IP Telephony Protocols and Architectures

Melinda Shore

Nokia IP Telephony

shore@ithaca-viennasys.com

Agenda

- Overview
- Scenarios
- Basic components of an IP telephony system
- Standards and standards bodies
- H.323 101
- Decomposing gateways (more components! more protocols!)
- Security (H.235)
- Numbering, addressing
- Wrap-up

- Various breaks for questions

Caveats

- Not talking much about
  - Mobility
  - Wirelessness
  - Multipoint/multiparty architecture
- SIP deserves a lot more attention than it's going to get today
- So does the PSTN switching hierarchy

Overview

Becoming mainstream

The big driver

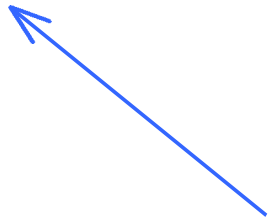
Plus, it’s pretty cool

IP Telephony - What is it?

- Several things, actually
- Widely used end-to-end, very often with video
  - NetMeeting
  - iVisit
  - CU-Seeme
- Increasingly popular to provide a gateway to traditional switched circuit networks
- Low-cost long distance services by trunking calls over an IP network
- Replace a PBX or key system with telephony on a LAN within an enterprise
- “IP Centrex”
- Call centers (CTI)
  - Screen pops
  - Predictive dialers
  - These usually use APIs and toolkits (TAPI, JTAPI, IBM CallPath)
- The protocols and architectures we’re talking about today cover all of these

Also known as ...

- Voice over IP (VoIP)
- Internet Telephony
- IP Telephony
- Computer Telephony Integration (CTI)



Not really - CTI can *use* IP, but is actually something else

Services

- IP telephony enables a variety of services
- Traditional telephony
- Video telephony
- Integration of voice and email
- Information kiosks (airports, hotels, supermarkets, etc.)
- Web browsing and other data stuff on your telephone (esp. wireless)
  - Palm VII is a step in that direction
  - Qualcomm has a new telephone that runs Palm OS
- WAP: Wireless Application Protocol
- Next-generation wireless **will** run over IP
- New stuff all the time

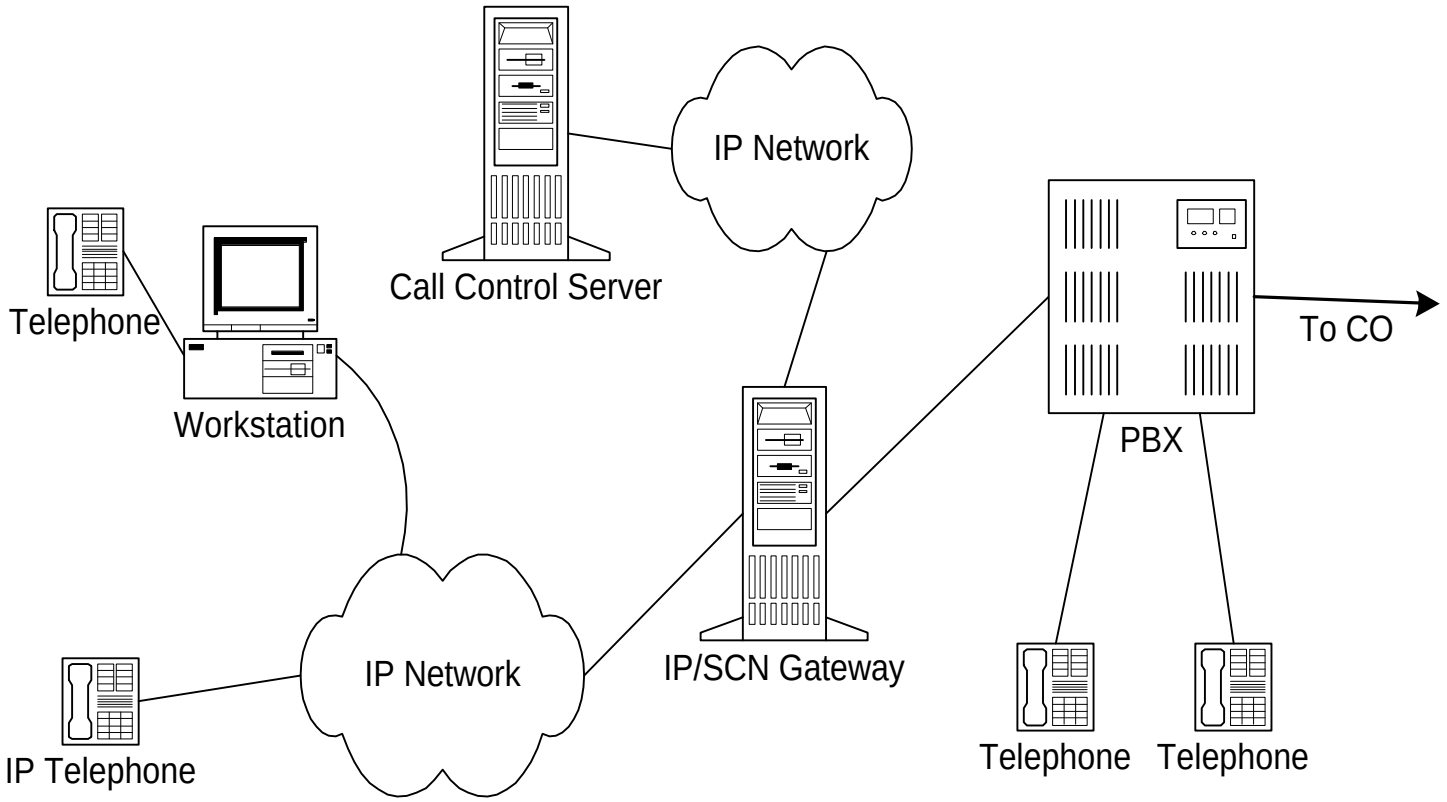
These are not yet IP-based, but are representative of the sorts of services and applications which will be IP-based in the future

A little terminology (more later)

- Traditional telephony, aka
  - POTS: plain old telephone system
  - PSTN: public switched telephone network
  - GSTN: general switched telephone network
  - CSN: circuit-switched network
  - SCN: switched circuit network (this is what we'll use, mostly)
- Black phone: a traditional dumb analog telephone device
- IWF: interworking function

Components

Typical enterprise configuration

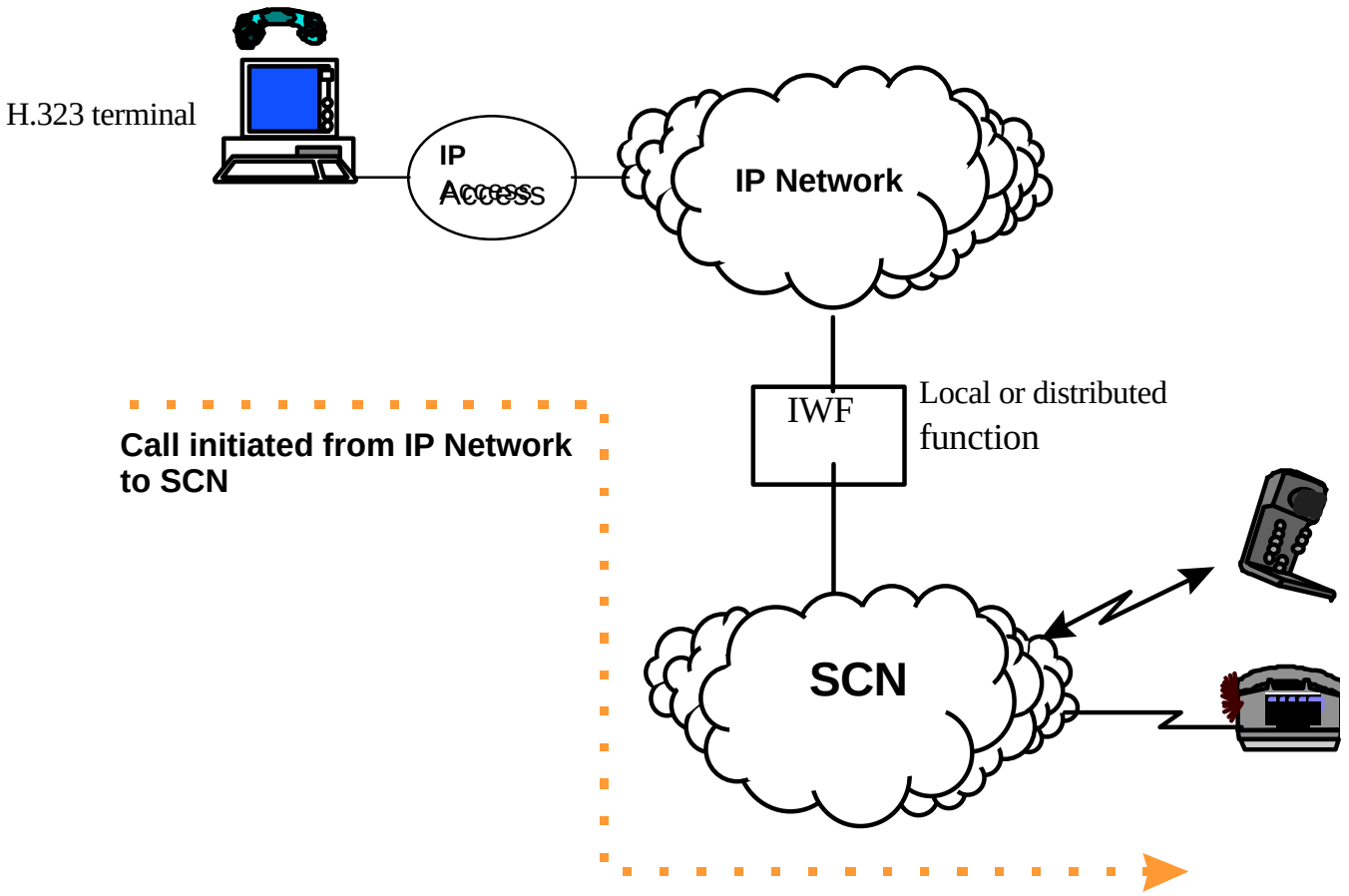


Scenarios

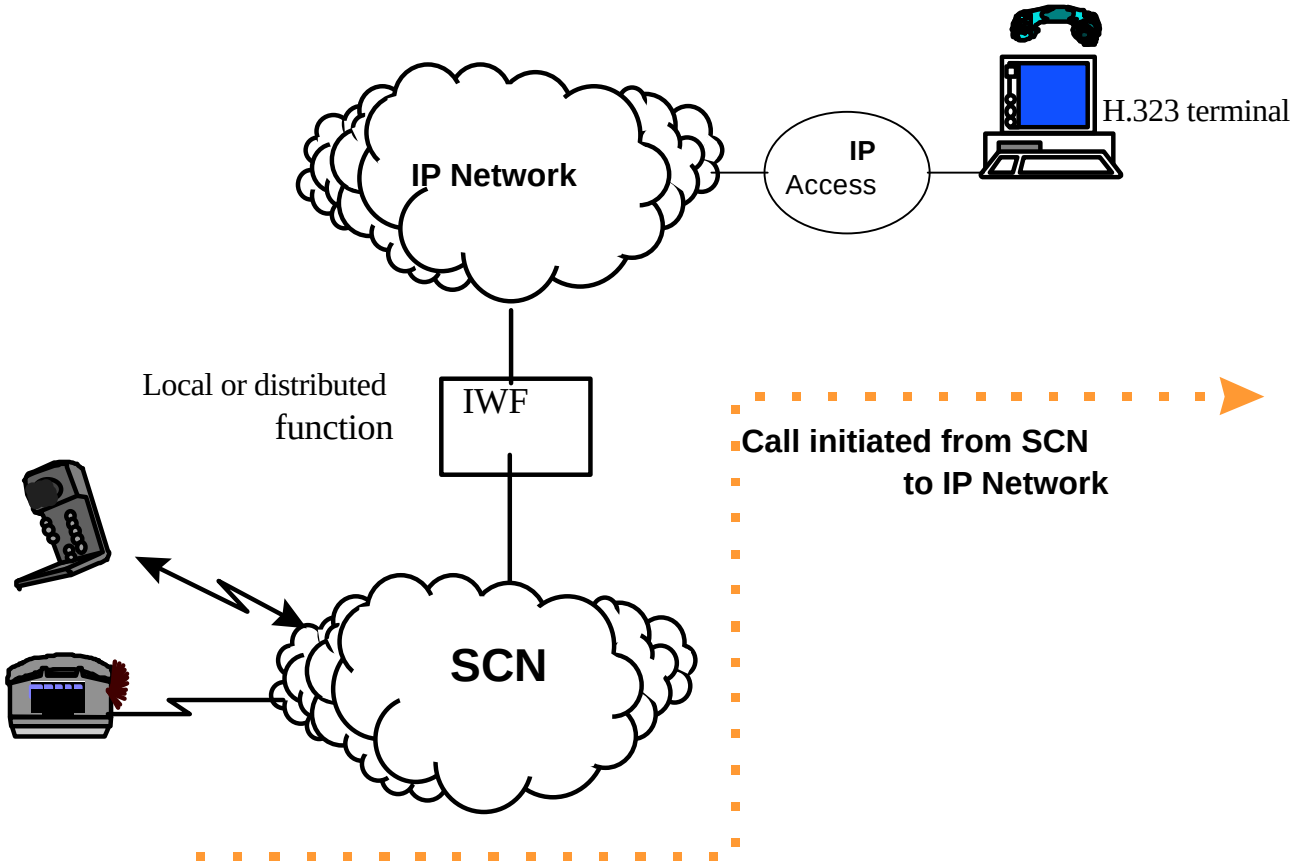
Scenarios

- End-to-end IP
- Calls originate in IP network and terminate in SCN
- Calls originate in SCN and terminate in IP network
- Calls originate in SCN, pass through an IP network and terminate in SCN
- Calls originate in IP network, pass through SCN, and terminate in IP network

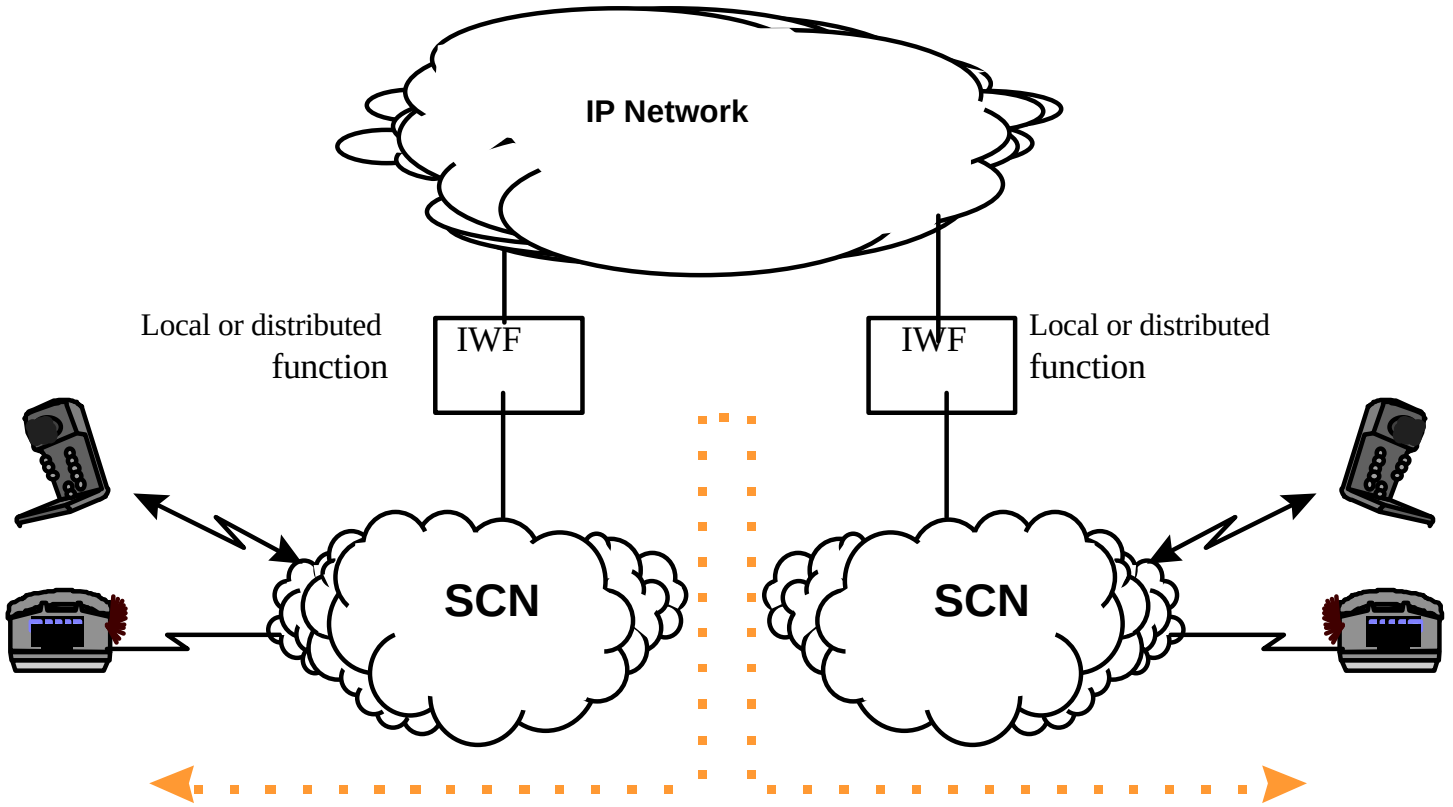
Calls originate in IP network



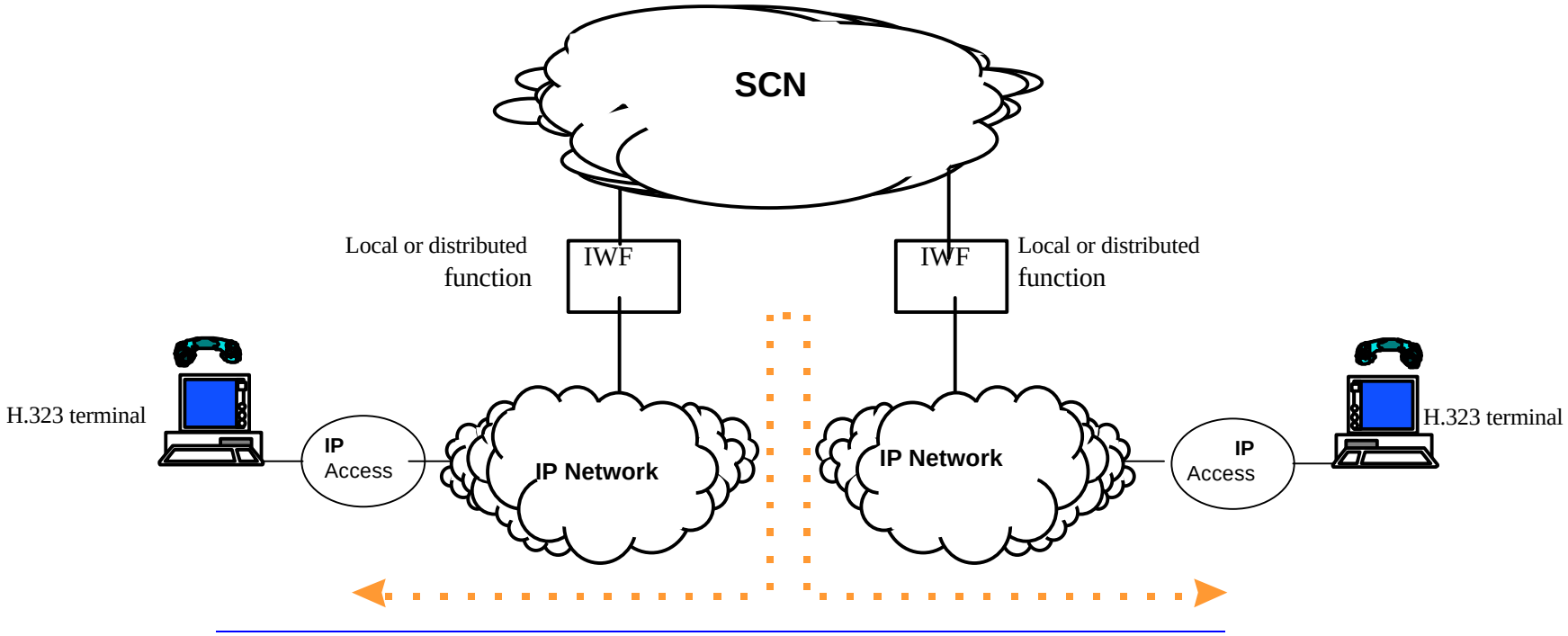
Calls originate in SCN



Calls originate and terminate in SCN, pass through IP network



Calls originate and terminate in IP network, pass through SCN



Standards

Different approaches

- IP telephony is heavily standards-driven (interoperability!)
- People working on standards for IP telephony come from two different communities
  - Traditional voice networks (bellheads)
  - IP networking (netheads)
- Centralized vs. decentralized models of call control
- Bellheads tend to see terminals as stupid and networks as smart
- Netheads tend to see networks as stupid and terminals as smart
- Reflected to a certain extent in H.323 vs. SIP
- Realities of building working telephone systems leads to some collaborations, some shared vision, occasional disagreements (“Your protocols suck.” “Your protocols suck more.”)

Standards: Who are they?

- ETSI - European Telecommunications Standards Institute
  - TIPHON - Telecommunications and IP Harmonization on Networks
  - SEC - Security
  - STQ - Speech Transmission Quality
  - NA2 - ETSI technical committee working on naming and addressing
  - NA8 - working on accounting and billing for IP
- ITU-T
  - SG 16 - multimedia applications
  - SG 2 - naming and addressing
  - SG 11 - signaling
  - SG 15 - transport equipment
- ATM Forum RMOA - Realtime Multimedia over ATM

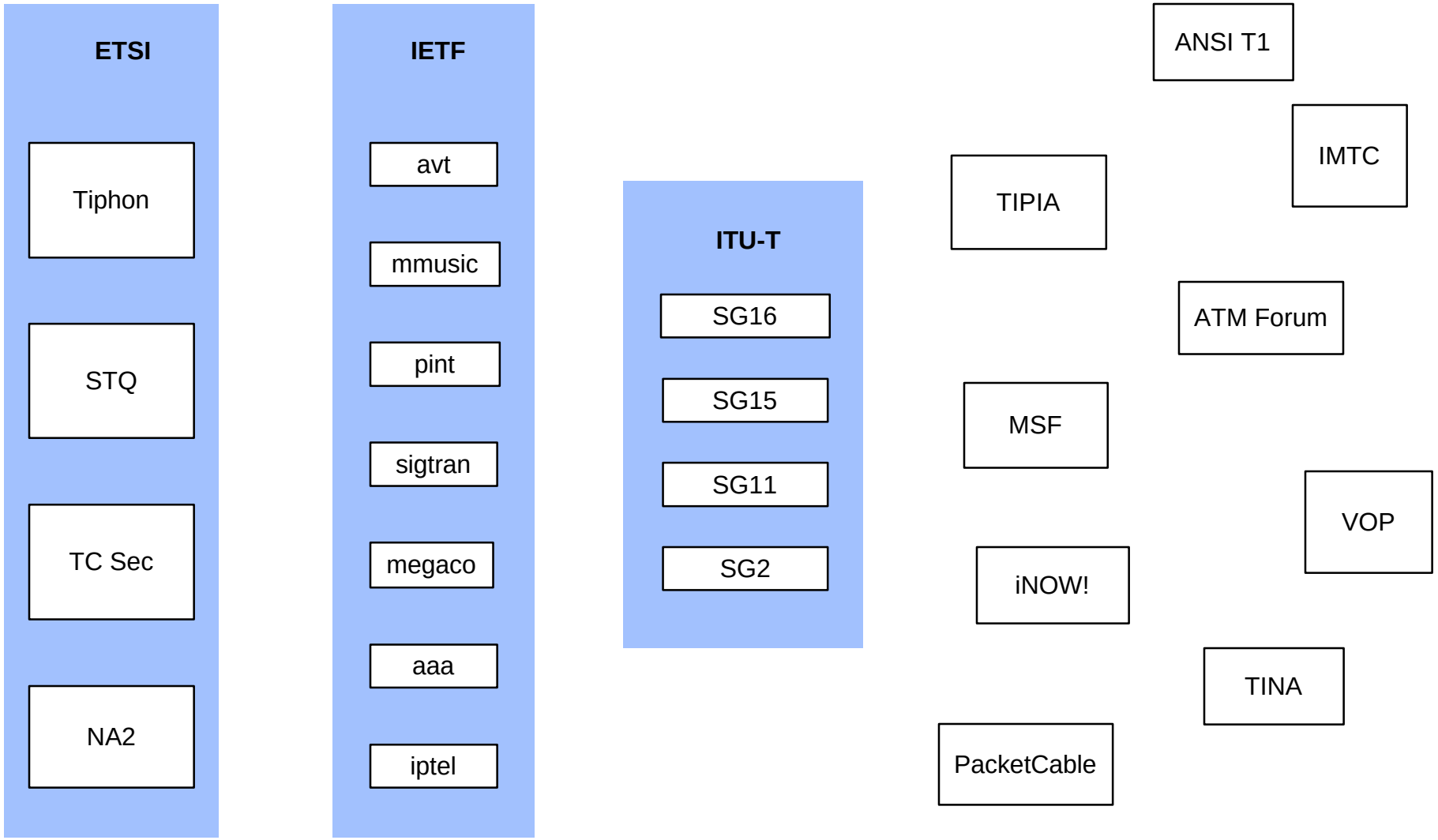
Standards - Who are they? (2)

- IETF - Internet Engineering Task Force
  - sigtran - signaling transport
  - megaco - media gateway control
  - iptel - IP telephony
  - pint - PSTN interworking (click-to-dial services)
  - aaa - authentication, authorization, and accounting
  - mmusic - multiparty multimedia control
  - avt - audio-video transport
- PacketCable - CableLabs (US) project to produce specifications for packet data over cable, including packet voice
- VOP - Voice Over Packet (Telcordia [Bellcore]-initiated)
- ANSI Committee T1
- MSF - Multiservice Switching Forum
- Softswitch Consortium

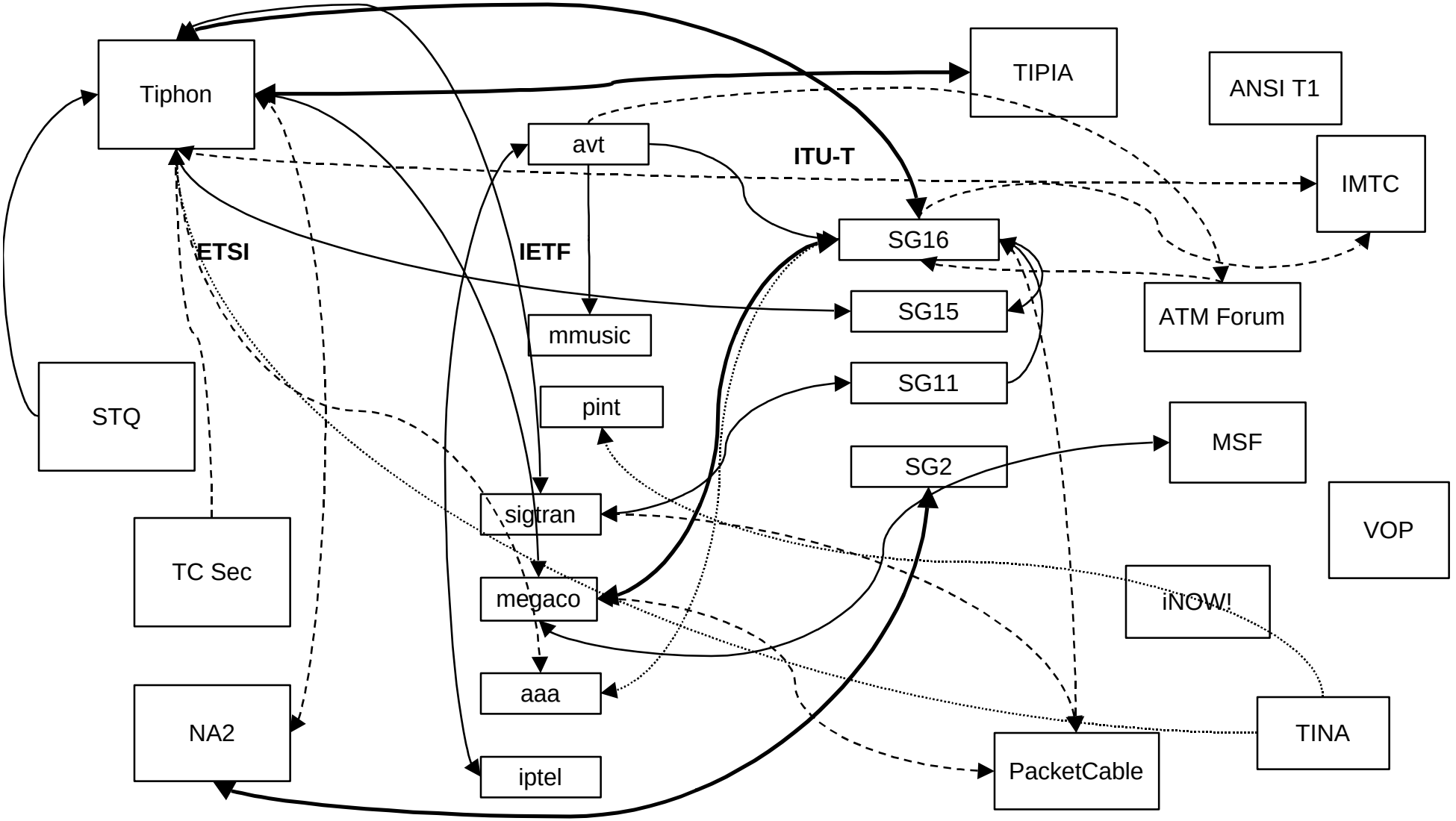
Implementation Agreements

- iNOW! - Interoperability implementation agreement
- TIPIA - TIPHON IP telephony Implementers Association
- IMTC - International Multimedia Teleconferencing Consortium
- TINA - a EURESCOM IP telephony project

IP Telephony Standards Groups



Standards Groups - the relationships



Good sources for standards documents

- http://www.etsi.org/Tiphon/Tiphon.htm - follow the “FTP area” link
- http://www.ietf.org - most of the relevant working groups are in the transport area
- http://www.itu.int - this is the ITU home site. No free access to documents, so try …
- ftp://standard.pictel.com/avc-site - has SG16 working (meeting) documents, as well as draft standards
- http://www.k1om.com/imtcftp.html - IMTC reflector
- http://standard.pictel.com/webftp.html - outstanding site with links to many groups working in this area
- http://www.inowprofile.com - home page for iNOW! interoperability agreement

H.323

What is H.323?

- H.323 is a multimedia conferencing standard produced by the ITU-T (Study Group 16 Questions 12-14)
- Umbrella specification describing how to build systems using other specifications (H.225, H.245, etc.)
- Built around traditional telephony common-channel signaling model
- Currently the most widely-supported IP telephony signaling protocol
- Very complex - stacks are available from a few vendors and tend to be expensive
- New open source H.323 project, includes an ASN.1 PER compiler: http://www.openh323.org

H.323 is an umbrella specification

- H.323: “Infrastructure of audiovisual services – Systems and terminal equipment for audiovisual services: Packet-based multimedia communications systems”
- H.245: “Control protocol for multimedia communication”
- H.225: “Call signalling protocols and media stream packetization for packet based multimedia communication systems”
- Q.931: “ISDN user-network interface layer 3 specification for basic call control”
- H.235: “Security and encryption for H-Series (H.323 and other H.245 based) multimedia terminals”
- H.450.1: “Generic functional protocol for the support of supplementary services in H.323”
- Codecs
  - G.711: “Pulse Code Modulation (PCM) of voice frequencies”
  - G.722: “7 kHz audio-coding within 64 kbit/s”
  - G.723.1: “Speech coders: Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s”

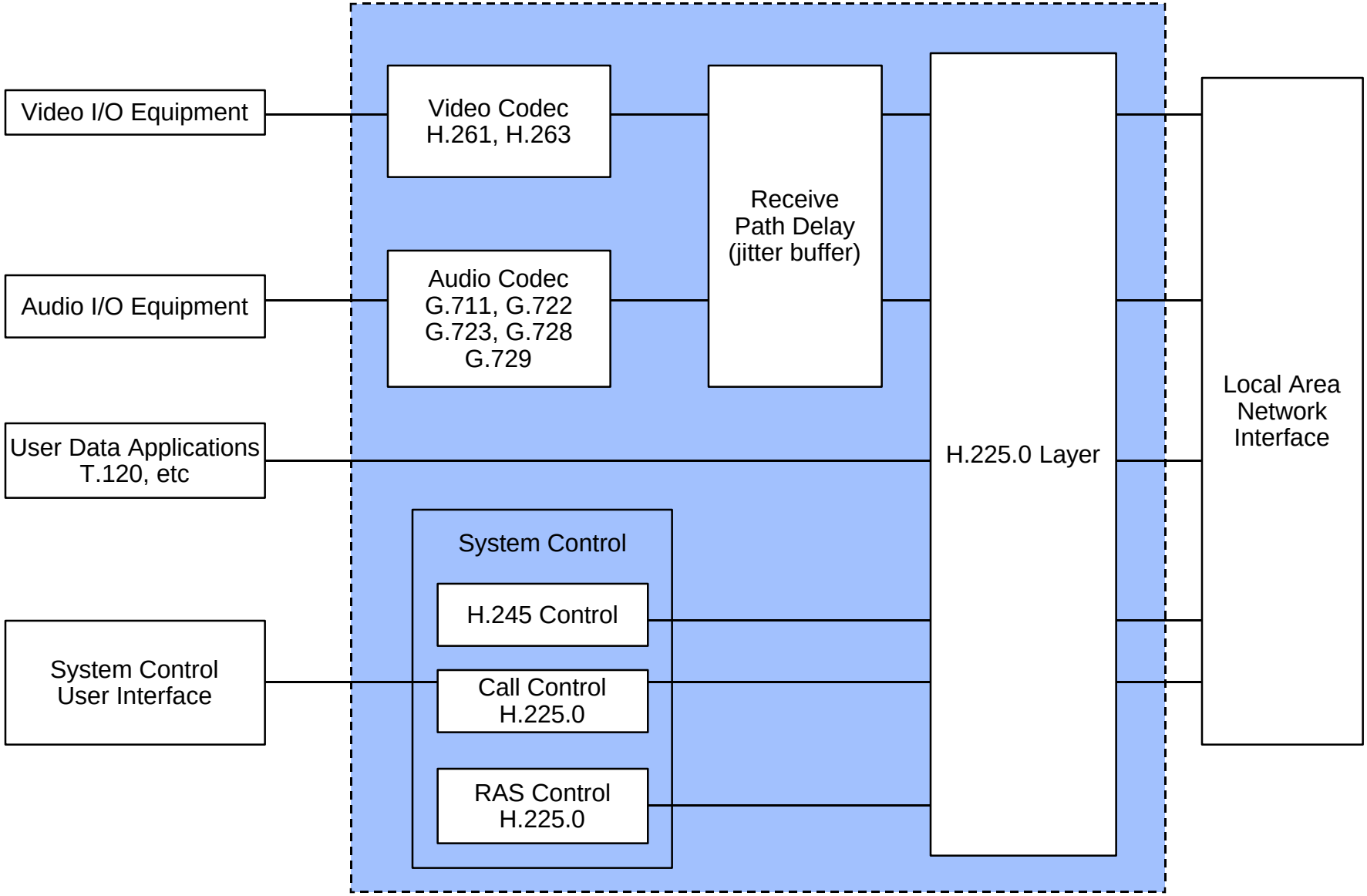
NOKIA

H.323 is an umbrella specification(2)

- More codecs:
  - G.728: “Coding of speech at 16 kbit/s using low-delay code excited linear prediction”
  - G.729: “Coding of speech at 8 kbit/s using Conjugate Structure Algebraic-Code-Excited Linear-Prediction (CS-ACELP)”
  - H.261: “Video codec for audiovisual services at p $\times$ 64 kbit/s”
  - H.263: “Video coding for low bit rate communication”
- T.120: “Data protocols for multimedia conferencing”
- X.680: “Information Technology - Abstract Syntax Notation One (ASN.1) - Specification of basic notation”
- X.691: “Information Technology - ASN.1 Encoding Rules - Specification of Packed Encoding Rules (PER)”

- At least one audio channel is required - video is optional
- Most of the codecs are encumbered - intellectual property issues abound
- Lots of work currently underway on the use of GSM codecs with H.323

Scope of H.323 (terminals)



Information streams

- Video
- Audio
- Data (T.120)
  - Whiteboarding
  - Pictures
  - Any sort of shared data
- Communications control (H.245)
  - Capabilities exchange
  - Open/close logical channels
  - Mode changes
- Call control (H.225)
  - Call establishment
  - Call tear-down

H.225

- TCP connection on a well-known port
- Used to perform call signaling
- Also specifies packetizationfor all H.323 communication
- Call signaling is based on ISDN signaling (Q.931)
- Media are packetized using RTP (including RTCP control channel)
- Work on optional UDP connection on well-known port underway

RAS signaling

- Registration, Admission, Status
- Separate UDP-based H.225 stream
- Used to:
  - register a user with a gatekeeper
  - indicate bandwidth changes
  - exchange status information
  - de-register

H.245

- Connection control function of H.323:
  - Master/slave determination
  - Capability Exchange
  - Logical Channel Signalling
  - Close Logical Channel Signalling
  - Mode Request
  - Round Trip Delay Determination
  - Maintenance Loop Signalling
  - May be used for transmitting user input, for example DTMF strings
- Encoded using ASN.1 PER

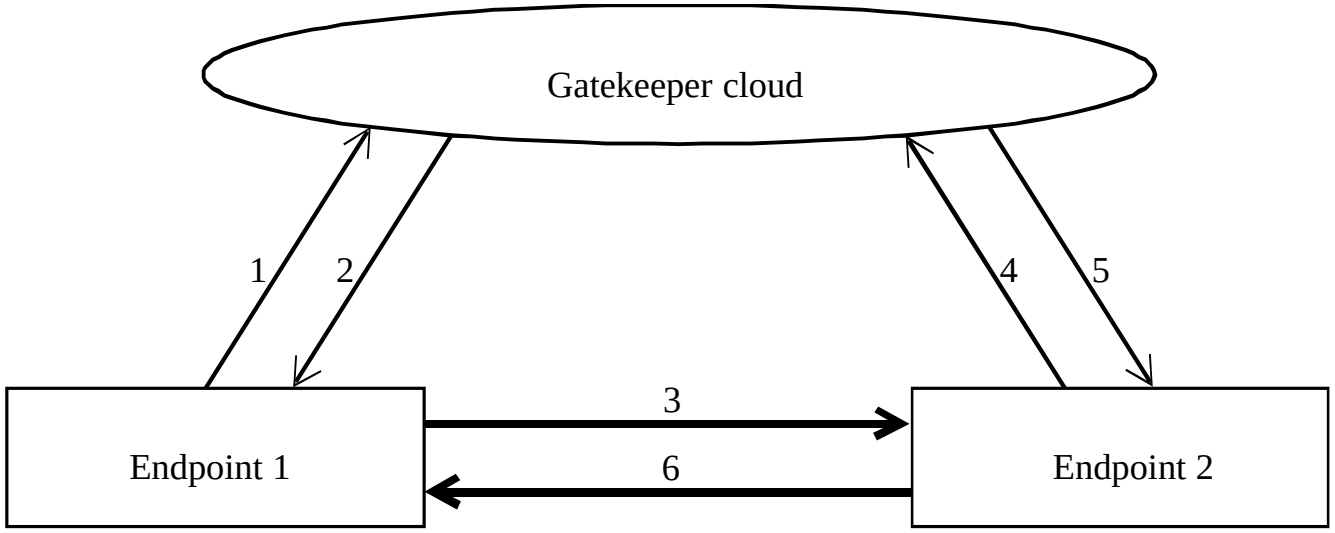
Gatekeeper

- “Brains” of IP telephony network
- One per zone
- Functions MUST include:
  - Address translation (E.164, domain name, other aliases)
  - Call admission control (based on identity, calling card account number, available resources, etc.)
  - Bandwidth control - this is allowed to be null (and in practice almost always is)
  - Zone management - must perform above functions for any endpoint registered with it
- Functions MAY include:
  - Call signaling (“gatekeeper-routed model”)
  - Call authorization
  - Bandwidth management
  - Directory services
  - Other stuff

Call signaling

- May be end-to-end (“direct call signaling”)
- May be routed through gatekeeper (“gatekeeper-routed”)
  - This is mandated by TIPHON and other organizations using H.323 as a base protocol
- Multiple phases:
  - Phase A: Call setup (RAS and H.225)
  - Phase B: Initial communication and capability exchange (H.245)
  - Phase C: Establishment of audiovisual communication
  - Phase D: Call Services
  - Phase E: Call termination
- With H.323v3, OpenLogicalChannel structures may be loaded into initial “connect” messages (AKA “Fast Connect”)
- H.245 messages may also be tunneled within Q.931 call signaling instead of being carried on a separate H.245 channel (“H.245 tunneling”)

Direct call signaling - Phase A



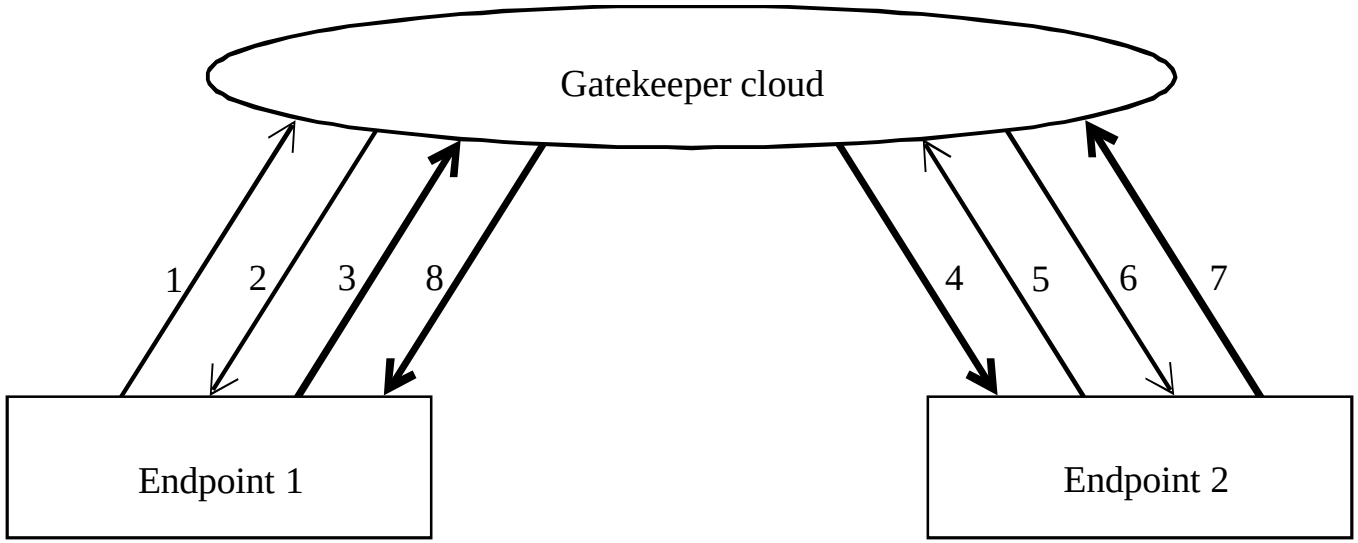
1 ARQ
2 ACF/ARJ
3 Setup
4 ARQ
5 ACF/ARJ
6 Connect

Call Signalling Channel Messages

RAS Channel Messages

Gatekeeper-routed call signaling - Phase A

1 ARQ
2 ACF/ARJ
3 Setup
4 Setup
5 ARQ
6 ACF/ARJ
7 Connect
8 Connect



Call Signalling Channel Messages

RAS Channel Messages

T1521280-96

Call signaling - Phase B and C

- Once Phase A is complete, the control signaling (H.245 channel) is setup
- First thing that happens is terminal capabilities (supported codecs, bandwidth, etc.) are exchanged
- Next order of business is master/slave determination
- Then Phase C is begun, and logical channels (i.e. media channels) are opened

Phase D - Call services

- Various signaling services are available throughout duration of call
  - Bandwidth changes
  - Status
  - Ad-hoc conference expansion
  - Supplementary services (H.450)
    - H.450.2: “Call transfer supplementary service for H.323”
    - H.450.4: “Call Hold Supplementary Service for H.323”

Phase E - Call termination

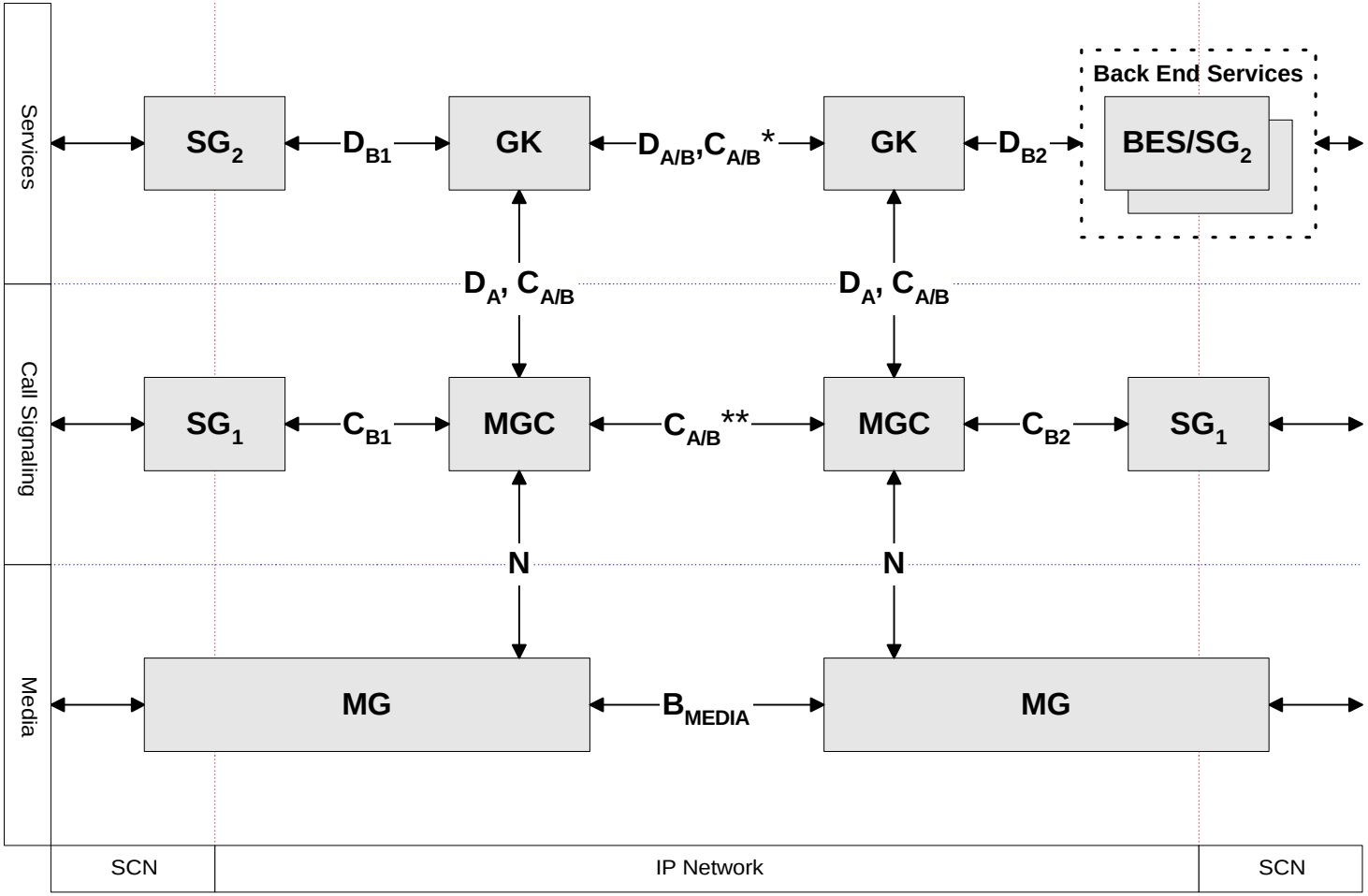
- Either endpoint may terminate a call
- Discontinue transmission of
  - video, then
  - data, then
  - audio
- Close all logical channels
- Send H.245 “end session” command, wait for replying “end session,” then tear down H.245 channel
- If H.225 channel is still open, send “Release Complete”
- If there’s a gatekeeper, additional procedures are required:
  - Send a “Disengage Request” to gatekeeper
  - Wait for “Disengage Confirm” from gatekeeper
- Gatekeeper may terminate a call by sending a DRQ to an endpoint

Gateway decomposition

What? Why?

- Gateways are being decomposed into
  - Gateways (usually referred to as “media gateways” and “signaling gateways”)
  - Gateway controllers
- Media gateway controllers manage multiple media and/or signaling gateways
- H.323 is a large, heavy protocol - it doesn’t scale well
- H.323 is a call control environment, and doesn’t do connection or resource control particularly well

The TIPHON architecture



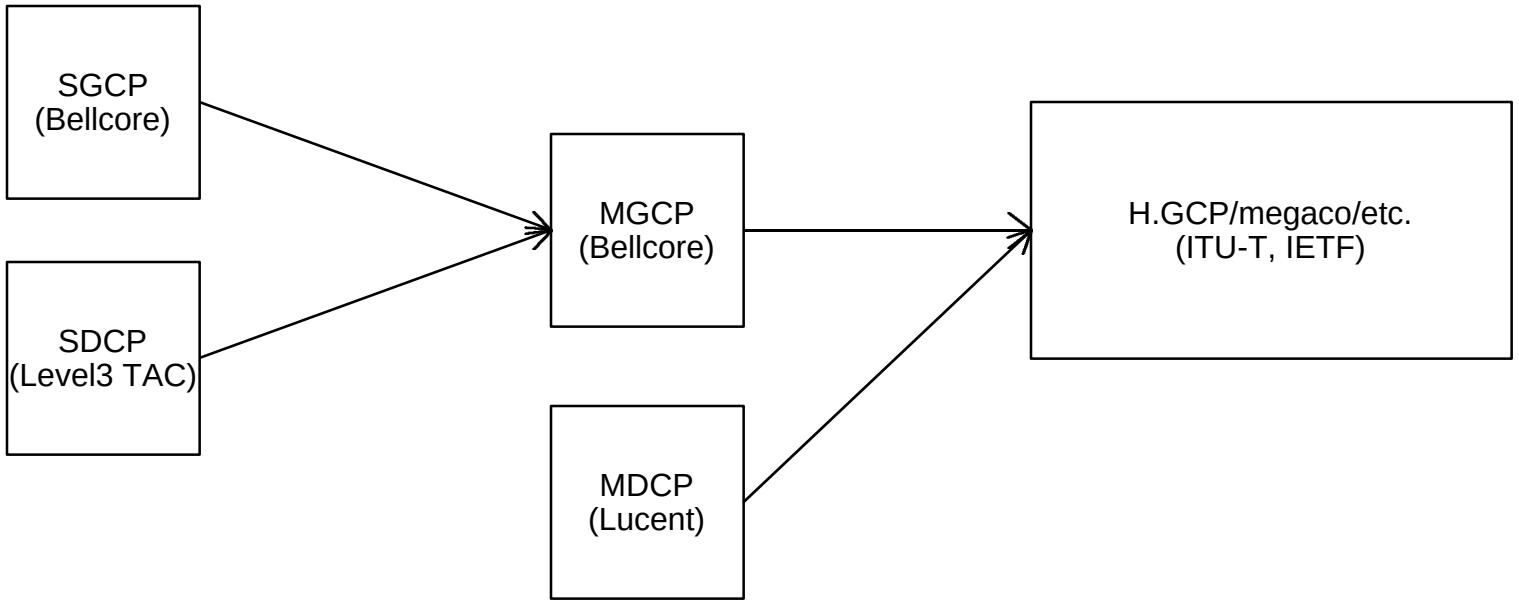
What media gateways do

- Connection control
  - Unicast
  - Multicast
  - Circuit to packet (IP)
  - Circuit to packet (ATM)
  - Packet to packet
  - Circuit to circuit
- Loopback testing
- The ability to identify/request endpoint attributes
  - The media protocol used (RTP, fax-protocol, ...)
  - The payload type (e.g. codec),
  - The codec-related attributes like packetisation interval, jitter buffer size and silence suppression where appropriate
  - The generation of comfort noise during silent periods.

What media gateways do(2)

- The ability to identify/request endpoint attributes
  - The application of encryption/decryption and identification of the encryption schemes.
  - The echo cancellation
  - The lawful interception of the content of a specified media stream
- Content insertion
  - Playing tone or announcement (IVR)
  - Mute request
  - Continuity testing, etc., as required by SS7 and others
- Event detection
  - On/off hook
  - DTMF
- Association management

Gateway control protocol evolution, roughly



A few words on signaling transport

- Two principal kinds of telephony signaling
  - In-band (“facility-associated”), for example T1
  - Common-channel, for example SS7
- In most models of decomposed gateways, signaling terminates in media gateway controller
- How to carry signaling from signaling gateway to MGC?
- sigtran (IETF signaling transport working group) adopting Motorola’s MDTP (Multi-Network Datagram Transmission Protocol) as base transport protocol

Questions?

Security

H.235

- H.235 is the security signaling framework for H.323
- Covers
  - Authentication
  - Call establishment (H.225) and call control (H.245) security
  - Media stream privacy
  - Trust relationships
- Allows call participants to signal choices of authentication and encryption mechanisms
- Interop agreements often provide “security profiles”

IMTC Security Profile 1 (SP1)

Security services	Call functions				
	RAS	H.225.0	H.245	RTP	Other(s)
Authentication	**HMAC-SHA1**	**HMAC-SHA1**	**HMAC-SHA1**		
Access control					
Non-repudiation					
Confidentiality				**Triple-DES/(40-bit) DES or RC2/IPSEC**	
Integrity	**HMAC-SHA1**	**HMAC-SHA1**	**HMAC-SHA1**		

Fun facts

- The European Union and (in the US) CALEA are *requiring* “lawful intercept” capabilities on all public telecommunications networks
- In Europe, this includes the internet, along with the ability to differentiate traffic types (email, web, etc., but also the ability to distinguish between signaling and data)

- It is extremely difficult to get H.323 through firewalls. NAT makes matters much, much worse. H.235 makes it just about impossible
- Several firewall vendors provide stateful inspection capabilities which understand H.323
- Proxies are also available
- Microsoft’s advice (concerning NetMeeting): Open all UDP ports > 1024

Numbering and addressing

Background

- Traditional telephony networks use combination of E.164 addressing and national numbering plans
  - E.164 is an ITU-T standard
  - Consists of
    - Country code
    - National destination code
    - Subscriber number
  - Should be dialable from any telephone on public network
  - 1-800 numbers and numbers like 911 and 411 are not E.164 numbers
- National telecom regulators are now mandating various levels of number portability
  - Local number portability (LNP) is required in major metropolitan areas in US, will be required nationwide over time
  - Service portability, number-for-life, etc. - these are being worked on

Background (2)

- IP uses a more layered approach to addressing and naming
  - MAC
  - IP
  - port (service)
  - “names”

Numbering and IP telephony

- Problem: How to locate a user/telephone number in IP networks
  - TIPHON/Tipia approach: Use E.164 address to locate gatekeeper
  - EP TIPHON and TIPIA working with ITU-T SG2 to allocate “country code” for IP telephony
  - Will be service-oriented
  - It is being argued that IP telephony will allow deployment of services (like number-for-life) which would be extremely difficult to do in traditional circuit networks
  - This assumes use of E.164 address
  - Lots of digits: 999 128.123.123.123
  - DNS probably can’t support transaction rate
  - It’s a really big database
  - Is it reasonable to tie telephone number to IP address?

To retrieve this presentation

`ftp://ftp.lightlink.com/pub/shore/usenix.ppt`