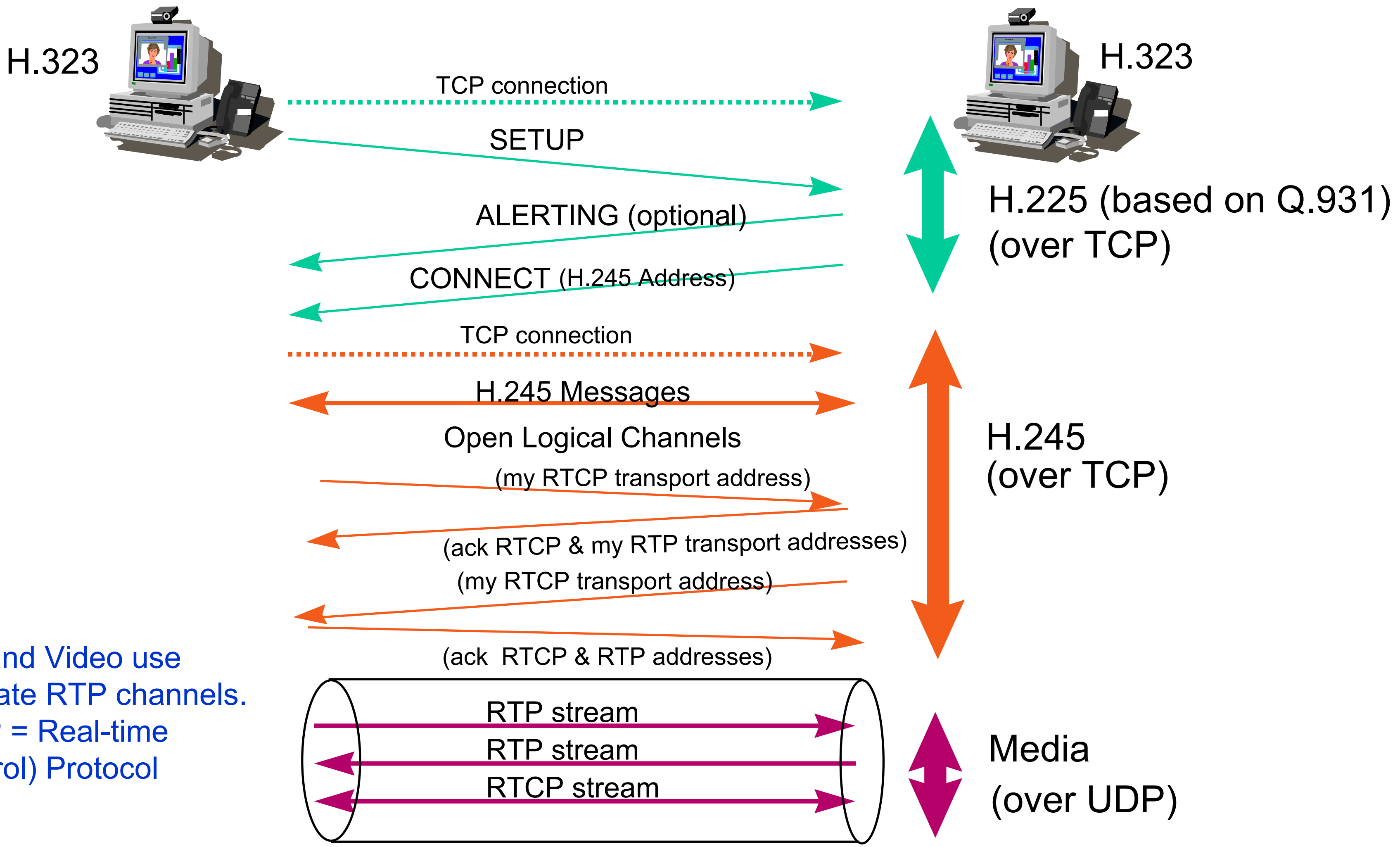
References

- ITU-T H.323, "Packet Based Multimedia Communications Systems," ITU-T, 11/2000
- RFC 2205, "Resource ReSerVation Protocol (RSVP) – Functional Specification," IETF, September 1997.
- RFC 3015, "MEGACO Protocol Version 1.0,", IETF, November 2000.
- RFC 3054, "MEGACO IP Phone Media Gateway Application Profile,"IETF, January 2001.
- RFC 3550, "A Transport Protocol for Real-Time Applications," IETF, July 2003 (Obsoletes 1889).
- RFC 3551, "RTP Profile for Audio and Video Conferences with Minimal Control," IETF, July 2003 (Updated 1890).
- RFC 3261, "SIP: Session Initiation Protocol," IETF, June 2002.
- RFC 3265, "Session Initiation Protocol (SIP)-Specific Event Notification," IETF, June 2002 (Updates RFC 3261).
- RFC 3435, "Media Gateway Control Protocol (MGCP)," IETF, January 2003 (Obsoletes RFC2705).
- "Computer Networks," Andrew Tanenbaum. 4th ed., Prentice Hall 2003.
- "IP Telephony Demystified" Ken Camp, McGraw-Hill Companies, 2003.
- "Voice-Enabling the Data Network: H.323, MGCP, SIP, QoS, SLAs, and Security," James Durkin, Prentice Hall, 2003.
- "Role of 3GPP/IMS in AT&T's RSOIP Network," AT&T whitepaper, V. 2.0, August 1, 2005.
- Sage Instruments has some informational whitepapers on voice measurements and impairments, http://www.sageinst.com/techPubs.html.

Appendix

- H.323 Call Exchange Scenario
- What Vendors Use Which Signaling Protocols?
- Emergency 911* & Legal Wiretap
- TCP/IP Concepts
  - IP
  - TCP
  - UDP

H.323 Call Exchange Scenario (Direct Call Model)



Audio and Video use separate RTP channels.
RT(C)P = Real-time (Control) Protocol

What Vendors Use Which Signaling Protocols?

Service Provider or Equipment. Vendor	SIP	MGCP or MEGACO	H.323	Notes
Net2phone		MGCP		
AT&TCallVantage & MCI Advantage	SIP			Has also gateway
Vontage	SIP			Phone adapter, mobility
GoogleTalk	SIP in near future			XMPP (eXtensible Messaging and Presence Protocol) for IM
Skype/eBay				P2P tech. like file sharing
Microsoft	SIP in IM / Messenger		H.323 for NetMeeting	XP Windows supports SIP
Cisco	SIP	MGCP		CallManager (SCCP, Skinny Client Control Protocol
Sonus & Vocaltec	SIP	MGCP/MEGACO	H.323	Uses gateway

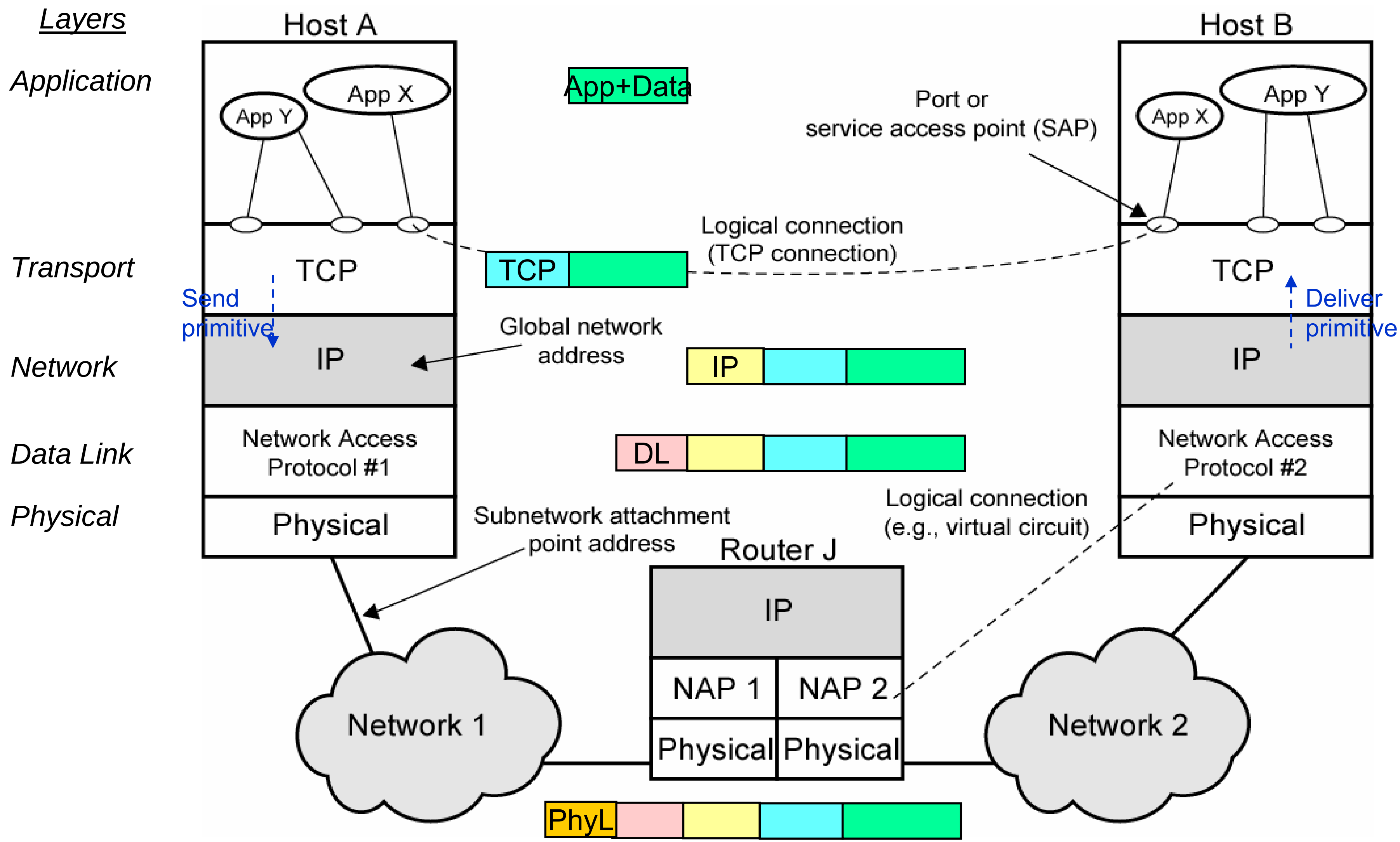
Emergency 911* & Legal Wiretap

- In May 2005, FCC required all providers of interconnected VoIP services to supply 911 emergency calling capabilities to their customers as a mandatory feature by Nov 28, 2005 Emergency
- Under FCC rules, interconnected VoIP providers must:
  - Deliver all 911 calls to the local emergency call center;
  - Deliver the customer’s call back number and location information where the emergency call center is capable of receiving it; and
  - Inform their customers of the capabilities and limitations of their VoIP 911 service

- August 5, 2005, FCC Required Certain Broadband and VoIP Providers to Accommodate Wiretaps
  - Because broadband Internet and interconnected VoIP providers need a reasonable amount of time to come into compliance <u>with all relevant CALEA requirements, the Commission established a deadline of 18 months from the effective date of this Order</u>, by which time newly covered entities and providers of newly covered services must be in full compliance

* "Interim VoIP Architecture for Enhanced 9-1-1 Services (i2)," National Emergency Number Association (NENA) VoIP-Packet Technical Committee, NENA Interim VoIP Architecture for Enhanced 9-1-1 Services (i2), NENA 08-001, Issue 1 December 6, 2005. CALEA = Communications Assistance for Law Enforcement Agencies

TCP/IP Concepts



TCP/IP Concepts (2)

IP (Internet Protocol)

- A connectionless Layer 3 protocol
- Allows the packet to route the network hop by hop
  - From source to destination
  - One at a time
  - Without any connection setup
- Packets carry full source and destination addresses
- Packets may choose different routes
  - Useful when a network node fails/congests
- Easier for internetworking among heterogeneous networks
- Provides a best-effort service
  - Not reliable (a problem)
  - Packet may get lost due to congestion

TCP/IP Concepts (3)

UDP (User Datagram Protocol)

- A connectionless Layer 4 protocol
- Takes packets from end-to-end
  - From one transport port to another
  - Quickly, low delay
    - Without any connection setup
    - Without acknowledgment (unreliable)
  - Good for real-time applications
    - Voice, video

TCP (Transmission Control Protocol)

- A connection-oriented Layer 4 protocol
- Takes packets from end-to-end
  - From one transport port to another
  - Reliably with delay
    - Need a connection setup
    - With acknowledgment
  - Good for non-real-time applications
    - File transfer, email, web browsing, signaling