Voice Over IP: Architectures, Applications and Challenges

Tom Chapuran
Telcordia Technologies
tc@research.telcordia.com
973 829-4186

April 15, 2002
**What Is VoIP?**

Initially, PC to PC voice calls over the Internet

- Multimedia PC
- Gateway
- IP Network
- PSTN (NY)
- Multimedia PC
- Gateway
- PSTN (DC)

Gateways allow PCs to also reach phones

...or phones to reach phones
Outline

- Why Voice over IP?
- Packet voice transport
- Signaling and control architectures
- Network applications
- Outlook and challenges
Origins of VoIP

Lessons learned

• Internet Telephony software for multimedia PCs (1995)
  – Users frustrated by poor QoS, difficulty of use, lack of interoperability

• Standards are critical for success
  – Coding/decoding (codec) between analog voice and digital packets
  – Locating the party you want to call
  – Signaling to set up, modify, tear down the voice call
  – Access to vertical services (call forwarding, 3-way calling, …)
  – Gateways to PSTN

• Media routing, quality of service (QoS) left to other IP mechanisms
  (not VoIP-specific)
Circuit-Switched Telephony
Traditional PSTN Approach

SS7 Signaling Network

Class 4 Switch

Most service logic in local switches, rest in SCPs

Class 5 Switch

Typically analog “loop”, conversion to digital at local switch

Circuit-based Trunks

64 kb/s digital voice

Media stream

• Data travels over a parallel (but separate) network
VoIP
Goals and Potential Benefits

• Consolidation of voice, data on a single network
  – Simplify infrastructure, operations; provide bundled services
• Support for intelligent terminals as well as phones
• Increased flexibility
  – Multiple bit rates, multiple media types, richer signaling
  – Distinguish calls from connections (add/modify streams during call)
• Separation of service control from switching/routing
  – Accelerate new service development, increase end-user control, evolve from VoIP towards advanced services
• Expansion of competition
Packet Voice Transport

- Key targets for voice call service quality:
  - Average packet loss: < 2%
  - Consecutive packet loss: < 200 ms burst
  - End-to-end (lip-to-ear) delay: < 150 ms for comfortable conversation

- Packet loss cannot be corrected by retransmission (TCP), because the packets arrive too late to be useful

- Use RTP (Real-time Transport Protocol) over UDP (User Datagram Protocol) for voice or video transport
  - Payload ID, sequence numbers, timestamps, monitoring via RTCP

- Packet and buffer lengths limited by constraint on end-to-end delay

- Typical codecs: G.711 (64 kb/s), G.729 (8 kb/s) G.723 (~ 6 kb/s)
  - Transmitted bit rates depend on overheads, optional silence suppression
H.323 Architecture
ITU-T

3 stages of signaling:
- RAS to Gatekeeper
- H.225 call signaling
- H.245 media stream control (can be simplified for VoIP)

- Telco-centric multimedia, multiparty conferencing (initially for LANs)
- Gatekeeper for network control, heavy-weight protocols
- Widely deployed in first wave of VoIP standardization
SIP (Session Initiation Protocol)
IETF Multimedia Architecture

- Internet-centric alternative, initially for large multicast conferences
  - SIP for call signaling, SDP (Session Description Protocol) for media
- Initially very simple, light-weight, loosely-coupled sessions; oriented towards direct signaling between endpoints
- Network servers for additional capabilities:
  - Registrar for terminal registration, aliases
  - Redirect returns contact address directly to end user
  - Proxy forwards signaling (requests, responses)
- Evolution towards greater use of proxy/registrar for locating users, vertical services, call tracking, network control
- Strong, rapidly growing support (e.g., Microsoft XP, 3GPP)
SIP Call Setup
Simplified View

INVITE SDP proposes media type(s), IP & ports to send to
200 OK SDP accepts/rejects media, gives IP & ports to send to
Where Do Services Live?

- Some implemented at the endpoints
  - Last-number redial, call hold...
- Others may be better supported from the network
  - Avoid need for PC or IP phone to be turned on (call forwarding)
  - More complex services, such as conferencing
  - Integration with web-based services (unified messaging)
- Example: SIP Proxy runs a script for each incoming call for Peter
  - Parallel forking: forward INVITE to multiple endpoints simultaneously
  - Sequential forking: try his office PC first, then lab, then cell phone, ...
SIMPLE (SIP for IM and Presence)
Simplified Example

Linda subscribes to notifications of changes in Peter’s status:
Off-line, on-line, busy, away, available, ...
NGN Architecture
Next-Generation Network

- Oriented towards application of VoIP (or VoATM) to large-scale public networks
- Focus on scalability, network control, support for traditional phones, sophisticated gateway (GW) to the PSTN and its services
- Media GW interfaces voice stream to PSTN trunk or phone line
- Signaling GW allows signaling directly to SS7 network
- Softswitch controls Media GWs and does call processing
  - Allows smaller, cheaper Media GWs (e.g., for individual homes)
  - Control via MGCP (Media Gateway Control Protocol) or H.248
NGN Example
Voice over DSL or Cable Modem

Can also use to interconnect PSTN clouds (long-distance), or PSTN switches (interoffice backbone)
International Voice Market
Calls Terminated on PSTN

Source: Telegeography 2001
(2001 figures were projections)
Carrier Applications of VoIP

- First major inroads for VoIP have been in long-distance
  - Avoid regulation, high international PSTN tariffs
  - VoIP invisible to end user, doesn’t rely on him to do anything
  - Installed base dominantly H.323, movement now towards NGN

- Local-carrier interest for interoffice connections
  - Consolidate voice and data networks (typically ATM)
  - Use NGN, or packet-enable existing switches

- Many trials of VoIP to residences, but deployments few
  - Cable TV has laid groundwork for NGN approach (DOCSIS 1.1)
  - Decline of CLECs likely to slow multi-line VoDSL
Enterprise VoIP

Location A

IP PBX

IP phone

Location B

PSTN

Centrex or PBX

Softswitch

GW GW

Core IP Network

Many possible combinations of VoIP and circuit-switched telephony
Enterprise Applications of VoIP

- Leverage spare data-network capacity, minimize phone bills, create platform for multimedia conferencing
- H.323 and SIP both being deployed, softswitches and IP-PBX options emerging, unclear which will prevail
- Examples: Telcordia/SAIC (H.323), Telia (SIP)
- Carrier-managed VPN networks last year from AT&T (H.323) and Worldcom (SIP)
- VoIP adoption slower than expected, partly due to plunging PSTN long-distance prices, QoS concerns
Peer-to-Peer VoIP
PC-to-PC

- Internet Telephony revisited, often facilitated by software or network servers from new types of voice service providers
  - Microsoft, Net2Phone, Dialpad, AOL, Yahoo!
  - Mass market alternative to telcos, requiring limited network infrastructure, capital costs, operating expenses
- What’s the business case for “free” VoIP?
  - Sell advertising, software, or enhanced services
  - Charge for PC-to-phone, phone-to-phone
  - Give away as a competitive differentiator
- Mostly H.323 today, likely to move towards SIP
- Could be key industry driver, even if penetration were limited
Outlook for VoIP

Current Status and Trends

- VoIP is not monolithic – many applications, with different drivers, will maintain a heterogeneous mix of technologies
- H.323 is most widely implemented today, but trends are towards SIP for intelligent terminals, NGN for most carrier networks
- Most success thus far in long-distance networks, perhaps with local carrier backbones to follow in next few years
- Footholds made in enterprise and access markets, but VoIP has not taken off as fast as initially expected
- Adoption being slowed by economic conditions, plummeting long distance rates, declining advertising market (peer-to-peer)
Continuing Challenges

• Quality of Service
  – Diffserv, MPLS, traffic engineering, bandwidth brokers, call admission…
  – What is really needed for consolidated voice and data networks?
• Security, reliability
• Extending SIP to provide conference control
• Operations (configuration of IP phones, version control and upgrading of highly distributed software, accounting/billing,…)
• Packet-level interconnection of VoIP islands which use competing architectures and protocols
• Controlling feature interactions in a distributed-services environment
  • Traversal of NATs and firewalls
  • Support for services beyond voice
NAT Traversal

• Network Address Translators (NATs) map a private IP address space to externally visible (public) IP addresses
  – Conserve scarce public IP addresses
  – Shield internal hosts from outside world
• Useful for enterprises, cable modem networks, broadband access routers, internet cafes…
• NATs interfere with peer-to-peer protocols such as SIP
  – SIP clients must identify the IP address and ports they will use to receive media streams (in payload of their signaling messages)
  – But they don’t know their externally visible addresses
• “One of the SIP community’s biggest problems”
**STUN – Simple Traversal of UDP Through NATs**
draft-rosenberg-midcom-stun-01.txt

- STUN client contacts STUN server, discovers NAT, address translation
- SIP client uses “external” address in signaling for setup of media streams
- This approach being implemented and tested at Columbia and LTS

*Source: P. Therms, Telcordia*
Advanced Services

• VoIP: natural platform for evolution to advanced services
  – Supports intelligent terminals and rich signaling
  – Separates calls from connections
  – Multimedia capabilities already in the protocols (SIP/H.323)
  – Removes bottleneck by separating call control from switching

• Thus far, focus is almost entirely on voice
  – For many players (but not all), voice is the killer app
  – Solve the simpler problem first

• This simplifies many network control issues, because of predictability of voice bandwidth, traffic patterns
  – But current solutions are likely to require significant extensions to accommodate more flexible advanced services
Moving Beyond Two-Party Voice
What’s Different About Advanced Services?

• Flexibility in media streams, participants, “ownership”; service not pre-defined at call setup
  – Multiple media per call, differing (and very wide range of) bandwidths
  – Dynamic reconfigurability during call
  – Potential for multicast conferencing, streaming

• Implications
  – Call admission control becomes more complex
  – Much less aggregation, localization of flows than with NGN voice
  – Usage, traffic patterns may be highly variable and hard to predict

• New approaches to traffic engineering, resource allocation and network control will be needed to address even a modest penetration of these new services
Acknowledgements

Ron Menendez
Stu Wagner
Tim Feustel
Peter Thermos
Dave Gorman
Nigel Dewdney
Gary Hayward