

Applied Networks & Security

VoIP – with Critical Analysis

<http://condor.depaul.edu/~jkristof/it263/>

John Kristoff
jtk@depaul.edu

Critical analysis disclaimer

Following this disclaimer are slides used in other versions of the course. We *mark up* some slides using ~~strikethroughs~~ and underlined red in comic sans ms 20pt font. This is not meant to slight other teachers or their material. Much of the material is good and helpful so we use it.

We do this to explore complex issues, refresh dated material, correct inaccuracies and stimulate critical thinking. In some cases we are pedantic where it seems useful, but we are not exhaustive and try to avoid being overly tedious when it is unnecessary.

IT 263
Applied Networks and
Security

VoIP

Lecture 15 Outline

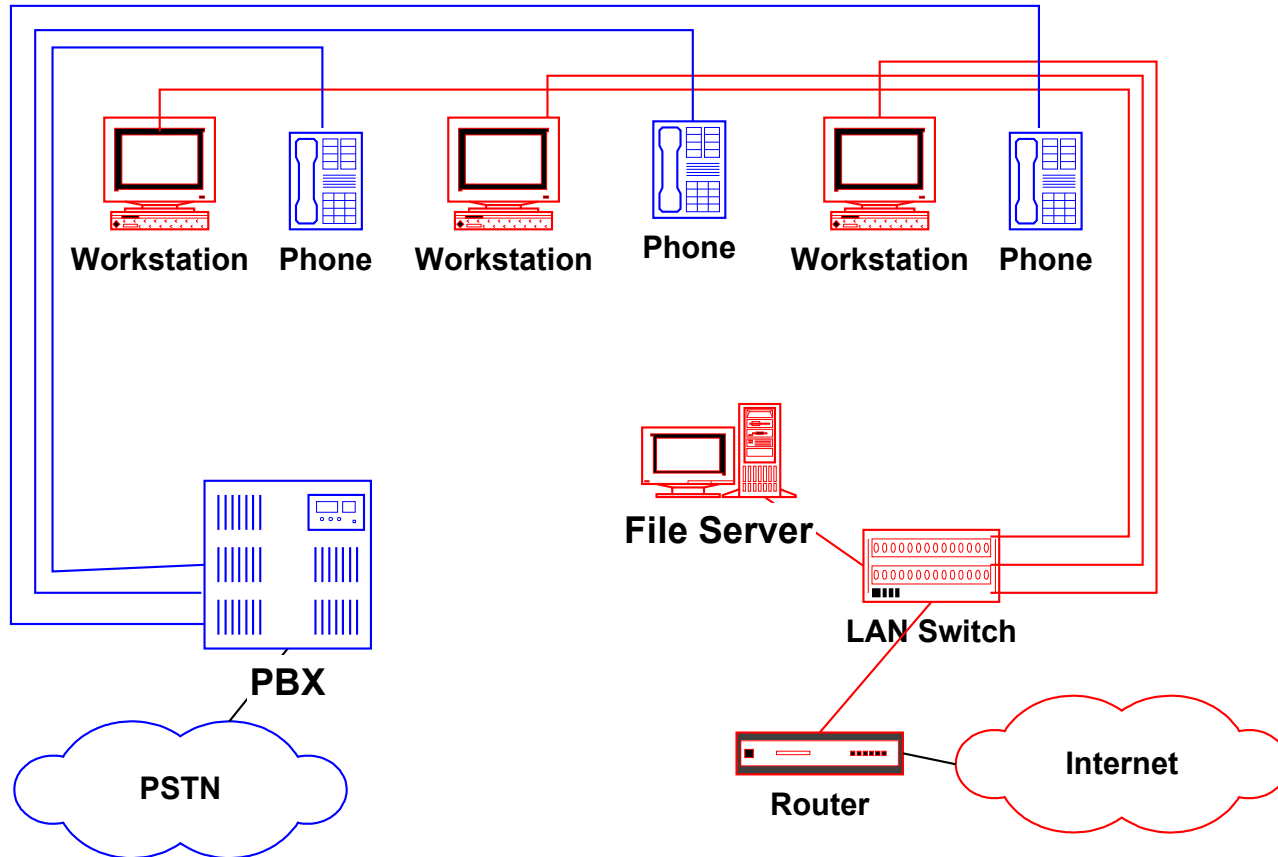
1. VoIP overview
 - a. Examples
2. Components
 - a. Terminals
 - b. Gateways and Gatekeepers
 - c. Multipoint Control Units
3. Packetized Voice
 - a. RTP
 - b. Network Performance Issues
4. Signaling
 - a. H.323
 - b. SIP

Why VoIP?

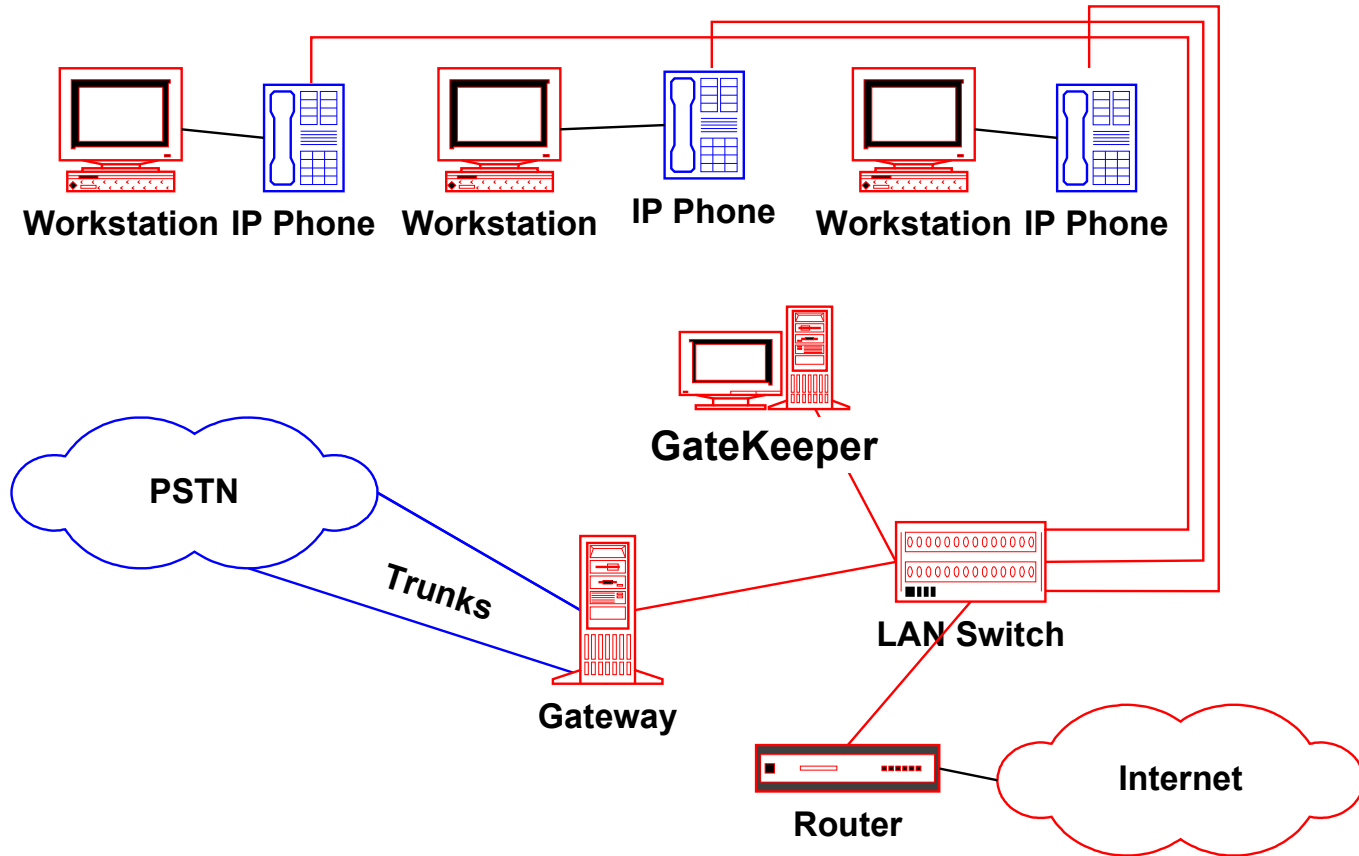
There should be a why not slide

- **More efficient use of existing networks** (can you prove it?)
 - Simpler cabling
 - Unified administration
- **New applications**
 - Click-through to Voice Call from Web
 - Improved Call Center Applications
 - Enhanced On-line Collaboration
- **Integration of voice/data on same transport** (ISP, Cable provider)
- **Savings on toll call charges** (toll bypass)
 - Local phone service charges
 - Long distance charges
- **Standards set by international organizations**
 - IETF (Internet Engineering Task Force)

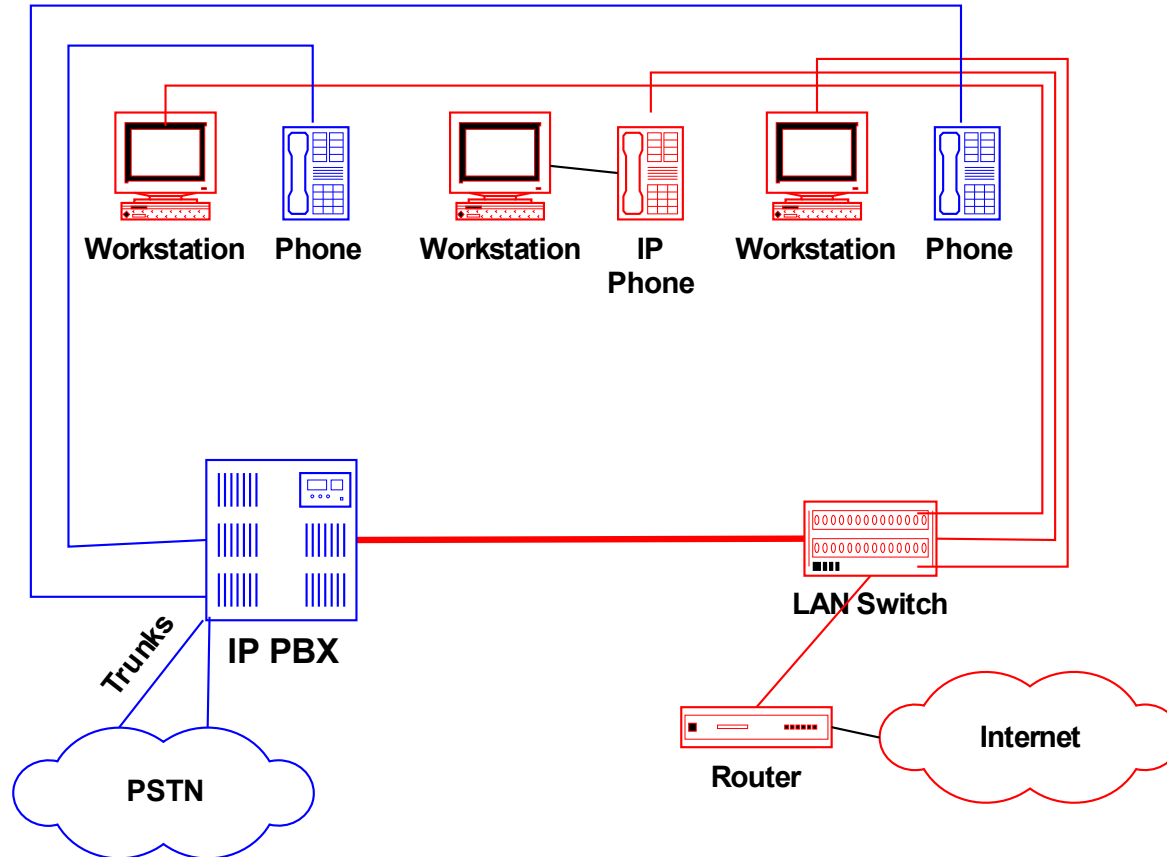
Traditional Data vs. Voice: Separate Cabling



Integrated Data & Voice: Unified Cabling



Mixed Environment with IP-PBX



VoIP Devices

- **Multimedia Terminals (IP phone, PC, etc.)**
 - Supports audio
 - May support video
- **Gateways**
 - Connects a VoIP network (such as a LAN) to a non-VoIP network, i.e. the **P**ublic **S**witch **T**elephone **N**etwork (PSTN)
 - May have telephony and data interfaces
 - Digitizes and packetizes audio

VoIP Devices

- **Gatekeepers** - provides centralized services:
 - Address translation
 - telephone number → IP address
 - Private dialing plans
 - Authorization and authentication of end terminals
 - Billing
 - Bandwidth management
 - Admission Control
 - Also may be called Telephony Server, Call

VoIP Scenarios and Devices

□ Peer-to-Peer Calling

- *Multimedia PCs* with appropriate software (i.e. MS NetMeeting) make direct connection
- Skype

□ IP Phone to IP Phone

- Requires *VoIP Gatekeeper* for telephone number lookup / translation to IP address and call management.

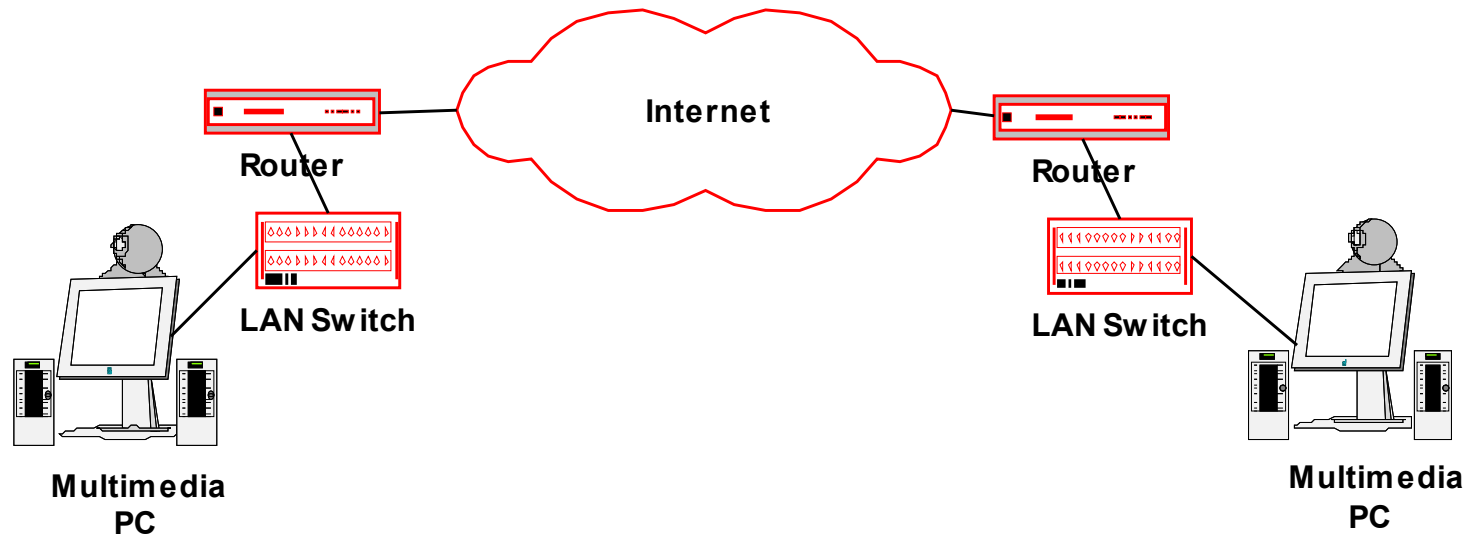
□ IP Phone to PSTN Phone

- Requires *VoIP Gateway* for translation between packet-switched voice and PSTN circuit-switched

Peer-to-Peer Calling

- First, let's consider a simple packet voice call across the Internet
- What issues need to be addressed to make this work?

Peer-to-Peer VoIP Calling

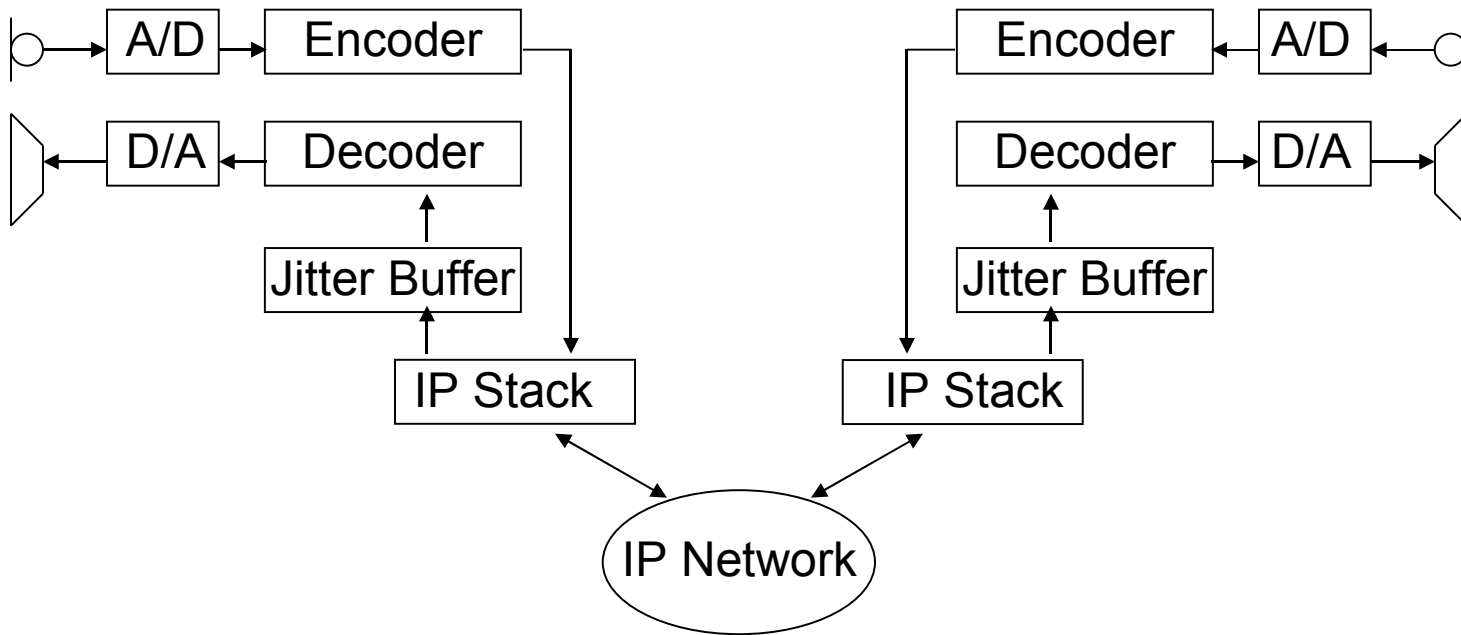


Packetized Voice Performance

- **Voice Coding and Playout**
 - Choosing voice coding method
 - Silence and activity detection
 - Ensuring sufficient bandwidth
- **Error Handling**
 - Error detection? Retransmission?
- **Network Delay**
 - Ensure reasonable network delay and jitter

Packet Audio

- In general, UDP is the transport protocol for multimedia



Voice Coding Standards

<i>Standard</i>	<i>Data Rate</i>	<i>Encoding Delay</i>	<i>Quality</i>
G.711 (PCM)	64 Kbps	0.1 ms	Excellent
G.727 (ADPCM)	32 Kbps	0.2 ms	Excellent
G.728 (CELP)	16 Kbps	5 ms	Good
G.729 (CS- ACELP)	8 Kbps	15 ms	Good
G.723.1 (vocoder)	5.3/6.3 Kbps	38 ms	Pretty Good

Voice Coding

□ Voice bit rate

- Quality decreases as bit rate decreases
 - Compressed channels cannot carry fax
 - Lowest bit rates can even obscure touch tones
- Lower bit rate increases coding delay
- Be sure to include protocol overhead when calculating total bandwidth required

□ Silence Detection can reduce bandwidth

- No bandwidth needed during silent periods
- Voice Activity Detection may need adjustment
- Comfort Noise Generation

Voice Packet Overhead

- Voice packets must be kept small to reduce packetization delay
 - Typically 30-40 bytes voice data in each packet
- Multiple packet headers required:
 - Ethernet header (for LAN transmission)
 - IP header (for routing to destination)
 - UDP header (for port numbers)
 - RTP header (for performance monitoring)
- So, packet contains more overhead

Voice Packet Overhead



- Rough Rule of Thumb:
 - Pkt Voice Bandwidth = (Voice Bandwidth * 3) !!!
- Example: to send G.728 (16 Kbps) voice, you should allocate approximately 48 Kbps of bandwidth per voice channel

Sources of Voice Quality Degradation

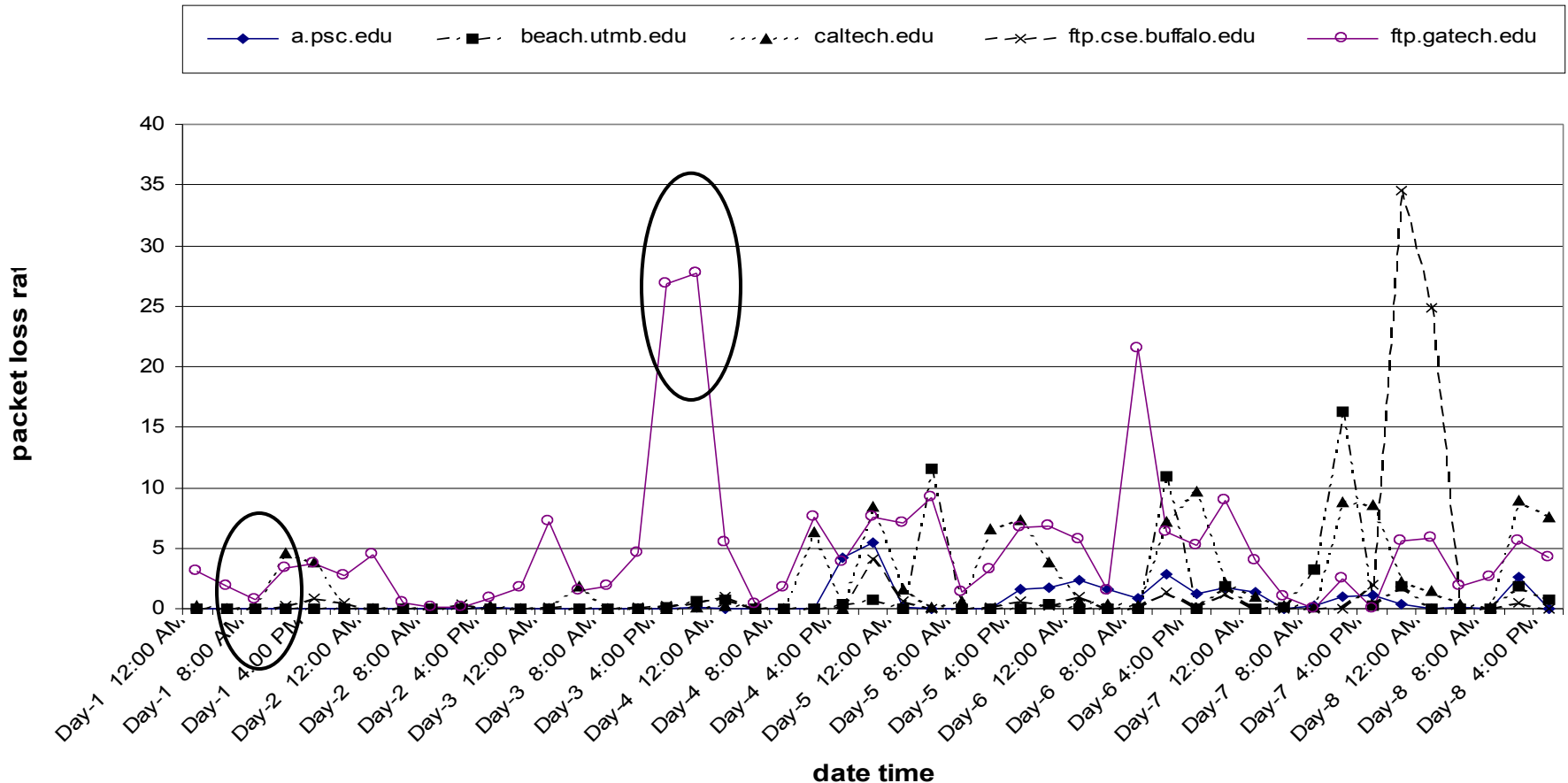
- IP Network introduces packet loss, delay and jitter (delay variation)
- Jitter buffers influence end-to-end delay and packet loss
- The delay should not exceed the maximum Mouth to Ear(M2E) allowable limit 400 ms
 - An ITU Standard

Network Delay Effects

<i>End-End Delay</i>	<i>Effect on User Conversation</i>
0-100 ms.	Fully interactive. No problems
100-200 ms.	Slightly perceivable delays, but few problems, if any.
200-400 ms.	Noticeable annoying delay, but somewhat normal conversation possible.
400-800 ms.	Interactivity is difficult. Speakers interrupt each other and must repeat information frequently.
Above 800 ms.	Fully interactive conversation is not possible. No longer "feels" like a telephone call at all.

Internet Packet Loss varies considerably

Figure 1: packet loss ratio - G.728 - U.S.
pkt size:720 bytes, #pkts sent:833, duration:5 mins

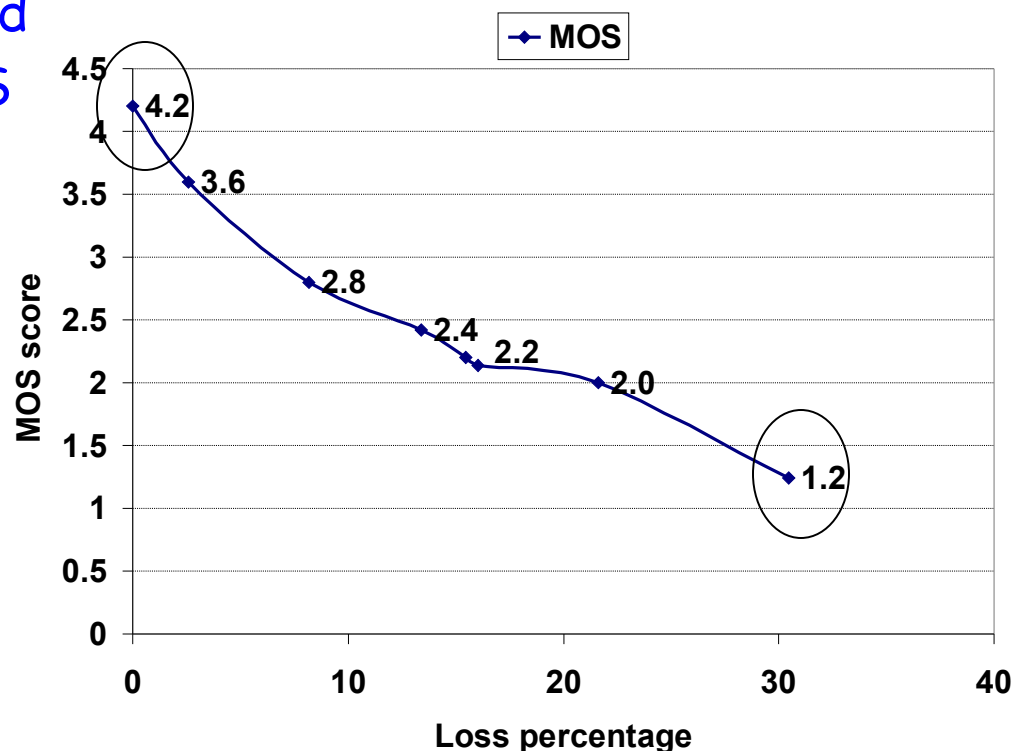


Loss Effect on Audio Quality

- Speech quality is described as a subjective score MOS (Mean Opinion Score - an ITU standard)
- Quality of an audio communication is highly sensitive to packet loss in

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

G.711 - audio quality vs. loss percentage



Real Time Protocol (RTP)

- RTP header includes the following:
 - Timestamp
 - Sequence Number
 - Coding Information
- RTP provides the following services
 - Monitors end-to-end delay and jitter
 - Provides resequencing of out-of-order data
 - Detects lost packets
 - Payload identification
 - Passes performance feedback from receiver back to sender (using Real-Time Control

VoIP Error Handling

- RTP allows us to detect lost packets.
- But what do we do about it?
 - Might be able to request and get retransmission from sender within jitter buffer playout time (but probably not)
 - Sender can transmit multiple copies of each packet
 - Receiver can't just ignore packet loss
 - receiver must play out some sound
 - White noise

Forward Error Correction

- ❑ Packet-retransmission may not be a viable option for real-time audio
- ❑ Packet loss for audio can be rectified using Forward Error Correction (FEC)

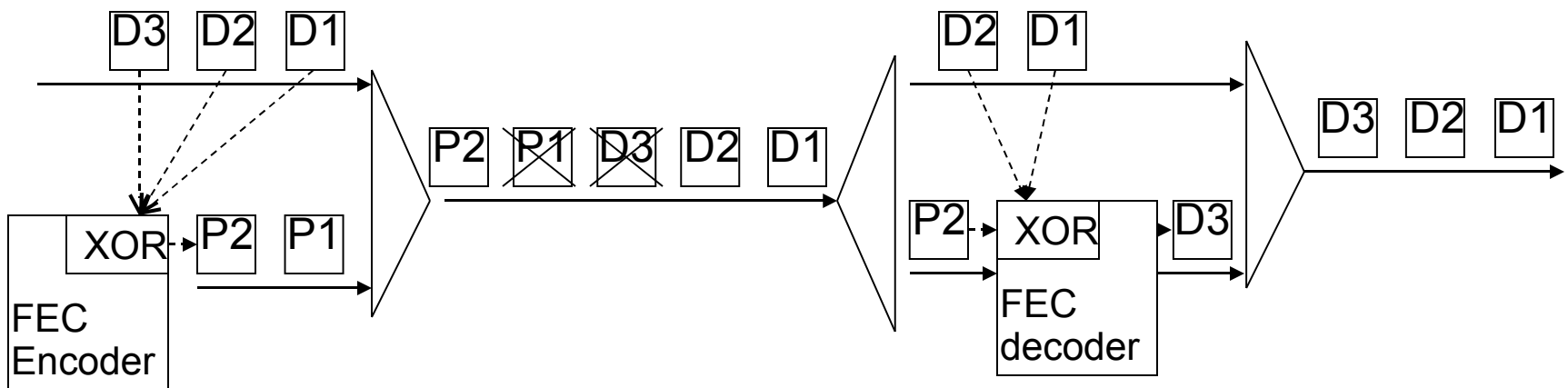


Fig. Operation of FEC

FEC Overhead

- ❑ But FEC has bandwidth and processing overhead
- ❑ Therefore, FEC should be used to maintain audio quality with optimal consumption of network bandwidth

VoIP Delay/Jitter

Objectives

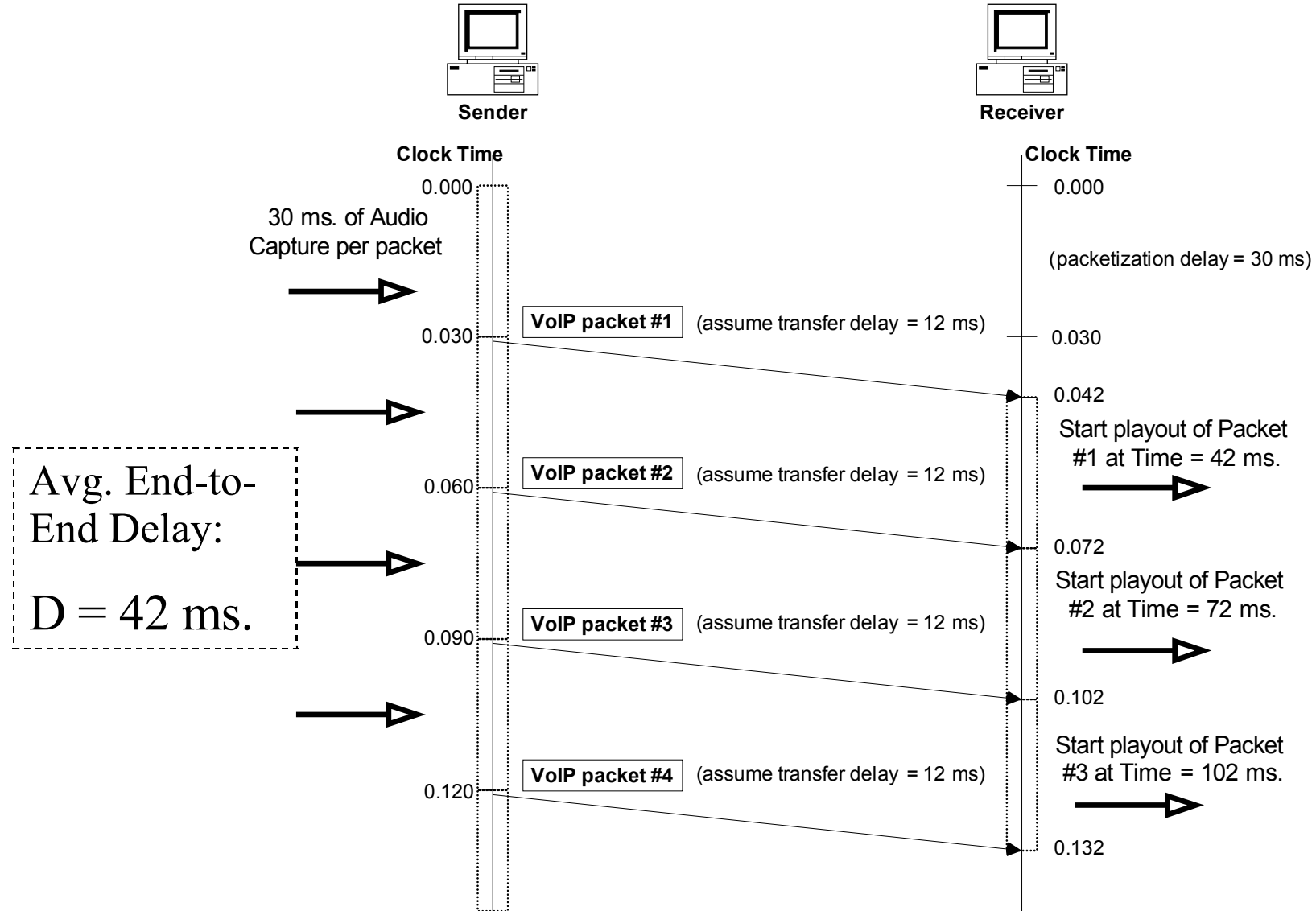
- For “toll-quality” VoIP (as good as wired telephone voice)
 - the end-to-end packet delay on VoIP networks should be 150 ms. or less
 - Delay is tolerated to 400ms for low bitrate codecs
 - The network jitter should be 10% or less (i.e. no more than 10% variation in transmission delays)

Jitter effect on Voice Quality

- A network with large delay and low jitter is perceived as providing *much better* voice quality than a network with low/medium delay but large jitter.
 - Reason: large jitter causes packet drops / disruptions in conversation.

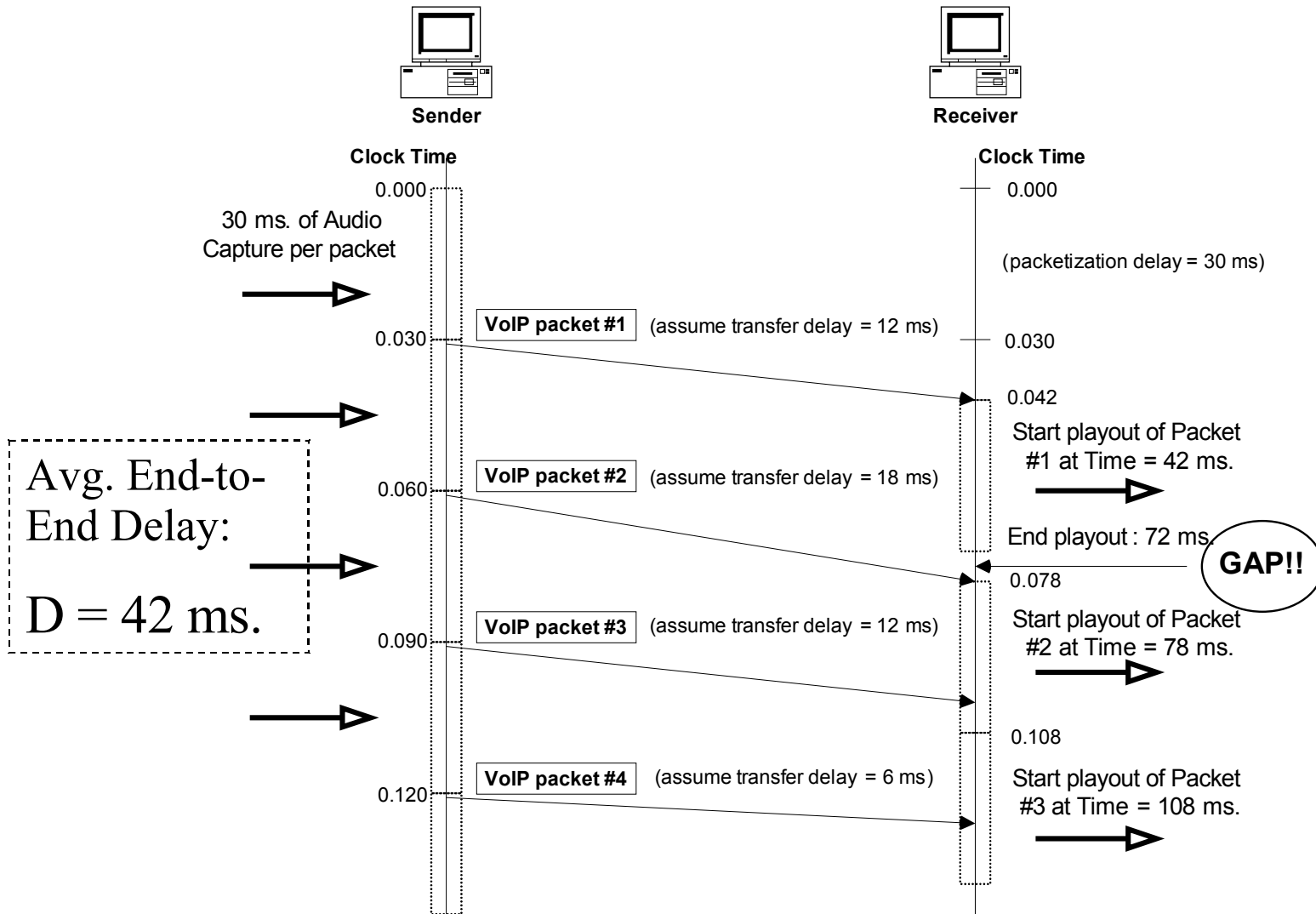
VoIP Playout

No jitter = perfect playout



VoIP Playout

6 ms. jitter but no receiver delay = Gap



The Jitter Problems

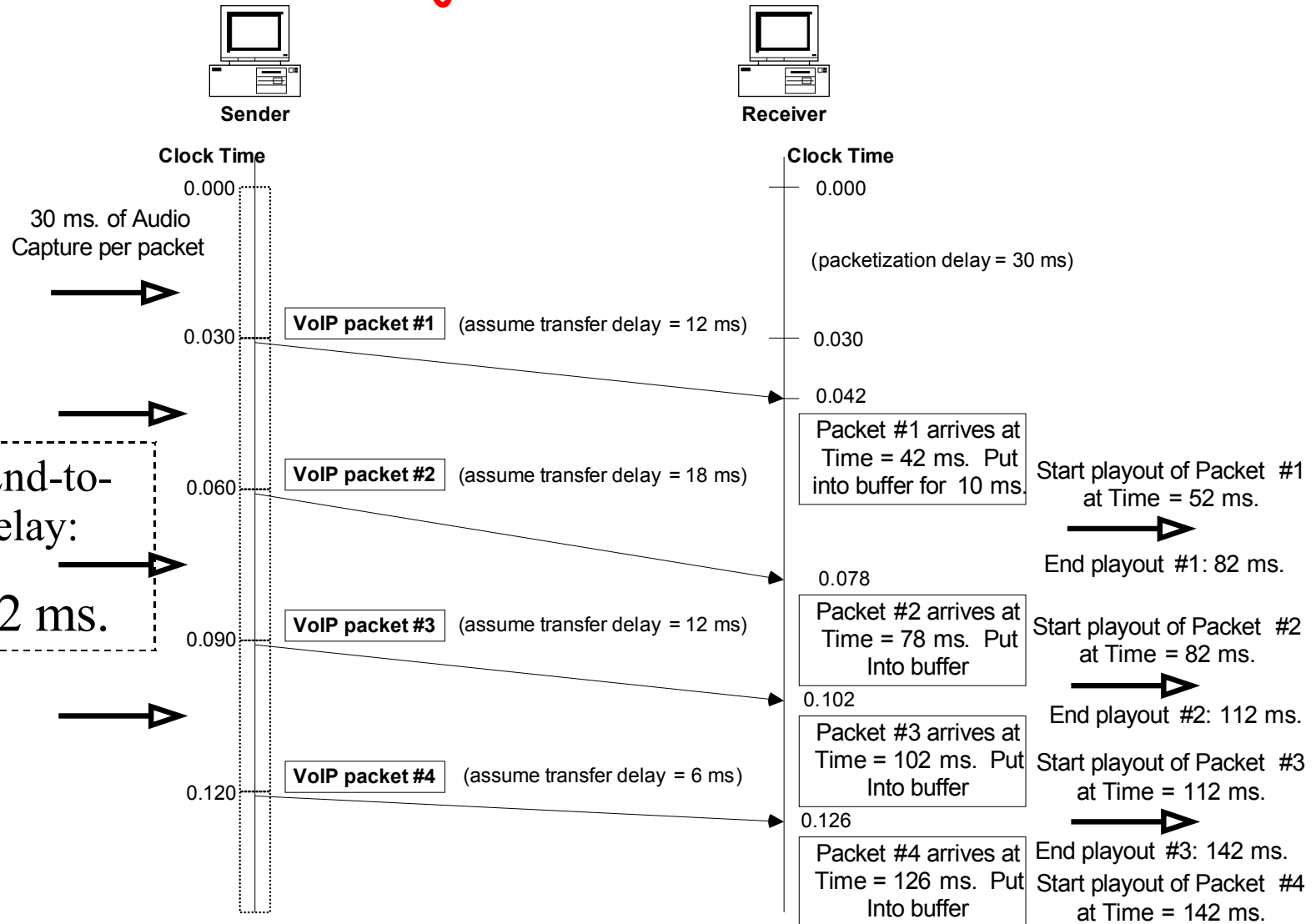
- Problem #1: If there's too much jitter, you may have gaps in audio playout at times when a packet arrives late.
 - Results can be improved by trying to intelligently compensate for the gap at the receiver by replaying/interpolating sounds.
- Problem #2: Jitter makes it hard to estimate how long you should wait for a packet at receiver before you assume it is lost.
 - If you wait too long, you delay the audio playout.
 - If you don't wait long enough, you miss packets.

The Playout Buffer Solution

- Receiver equipment assumes some maximum acceptable jitter time - call it **B**.
- First packet at receiver is put into a buffer and playout is delayed for time **B**.
- The second, third, etc. packets are added to buffer when they arrive and buffer plays out continuously.
- If packet is more than **B** ms. **late** in arrival, then it is assumed **lost** and the receiver inserts one packet time (for example: 30 ms.) worth of compensatory audio for listener.
 - Compensatory audio = repeat previous sound or interpolate what sounds should have been in missing

VoIP Playout

6 ms. jitter and $B = 10$ ms.



Playout Buffer Ads/Disads

□ Advantage of Playout Buffer

- Eliminates gaps: Every packet plays out correctly (no gaps) if all jitter is less than B .

□ Disadvantage of Playout Buffer

- Increases the end-to-end delay: End-to-end delay is increased by time B .

How to reduce jitter?

- *Provide higher priority to VoIP traffic compared with other traffic on the network.*
- *How to implement priorities?*
 - Ethernet priorities - IEEE 802.1p and 802.1q
 - IP precedence field
 - Multi-Protocol Label Switching (MPLS)
 - DiffServ point codes
- *Discussing these specific protocols is beyond the scope of this class.*

VoIP Signaling

- OK, so now we have discussed the issues in just sending packetized voice across a network.
- Let's move on to signaling.

VoIP Signaling

- What do we need signaling for?
 - Setting up calls
 - Destination address lookup
 - Telephone number → IP address translation
 - Indicating terminal capabilities
 - Call Routing
 - Offering call (ringing)
 - Accepting calls
 - Calling features - conference, transfer, etc.
 - Tearing down calls
 - Administration
 - Billing and auditing
 - Admissions control

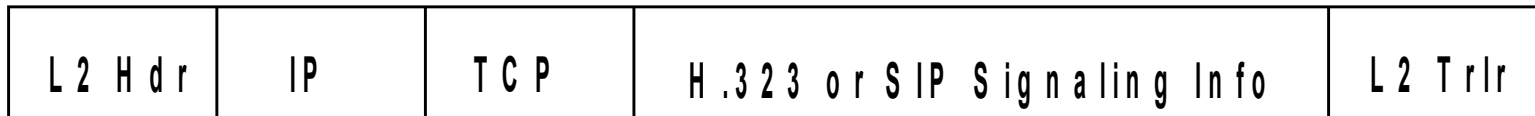
VoIP Signaling Packet Flows

- Terminal ↔ Gatekeeper
 - Signaling packets exchanged to do destination address lookup, admission control, and call administration
- Terminal ↔ Gateway
 - Signaling packets exchanged to set up, manage and tear down calls to non-VoIP devices
- Terminal ↔ Terminal
 - Signaling packets may be exchanged directly to set up, manage and tear down

Dueling Standards

- H.323 suite
 - Developed by ITU
 - Most mature and most widely implemented
- Session Initiation Protocol (SIP)
 - Developed by IETF
 - More efficient and “IP-friendly”

Signaling Packets



- Signaling packets are typically sent over TCP connections for reliability, but UDP is also an option.

Let's talk SIP

- **Session Initiation Protocol** is one of a set of protocols defined by IETF since 1999 to do VoIP call session signaling and management.
- SIP protocols:
 - Designed to be simple and efficient
 - Based on simple text commands and responses as with HTTP
- SIP model: places most call intelligence in end-user terminal rather than in the gatekeeper.

SIP Advantages

- ❑ SIP generally uses fewer messages and runs over UDP - so it is fast and lightweight
- ❑ SIP operations can easily be extended (new call features, etc.) by any competent programmer
- ❑ SIP servers keep no state information about calls. End user devices responsible for all call status.
- ❑ SIP designed to fully utilize existing Internet infrastructure (DHCP, DNS, IP routing, IP multicasting)

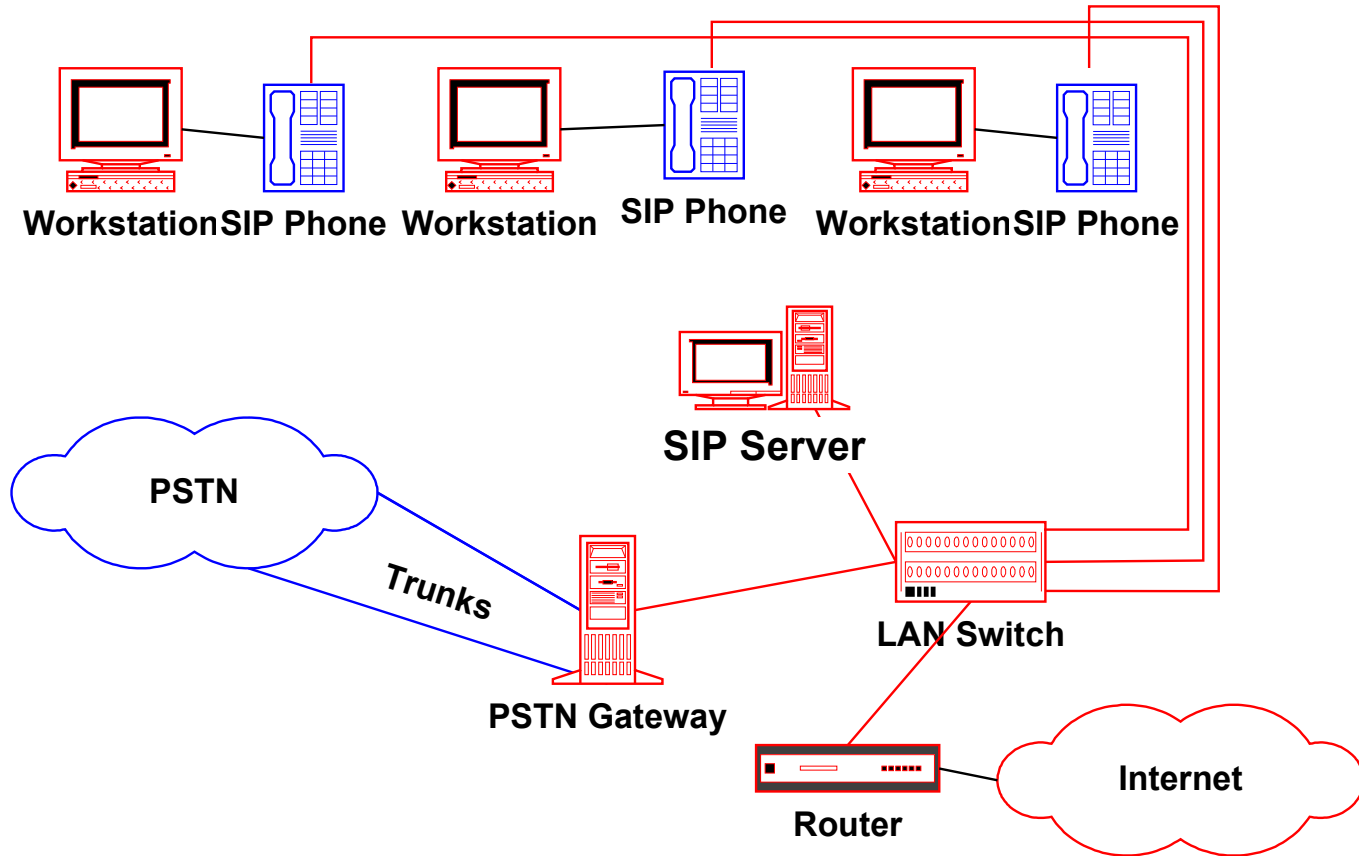
SIP Protocols

- Session Initiation Protocol (RFC 2543)
 - Signaling to set up call, transfer call, conference call, etc.
- Session Description Protocol (RFC 2327)
 - Signaling to determine what features are supported by another device.
- Session Announcement Protocol (RFC 2974)
 - Signaling to register presence of particular device/session on the network.

SIP Components

- **User Agent** is the user telephony application
 - User Agent Client (UAC) can initiate calls.
 - User Agent Server (UAS) can receive calls.
- **SIP Server** acts as central point to relay signaling information
- **SIP Registrar** is server that accepts registration messages
- **Location Server** stores user device

SIP Components



SIP Registration

- When SIP device first powers up, it registers itself with Registration Server.
- This information can then be used for many services:
 - VoIP call
 - Videoconferencing
 - Instant messaging

SIP Addressing

- SIP users are addressed by a SIP URL, in the format sip:user@host.
- This is very flexible
- Examples:
 - sip:lopa@roychou.depaul.edu
 - sip:lopa-roychou@depaul.edu
 - sip:Ext-25040@depaul.edu
 - sip:13123625040@chicago.il.us
 - sip:13123625040@pstn.org

SIP Calling

- If SIP User Agents already know each other's addresses, they just communicate with each other directly (via INVITE message peer-to-peer).
- Otherwise, INVITE message is sent to a SIP Server, which may operate in one of two modes:
 - Proxy Server mode: looks up destination and forwards INVITE
 - Redirect Server mode: looks up destination and

SIP Signaling Messages

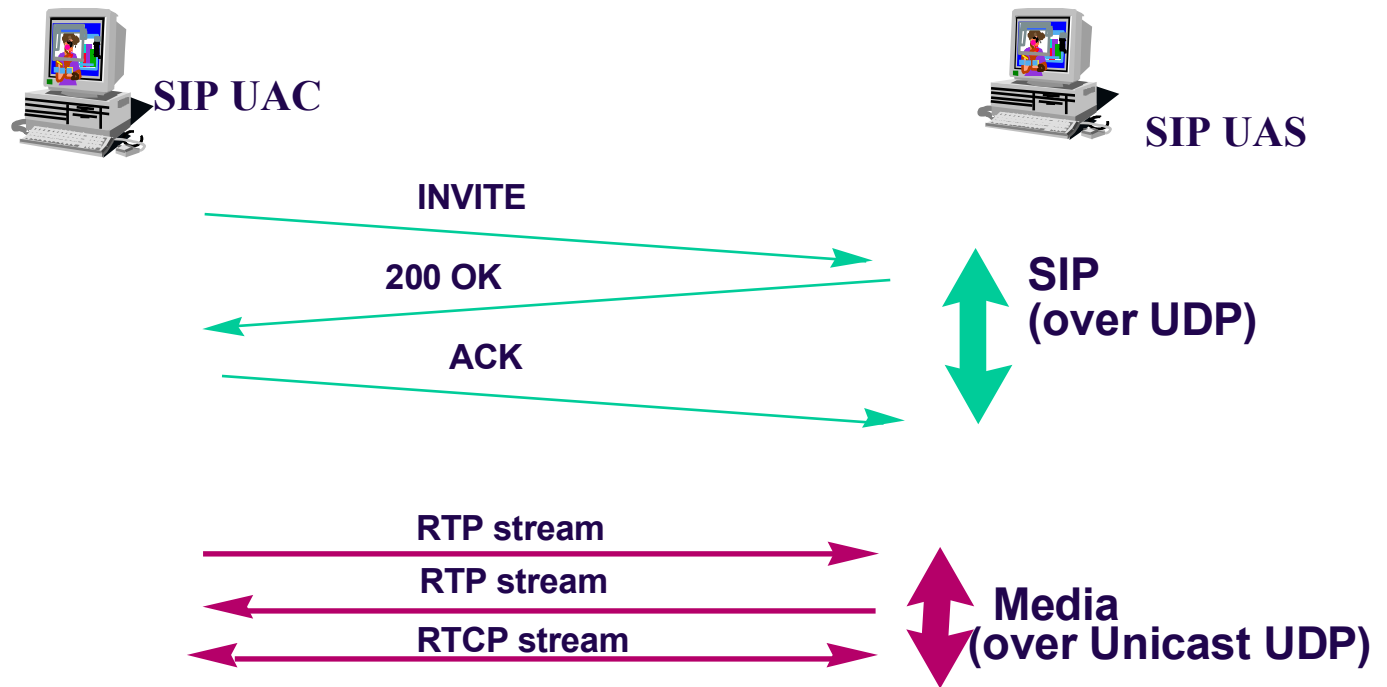
□ SIP Requests:

- **INVITE** - Initiates a call by inviting user to participate in session.
- **ACK** - Confirms that the client has received a final response to an INVITE request.
- **BYE** - Indicates termination of the call.
- **CANCEL** - Cancels a pending request.
- **REGISTER** - Registers the user agent.
- **OPTIONS** - Used to query the capabilities of a server.
- **INFO** - Used to carry out-of-bound information, such as

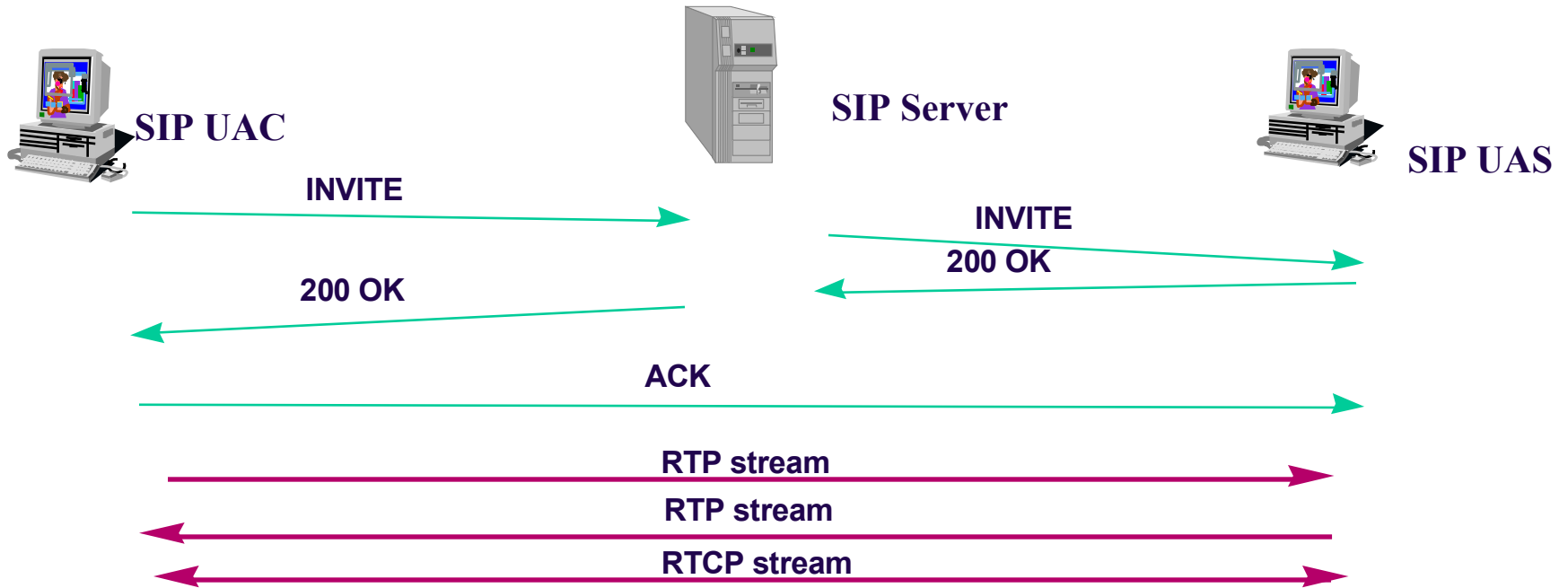
▪ SIP Responses:

- **1xx** - Informational Messages.
- **2xx** - Successful Responses.
- **3xx** - Redirection Responses.
- **4xx** - Request Failure Responses.
- **5xx** - Server Failure Responses.
- **6xx** - Global Failures Responses.

SIP Peer-to-Peer Call



SIP Call – Proxy Server



Related IETF Working Groups

- ❑ Internet Telephony WG
- ❑ Audio / Video Transport WG
- ❑ Firewall Traversal WG
- ❑ Instant Messaging WG
- ❑ Quality of Service WG
- ❑ PSTN Interconnection WG
- ❑ IP / PSTN Interaction WG

SIP Web Pages

- IETF SIP Charter
 - <http://www.ietf.org/html.charters/sip-charter.html>
- SIP Forum
 - <http://www.sipforum.org>
- SIP Pages
 - <http://www.cs.columbia.edu/sip/>
- Many SIP product manufacturers have good information as well.

VoIP Security

- VoIP security concerns:
 - The usual PC problems (spam, viruses, worms, etc.) can now attack your telephone.
 - Spam phone calls
 - Denial of service attacks
 - Confidentiality - can someone tap your phone?
 - **Communications Assistance for Law Enforcement Act (CALEA)** law requires that law enforcement be able to tap your phone - how to comply?
 - Integrity - can someone modify your message?
 - Can be resolved by digital signatures

VoIP Security

- VoIP security concerns:
 - Availability - needs to match 5-9s (99.999%) availability provided by PSTN
 - Denial of Service attacks on VoIP may be easy
 - Authenticity - how can you be sure who you are talking to?
 - E911 services
 - VoIP services cannot always locate user and/or appropriate E911 service center for 911 calls.
 - Currently, VoIP providers are required by law to tell subscribers **NOT** to use VoIP phones for 911 calls.

VoIP and Network Security

- VoIP is hard to configure with NAT and firewalls
 - VoIP software chooses random port numbers to set up new calls.
 - Which ports should firewall open up?
 - IP addresses and port numbers are buried deep within VoIP packets.
 - NAT or Firewall may not find them all.
 - Even if they can - what if the user is encrypting the packet?

VoIP Server Compromise

- If a VoIP Call Manager (Gatekeeper) or other VoIP server is compromised by a hacker, then security is lost.
- If a VoIP Gateway is compromised, then security is lost.
- You should authenticate:
 - The person calling you
 - The gatekeeper registering your VoIP identity
 - Other VoIP servers along the way
- This is a lot of work!

Summary

- VoIP is an emerging technology to use IP as the carrier for voice communication (we've been saying that for years! literally!)
 - May use the Internet
- Many issues still being ironed out, but growing in popularity (been saying that too)
- Many security issues missing, for example, ever wonder about potential for VoIP-based spam? Hmm...