VoIP, and Multimedia over Broadband Networks

Speaker: Nader F. Mir

- **Professor of Electrical Engineering,** San Jose State University, California
- **Director,** Wireless/Optical Sensor Networks, MSE Program, Lockheed Martin Space, Sunnyvale, California
- **Editor,** IEEE Communication Magazine

Reference: Text Book

- **Computer and Communication Networks**, Authored By: Nader F. Mir, Published By: Prentice Hall Inc. 1st Ed. 2006, 2nd Ed. 2010 ISBN: 0-13-138910-6

Note: **This power point originally had 126 slides out of which 70 slides containing drawing figures have been removed from this power point since they belonged to the following copy-righted text book:**

Computer and Communication Networks

Authored By: Nader F. Mir

Published By: Prentice Hall Inc. 1st Ed. 2006, 2nd Ed. 2010

ISBN: 0-13-138910-6

If you are interested in learning more about the removed slides, please purchase the text from Prentice Hall or various on-line book stores such as Amazon or Barnes&Nobles.

Talk Agenda

- **Part 1:** Control Signaling in VoIP, and Multimedia Networks

<u>20 minutes break (we may skip)</u>

- **Part 2:** Preparation and Compression of Media for Transport

- **Part 3:** Packetization and Real Time Transmission

Part 1:
Control Signaling in VoIP and Multimedia Networks

Text:

Computer and Communication Networks
Nader F. Mir, 2nd Edition, Prentice Hall

Multimedia Networks

<u>Multimedia Networks</u>:

- A networking technology that handles Voice, Video and Data over a packet switched network

<u>Type of Networks Engaged</u>:

- Packet Switched networks (PSN)
- Public Switched telephone networks (PSTN)

Differences Between PSN an PSTN in Handling Multimedia

<u>In PSTN</u>:

- Raw Voice → Digital Form → Channel of a Frame

<u>In PSN</u>:

- Raw Voice → Digital Form → Packets

VoIP and Multimedia Protocols

<u>Control Signaling Protocols (Part 1)</u>:

- Traditional PSTN
- H.323
- Session Initiation Protocol (SIP)
- Media Gateway Control Protocol (MGCP)

<u>Media Exchange Protocols (Parts 2&3)</u>:

- Real Time Protocol (RTP) – Layer 5
- Stream Control Transmission Control (SCTP) – Layer 4

Public Switched Telephone Network (PSTN)

Main Network:

- Originally designed for telephone systems
- Backbone entity for all communications
- Nodes called “Central Office” (CO) and are hierarchical

Private Branch Exchange (PBX):

- Local entity for small size communication systems

PSTN Nodes

- **Nodes** are hierarchical called "Central Office" (CO):
  - CO5 (Local Center)
  - CO4 (Toll Center)
  - CO3 (Primary Center)
  - CO2 (Sectional Center)
  - CO1 (Regional Center)

- **Example 1**: A local connection: CO5-CO5
- **Example 2**: A domestic connection: CO5-CO4-CO4-CO5

Private Branch Exchange (PBX)

<u>Nodes</u>:

- An MxM Space Division Multiplexing (SDM)
- Followed by an MxN (M>N) Time Division Multiplexing (TDM)
- Typical values: M=20,000 (N =2-200)
- Smaller version of PBX is called <u>Key System</u> to handle up to 40 units

<u>Links</u>

- Line: carries one phone data
- Trunk: carries more than one phone data

Standard Voice Multiplexing

<u>TDM Channel</u>:

- Channel Bit Rate

  = 4 KHz

  = 4000 Cycles/S

  = 4000 x 2 = 8000 samples/s

  = 8000 samples/s x 8 b/samples

  = 64 kb/s

<u>TDM Frames</u>

- 23 channels + control channel + guard bit
- 24 x 8 b/channels + 1= 193 b/frame

Organization of PSTN

- **Architecture of PSTN:**
  "Control Signaling Network" and "Voice Network" are separated.

- **Reasons for Separation**:
  1. Control signaling messages are normally more complicated to process
  2. Control messages are much smaller than data messages and they should not be mingled with data

The Two Distinct Networks in PSTN

1) Control Signaling Network: "Signaling System 7 (SS7) Network":

- For Control Signaling
- Nodes to transfer message signal units (MSUs):
  1. Signaling Transfer Points (STPs): To transfer MSUs
  2. Service Control Points (SCPs): To transfer MSUs for advanced services such as 800 toll free numbers.

2) Voice Network: "Circuit-Switched Network":

- For voice and media exchange
- Nodes are traditional switches

Message Signal Unit (MSU) Format

- Header: 11 bytes, including
  - Service indicator
  - Destination Number
  - Origination Number
  - Circuit Identifier

- Content: A value to indicate type of control.

Signaling Protocol Stack

1. Level 1: Message Transfer Part 1 (MTP1), Physical Layer
2. Level 2: Message Transfer Part 2 (MTP2), Data Link Layer
3. Level 3: Message Transfer Part 3 (MTP3), network layer
4. Level 4: Signaling Connection Control Part (SCCP), transport layer
5. Level 4-5: ISDN User Part (ISUP), transport and application layer

ISUP Message Types

- Initial address message (IAM)
- Address complete message (ACM)
- Call progress message (CPM)
- Answer message (ANM)
- Release message (RLM)
- Release complete message (RCM)

Routing with ISUP Messages

- Message Length: Max 279 B
- Main Fields:
  - **Circuit Identification Code** (CIC) 14b, to identify a circuit at an STP, so up to 16,384 circuits
  - **Originating/Destination Point Code** (OPC/DPC) 3B. Each node such as SSP, STP, or SCP has a unique PC, so there will be up to 2.8 x 10^11 nodes can be identified.

VoIP and Multimedia Control Signaling Protocols over Packet-Switched Networks

Commonly Used Protocols:

- H.323
- Session Initiation Protocol (SIP)
- Media Gateway Control Protocol (MGCP)

H.323 Protocol

- One of the earliest protocols in 1980s
- One of the most powerful VoIP protocols
- Used substantially by ISPs
- Requires too many commands for each action

H.323 Protocol Main Components

1. **Terminals:** Users' endpoints

2. **Gateway:** A voice-(multimedia)-enabled router to interface with PSTN

3. **Gatekeeper:** To provide translations and to authorize access to the domain it belongs to, and also to communicate with all end-points to manage QoS.

4. **Multipoint Control Unit (MCU).** This unit provides multipoint services such as conference calls.

H.323 Control Signaling Protocol Organization

A. **Registration, Admission, and Status (RAS).** (Plug-in) (over UDP) signaling between endpoints and gatekeepers for registration with gatekeeper, so that a gatekeeper allows an endpoint to access its network resources.

B. **Call signaling.** (Dialing) (over UDP or TCP whichever is faster) to establish and terminate connections between endpoints.

C. **Control signaling** (Tone) (over TCP) portion of the H.323 is mandated by H.245 protocol. This protocol applies media type or bit rate limits if there is any restriction at a receiving media. H.245 operates by the establishment of one or more logical channels between endpoints.

D. **Data Signaling** (over TCP) T.120 is the protocol that provides services for data transmission using TCP connections.

H.323 Protocol Hierarchy

A) RAS (plug in)

B) Call signaling (dialing)

C) Control signaling (tone)

Media Exchange

C) Control signaling

D) Data signaling (hang-up)

A) RAS Signaling

A1) Gatekeeper Discovery. When plug in, a process once and for all times starts to <u>find (discover) a Gatekeeper</u>.

A2) Endpoint Registration.

- ✓ This process is done once and for all times to register with a Gatekeeper.
- ✓ Any registration has a limited life time up to 136 years equivalent to hexadecimal FFFFFFFF seconds but it can be a lower number requested by the endpoint or assigned by the gatekeeper

A3) Admission. At the beginning of call a user can request <u>permission from the gatekeeper</u> to participate in a call

A) RAS Signaling …

A4) Bandwidth Modification. An endpoint can modify its bandwidth. If the bandwidth modification does not exceed the limit specified by the gatekeeper, it is granted.

A5) Status. The status mode allows a gatekeeper to be informed whether an endpoint is still functional or active.

A6) Disengage. At the end of call Once a call or transmitting media ends, each endpoint should disengage

A) RAS Signaling …

A7) Resource Availability. A process that informs the level of call capacity. A gateway can send a message to inform the gatekeeper of the currently available call capacity and bandwidth for each protocol supported by the gateway.

A8) Service Control. To enable some "advanced features". Any entity can initiate this process. It can also be used to implement the specific capability of a vendor equipment.

A9) Request in Progress. An entity can express that the response to a given request may take longer than expected.

RAS (A1 - Gatekeeper Discovery)

- **gatekeeper request (GRQ)** message . Normally, this message is multicast to all gatekeepers with the multicast address 224.0.1.41 at port 1718. Any gatekeeper that is available and willing to control the endpoint to be the gatekeeper for a given endpoint may respond.

- **gatekeeper confirmation (GCF)** message. When a gatekeeper is available, it sends this message.

- **gatekeeper reject (GRJ)** message. In case a gatekeeper is not available due to variety of reasons as lack of resources.

RAS (A2 - Registration)

- **registration request (RRQ)** message. An endpoint joins a gatekeeper starting with the issuance of this message. The port number to be used for the message is still the RAS signaling port of 1719.
- **registration confirmation (RCF)** message. the gatekeeper responds with an RCF message and an alias will be assigned to the endpoint.
- **registration reject (RRJ)** message. Otherwise, the gatekeeper can choose to reject a registration by responding with an RRJ message.
- **unregistration request (URQ)** message. The registration can be cancelled by the endpoint through sending an URQ.
- **unregistration confirmation (UCF)** message. To be confirmed by the gatekeeper through sending an UCF message.

RAS (A3 - Admission)

- **admission request (ARQ)** message. An endpoint can then request permission from the gatekeeper to participate in a call. The ARQ message indicates parameters such as:
  - ✓ type of call being two-party or multi-party,
  - ✓ endpoint identifier, a unique string as call identifier,
  - ✓ a call reference value, and other party aliases and signaling addresses.

- **admission reject (ARJ)** message. A gatekeeper may choose to deny (lack of available bandwidth) or an endpoint not being registered.

- **admission confirm (ACF)** message. The admission confirm message also specifies the amount of bandwidth required in units of 100 b/s.

RAS (A6 - Disengage)

- **disengage request (DRQ)** message. Once media exchange ends, each endpoint should send a DRQ to its associated gatekeeper. This message must contain disengage reason.

- **disengage confirm (DCF)** message The associated gatekeeper responds either with a DCF message, or

- **disengage reject (DRJ)** message. A gatekeeper could also decide to terminate a call in which case the gatekeeper sends a DRQ message to the endpoint. The endpoint must stop transmitting media and must bring the session to a close using H.245 control messages and call-signaling messages.

H.323 Protocol Hierarchy

A) RAS (plug in)

B) Call signaling (dialing)

C) Control signaling (tone)

Media Exchange

C) Control signaling (tone)

D) Data signaling (hang-up)

B) Call Signaling

- **Setup.** Upon admission ,the Setup message contains several parameters such as a call reference and the user-to user information element.
- **Call Proceeding.** Optional and sent by the recipient of a Setup message to indicate that the Setup message has been received.
- **Alerting.** Sent by an entity to indicate that the called point is being alerted.
- **Progress.** From a called gateway to indicate that a call is in progress.
- **Connect.** From a calling endpoint to indicate that the called party <u>has accepted the call</u>.
- **Release Complete.** From an entity that releases (terminates) the call, it should also provide the other terminal with a reason for the release.
- **Facility.** Used to <u>redirect a call</u>. For example, consider a case when a Setup message is sent to an endpoint where the endpoints gatekeeper indicates that it wants to intervene. Thus, the user who receives this message releases the call and attempts to set up the call again via the called gatekeeper.

H.323 Protocol Hierarchy

A) RAS (plug in)

B) Call signaling (dialing)

C) Control signaling (tone)

Media Exchange

C) Control signaling (tone)

D) Data signaling (hang-up)

C) Control Signaling

C1) Capabilities Exchange. To ensure the mutual understanding of communication capabilities between the two endpoints

C2) Master-Slave Assignment. A method through which the system decides who the controlling entity should be.

C3) Media Exchange Establishment Control. Immediately before and after a media exchange. This signaling requires the construction of "logical channels" between endpoints before media can be exchanged. A logical channel is a combination of IP address and port number.

Control Signaling (C1 – Capability Exchange)

- **terminal capability set request (TCS-Req)** message . A calling endpoint releases its capabilities in a terminal capability set request TCS-Req indicating a sequence number, the types of audio/video formats that the endpoint can send and receive, and what media can be handled simultaneously.

- **capability set ACK (TCS-Ack)** message. A confirmation response containing a sequence number that matches the sequence number received in the original request.

- **terminal capability set reject (TCS-Rej)** message. If there is no capabilities match, the called endpoint responds with a TCS-Rej.

Control Signaling (C2 – Master Slave Assignment)

- **master-slave assignment (MSA)** message. The assignment of master in a communication session starts with each endpoint sending the MSA . This message contains the terminal type value and its random number. Once an MSA message is received at an endpoint.

- **master-slave assignment ACK (MSA-Ack)** message. the endpoint compares the terminal type value with its own value and makes the decision as who the master is. It then returns a MSA-Ack. This message delivers the ranking view of the endpoint as whether it is a master or slave.

Control Signaling (C3 – Media Exchange Establishment Control)

- **open logical channel request (OLC-Req)** message. Each endpoint sends an OLC-Req to others containing the information such as the <u>type of data</u> to be sent, an RTP session ID, and an RTP payload type.

- **open logical channel ACK (OLC-ACK)** message. If the receiving endpoint is prepared for the media, it responds with an OLC-ACK containing information such as the same logical channel number as received and a transport address to which the media stream should be sent.

- **open logical channel reject (OLC-Rej)** message. An endpoint has an option to deny a request for a reasons such as an incapability to handle the suggested media format.

Control Signaling (C3 – Media Exchange Establishment Control …)

- **open logical channel confirm (OLC-Con)** message. Upon receipt of OLC-ACK, the calling endpoint responds with OLC-Con to indicate that the bidirectional setup is fine.
- **close logical channel request (CLC-Req)** message. Sent from either endpoint to close a logical channel.
- **close logical channel ACK (CLC-ACK)** message. An acknowledgement to CLC-Req.
- **end session (END-Ses)** message. Upon closure of all logical channels in a session, the session must be terminated
- by both endpoints each sending an END-Ses.

Main VoIP and Multimedia Control Signaling Protocols over Packet-Switched Networks

- **H.323**
- **Session Initiation Protocol (SIP)**
- **Media Gateway Control Protocol (MGCP)**

Session Initiation Protocol (SIP)

- Newer than H.323
- One of the most powerful multimedia protocols
- Provides many pieces of information in its messages (more intelligent)
- A SIP address is of type HTTP, so it can be included in the WEB content
- SIP is client-server protocol:
  - Client: user agent program
  - Server: program to respond
  - Request: call
  - Response

SIP - Main Servers

1. **Proxy Server.** Forwards requests from a user agent to a different location and handles authorizations by checking if the caller is authorized to make a particular call, similar to the ones of proxy servers used for web access.
2. **Registrar Server**. For registering users with their available addresses. The server also frequently updates the user information database that the location server consults. The requests for registration must be authenticated before the registration.
3. **Redirect Server.** For call forwarding when a user goes to a different location.
4. **Location Server.** For address resolution for proxy and redirect servers. It interacts with the databases of proxy and redirect servers during call set-up.

SIP Message Structure

- **UDP Segment:** IP Header + UDP Header + SIP Message
- **SIP Message:** Start-Line (version and code) + Header + Body

Example:

- Start Line:
  - SIP/2.0   200 OK
- Header
  - Via: SIP/2.0/UDP
  - From: SIP: user1@isp1.com
  - To: SIP: user2@isp1.com
  - Call-ID: 1234@station.isp1.com
  - CSeq: 1 INVITE
  - Subject: Vacation
- Body:
  - Audio 4444 RTP

SIP Messages

- **Provision (100s):** TRYING, RINGING, CALL-IS-BEING-FORWARDED, …
- **Success (200s):** OK, …
- **Redirection (300s):** MOVED-TEMPORARILY, MOVED PERMAMENTLY, …
- **Request Features (400s):** BAD-REQUEST, UNUTHORIZED, PAYMENT-REQUIRED, BUSY HERE…
- **Server Failure (500s):** SERVER INTERNAL ERROR, BAD GATEWAY…
- **Global Failure (600s):** BUSY ANYWHERE, DOES NOT EXIST ANYWHERE …

Registration and Call Establishment

Registration: Before any call, a user-agent must register with a Registrar Server:

- The user sends a multicast REGISTER message to Registrar Servers, the multicast address is 224.0.1.75
- Any Registrar Server responds.
- Once registered:
  - a registration by default is active for 136 years (FFFFFFFF) minutes
  - multiple registrations from different locations is allowed so that a coming call can be sent to multiple registered terminals

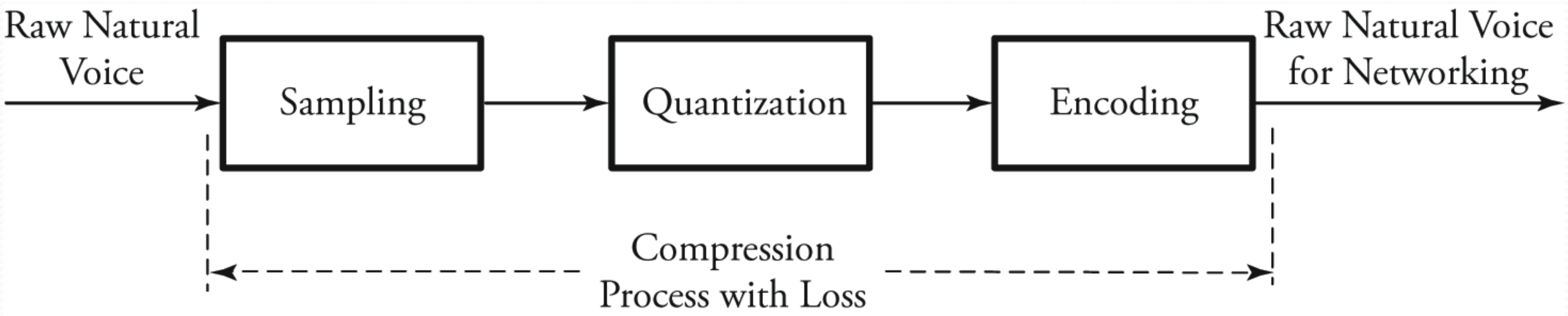
Call Establishment (and Termination): Starts with INVITE, and ends with termination using BYE :

Part 2
Preparation and Compression of Media for Transport

Text:

Computer and Communication Networks
Nader F. Mir, 2nd Edition, Prentice Hall

I) Digital Voice Process



Sampling Methods:

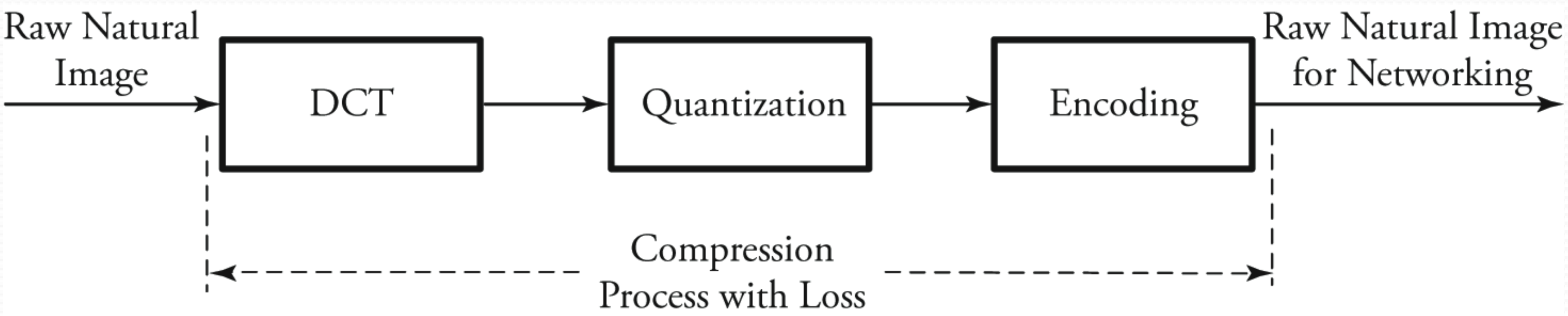
- **Pulse amplitude modulation (PAM)**: sampled values are translated to pulses with corresponding amplitudes.
- **Pulse width modulation (PWM):** sampled valued are translated to pulses with corresponding widths.
- **Pulse position modulation (PPM):** sampled values are translated to identical pulses but with corresponding position to sampling points.

Quantization:

- Samples are real numbers, so up to infinite bits are required for transmission of raw sample. In practice, sampled values are rounded off to available quantized levels. However, by rounding off the values we lose data and generate distortion.

Encoding: quantized values are converted to bits.

II) Still Image and Joint Photigraphic Experts Group (JPEG) Sources



Discrete cosine transform (DCT):

converts intensity of light to values by dividing a raw natural image into a series of standard N x N pixel blocks.

Quantization:

Samples are real numbers, so up to infinite bits are required for transmission of raw sample. In practice, sampled values are rounded off to available quantized levels. However, by rounding off the values we lose data and generate distortion.

Encoding:

quantized values are converted to bits and are compressed.

Part 3
Packetization and Real Time Transmission

Text:

Computer and Communication Networks
Nader F. Mir, 2nd Edition, Prentice Hall

Broadband Network Applications

- Voice over Internet Protocol (VoIP)
- Data steaming over IP networks such as cases for financial corporations
- Video steaming over IP networks
- Live multicasting television programs over IP networks (IPTV)
- Video-on-demand (VoD)
- Premium and on-line gaming

Media Exchange

Examples of Strategies:

- Content Distribution Network (CDN)

Layer 4 Protocols:

- UDP/TCP
- SCTP

Layer 5 Protocols:

- RTP
- RTCP

Content Distribution Networks (CDNs)

<u>Applications</u>:

- Video Streaming Strategies

<u>Protocol</u>:

- CDN is a network of Proxy Servers located at certain locations to provide video streaming:
  - A streaming owner company uses its CDN servers
  - Downloads videos on various dispersed servers
  - A user is directed to the most convenient location

Real-Time Protocol (RTP)

Applications:
- Layer 5 protocol (Application Layer)
- Runs over UDP for each session

Session:
- A logical connection between a client and a server
- A session is defined by:
  - Destination port no. from UDP header
  - IP Address

Segment (Packet):
- Standard Size: 1,450 B up to the largest packet size
- Data (media) + RTP header (12B + up to 64B optional) + UDP header (8B) + 20B (IP header)

RTP Packet Format

- **Version (V).** 2b.
- **Padding (P).** 1b, indicates the existence of padding field at the end of the payload. Padding is required in applications which require the payload to be a multiple of some length.
- **Extension (X):** 1b to indicate the use of an extension header for RTP.
- **Contributing source count (CSC).** 4b, indicates the number of contributing source identifiers.
- **Marker (M).** 1b, to indicate boundaries in a stream of data traffic. For video applications, it can be used to indicate the end of frame. the contributing sources for the data.

RTP Packet Format …

- **Payload type.** 7b, to specify the type of RTP payload. It also contains information on the use of compression or encryption.
- **Sequence number.** 16b, used by a sender to identify a particular packet within a sequence of packets. It is used to detect packet loss and for packet re-ordering.
- **Timestamp.** 32b, to enable the receiver to recover the timing information. It indicates the timestamp when the first byte of data in the payload was generated.
- **Synchronization source identifier (SSI).** A randomly generated field to identify the RTP source in an RTP session.
- **Contributing source identifier (CSI).** 32b, to indicate the ID contributing sources, up to 16 sources simultaneously.

Stream Control transmission Protocol (SCTP)

Principles:

- An error free Layer-4 protocol for streaming
- More features than TCP and UDP
- Several Streams within a connection
- SCTP can be used as Transport protocol without UDP or with encapsulated UDP

SCTP Chunk Type

Sample Values for Field "Type":

- 0 = Data (chunk carries data)
- 1 = INIT (initiation of a session)
- 7 = SHUTDOWN (to terminate session)

Thanks for Listening!

VoIP, and Multimedia over Broadband Networks

Speaker: Nader F. Mir

- **Professor of Electrical Engineering,** San Jose State University, California
- **Director,** Wireless/Optical Sensor Networks, MSE Program, Lockheed Martin Space, Sunnyvale, California
- **Editor,** IEEE Communication Magazine

Reference: Text Book

- **Computer and Communication Networks**, Authored By: Nader F. Mir, Published By: Prentice Hall Inc. 1st Ed. 2006, 2nd Ed. 2010 ISBN: 0-13-138910-6