CS 414 – Multimedia Systems Design
Lecture 36 – Voice-over-IP/Skype

Klara Nahrstedt

Spring 2009

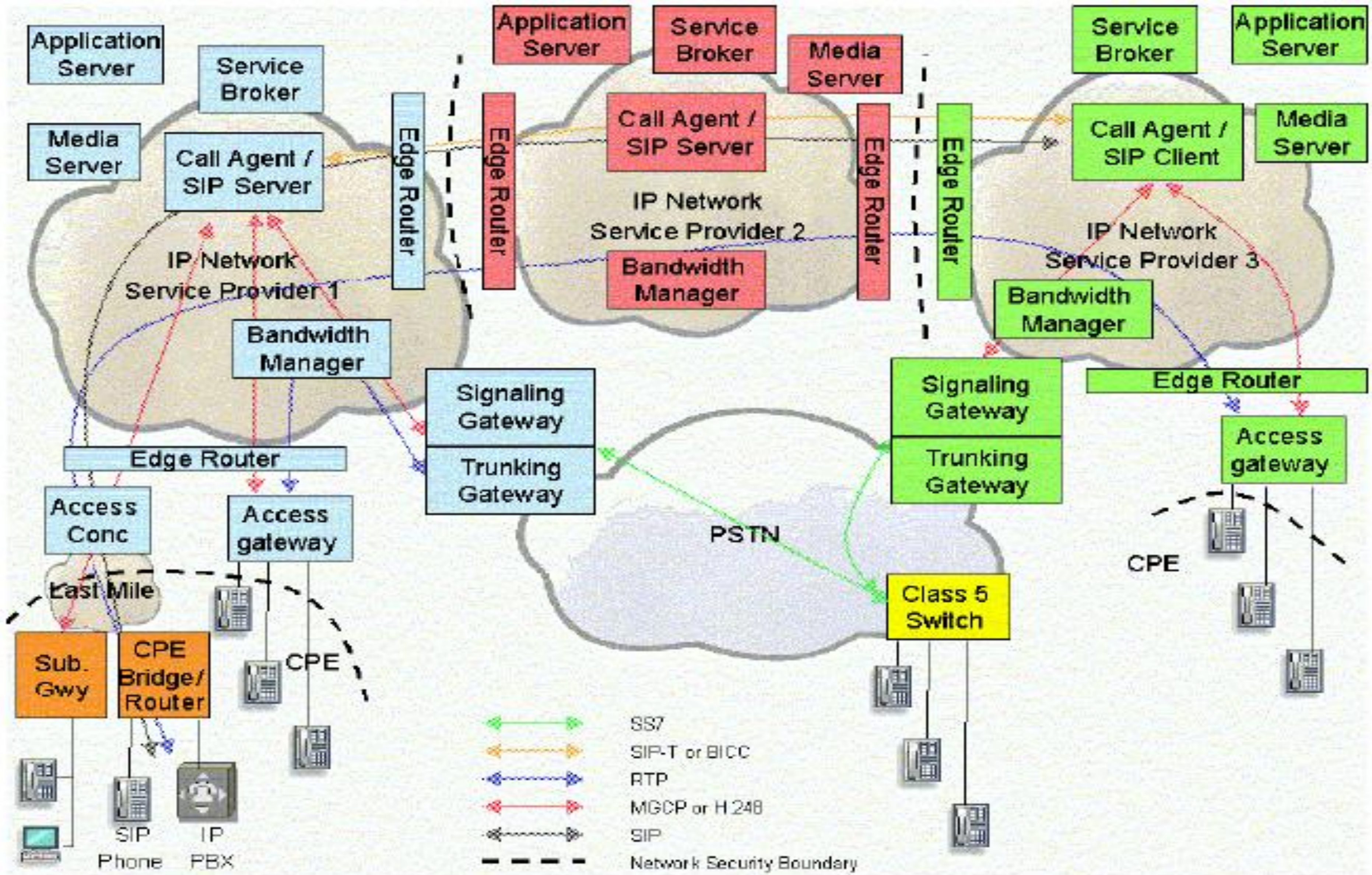
Outline

- Voice-over-IP Basic Principles
- Skype – first VoIP over Peer-to-peer Infrastructure

Voice over IP (VoIP)

- VoIP – transport of voice over IP-based networks
- Complexity ranges from
  - Hobbyists using Internet to get free phone calls on peer-to-peer basis to
  - Full scale PSTN replacement networks
- VoIP must address
  - Types of end user terminals - IP phones, PC clients
  - Quality of Service – ensure agreed quality
  - Security risks must be clearly identified
  - Last mile bandwidth – which affects codec, packetization period and where to use compression to best meet service goals
  - Signaling protocol must support service set required

Next Generation VoIP Network (MSF Example)



MSF VoIP

- Access Services Signaling protocol and network service signaling protocol: SIP
  - Use RTP packets for telephony events
  - Transport DTMF tones out of band using the signaling protocol such as SIP
- Quality of Service (Delay, Jitter, Packet loss)
  - Use RSVP, DiffServ, MPLS, even ATM
  - RTP is used for media traffic

Skype

Source: An Analysis of the Skype Peer-to-peer Internet Telephony Protocol, S. Baset, H. Schulzrinne,

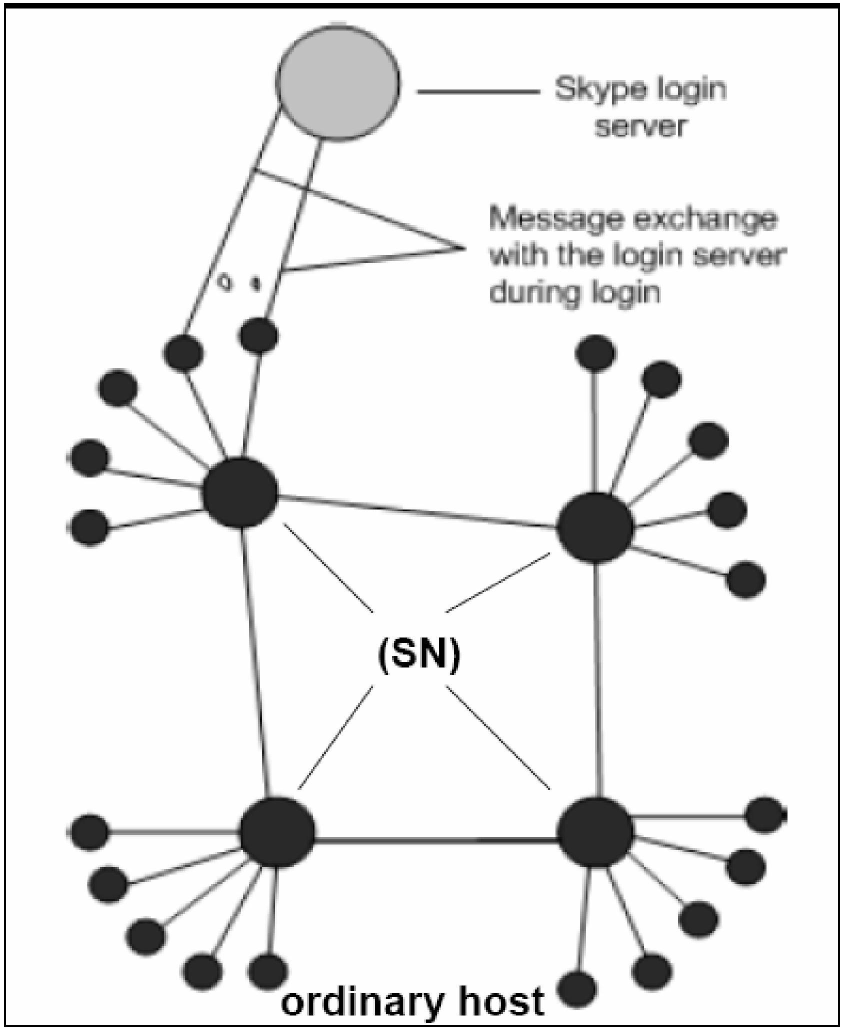
2004

Skype Overview

- Peer-to-peer (P2P) overlay network for Voice-over-IP (VoIP) and other application
- Developed by Niklas Zennstrom and Janus Friis (founders of KaZaA, file-sharing company)
- Users see Skype as an Instant Messaging (IM) software
- Free on-net VoIP service and fee-based off-net SkypeOut service (allows calling to PSTN and mobile phones)
- Runs on Windows, Linux, Pocket PC, …

Skype Network

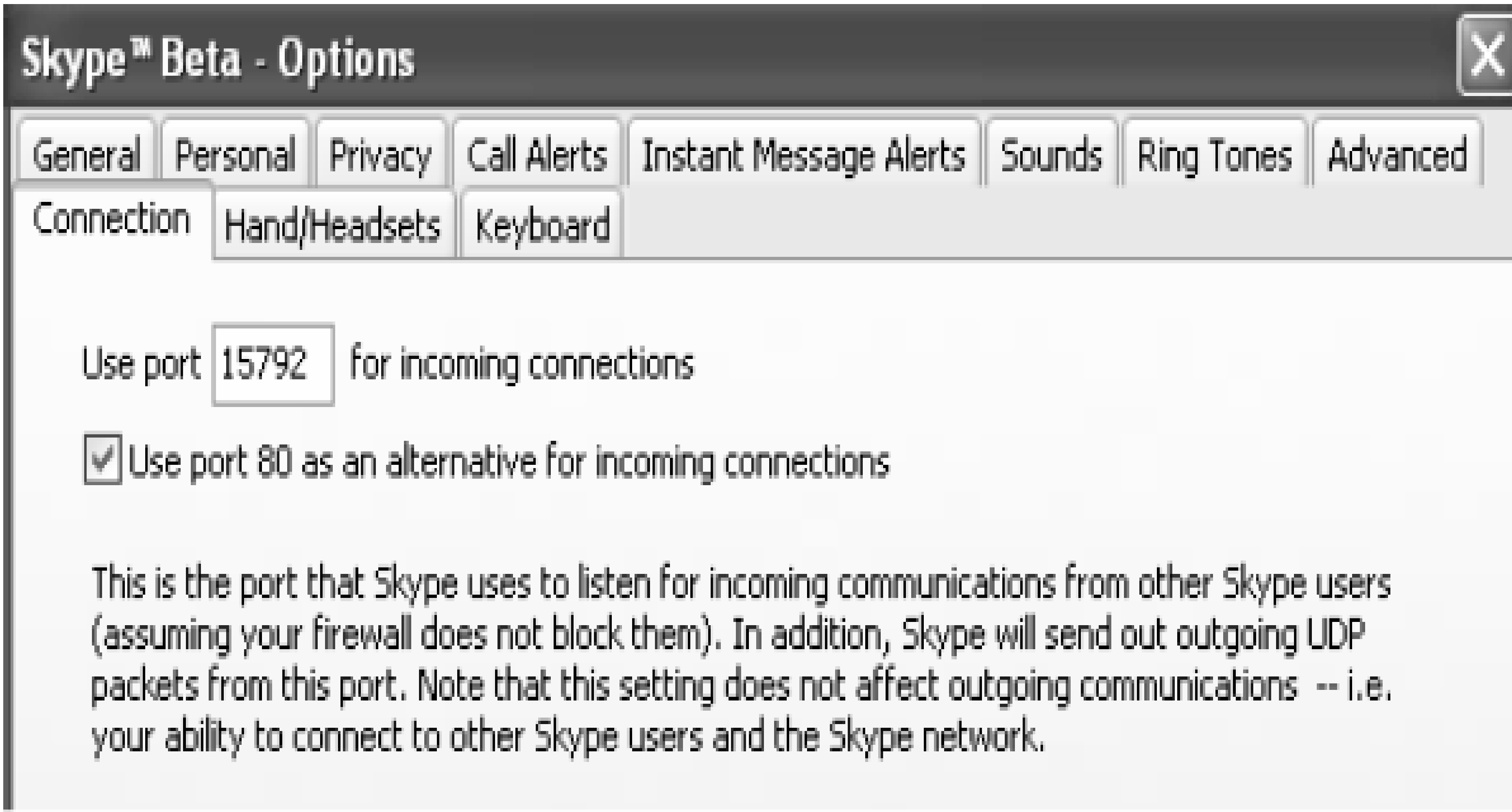
- Super Nodes: Any node with a public IP address having sufficient CPU, memory and network bandwidth is candidate to become a super node
- Ordinary Host: this host needs to connect to super node and must register itself with the Skype login server



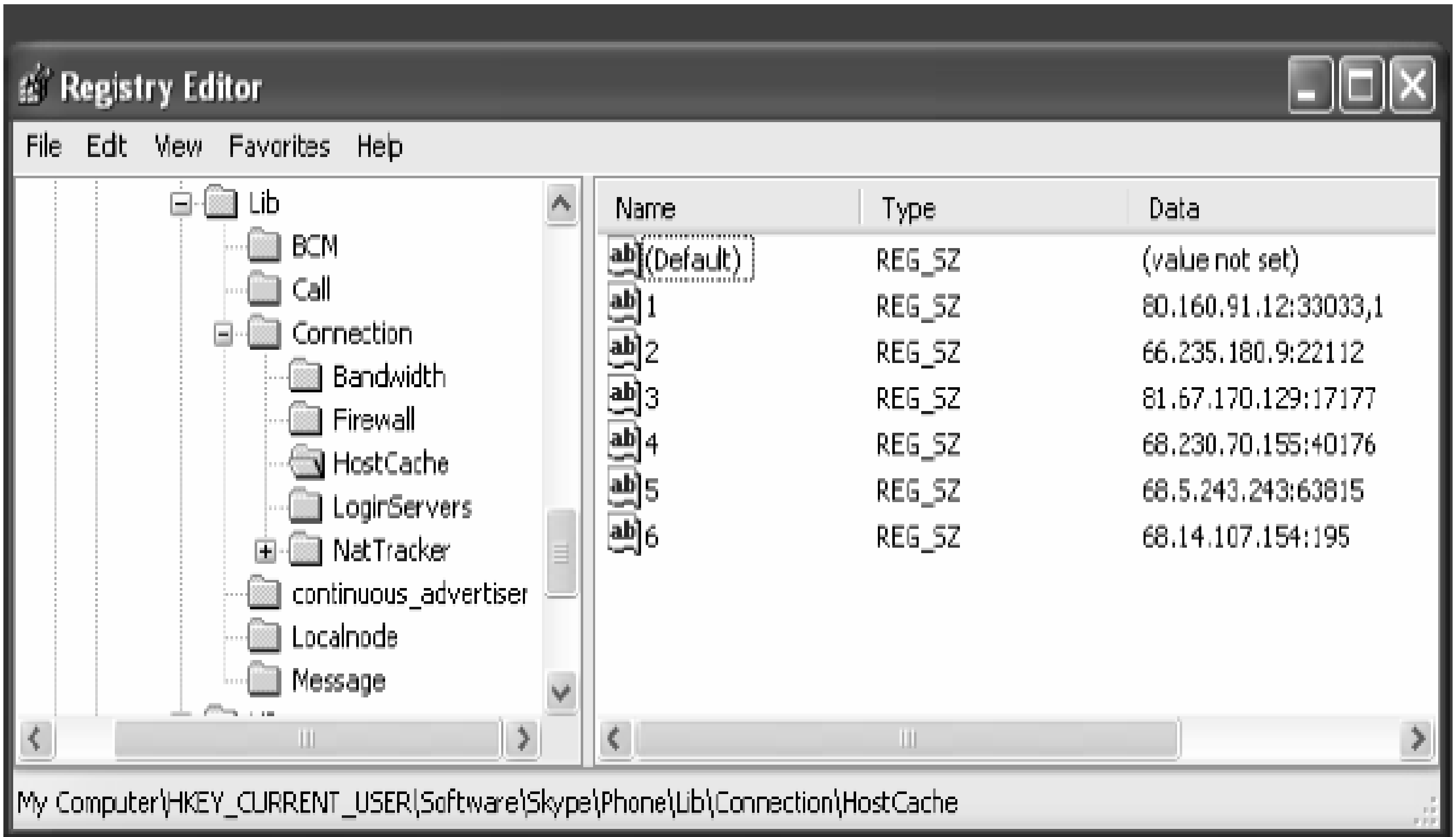
Components of Skype

- Ports
  - Skype client (SC) opens TCP and UDP listening port configured in its connection dialog box
- Host Cache (HC)
  - List of super node IP address and port pairs that SC builds and refreshes regularly
  - SC stores HC in the Windows registry
- Codecs
  - Wideband coded allowing frequencies between 50Hz-8KHz (one of the codecs is implemented by Global IP Sound)

Skype Ports on which Skype listens for incoming connections



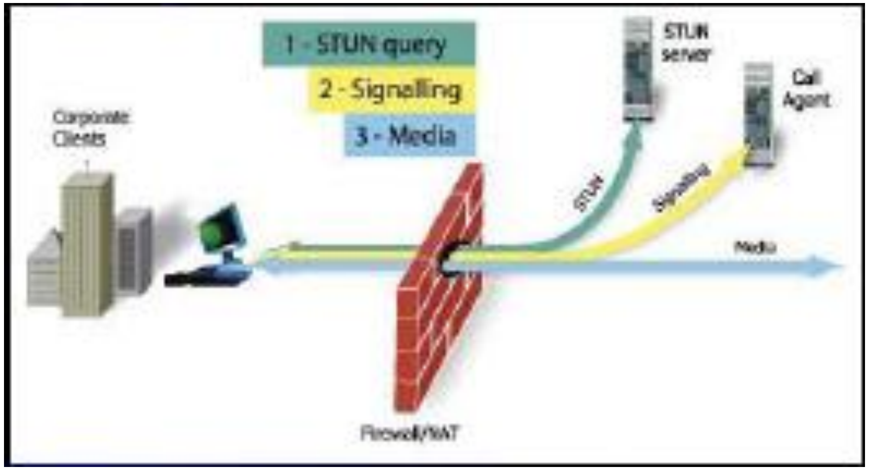
Skype Host Cache List

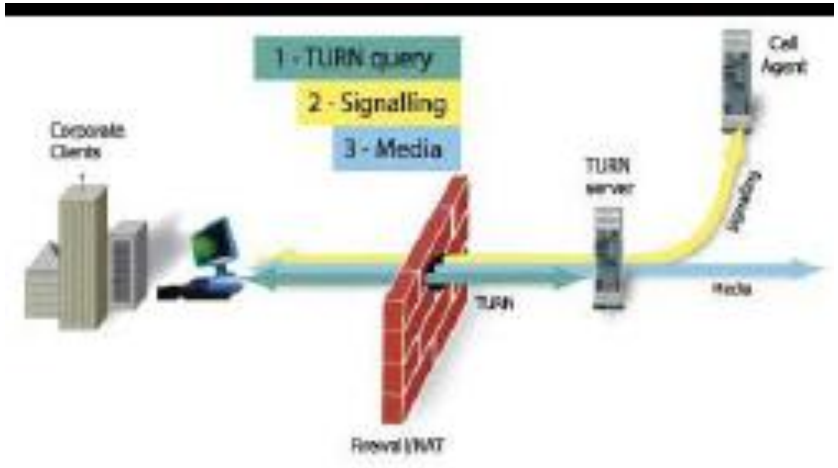


Components of Skype

- Buddy List
  - Skype stores buddy information in Windows registry
  - Buddy list is digitally signed and encrypted, local to machine and not on a central server
- Encryption
  - Skype uses 256-bit AES encryption
  - Skype uses 1536 to 2048bit RSA to negotiate symmetric AES keys
- NAT and Firewall
  - SC uses variations of the STUN and TURN protocols to determine type of NAT and firewall

STUN and TURN





- STUN (Simple Traversal of UDP through NAT)
  - Does not work through symmetric NAT
- TURN (Traversal Using Relay NAT)
  - Increase latency and packet loss

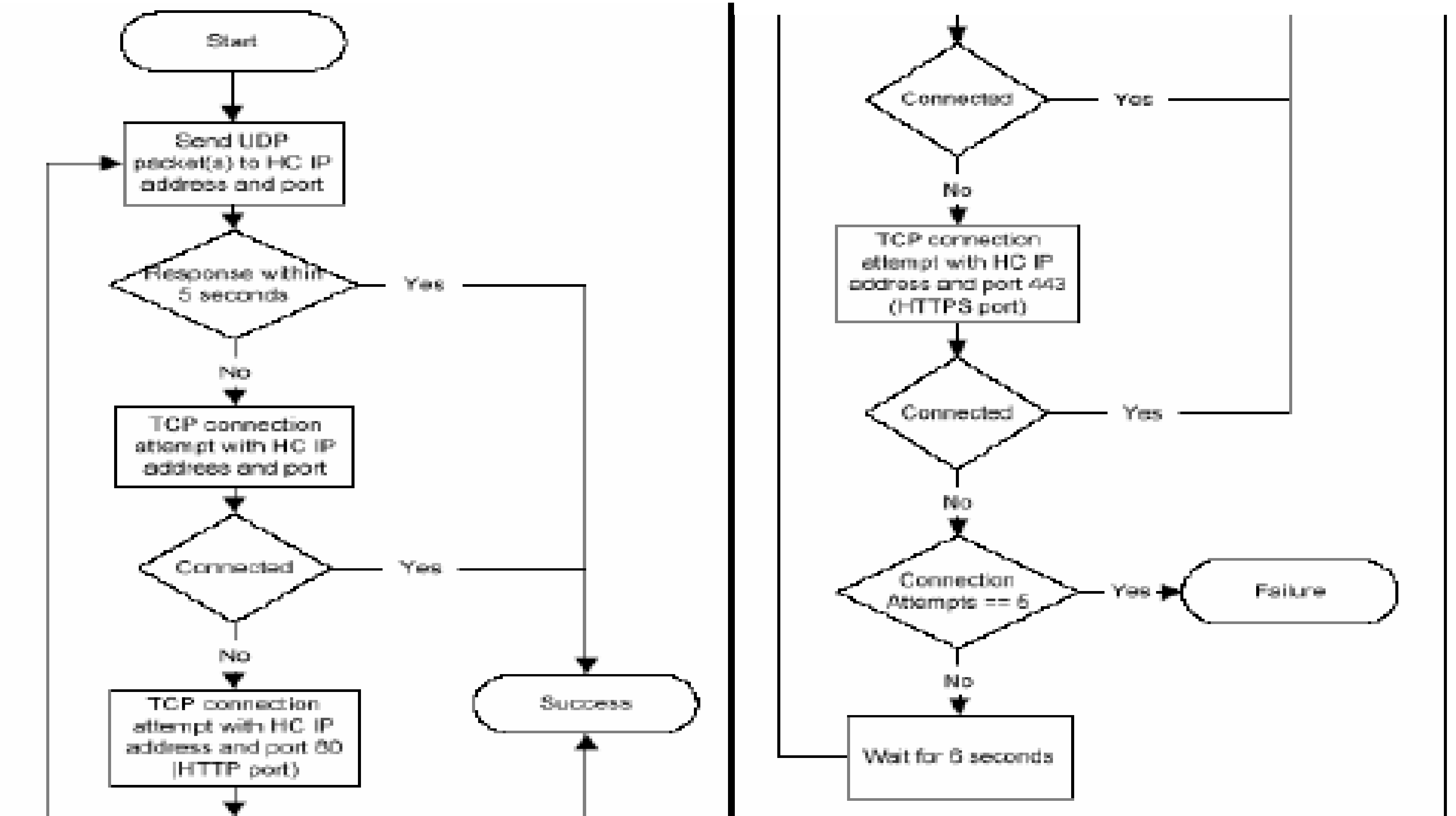
Techniques used in Skype

- Firewall and NAT traversal
- Global decentralized user directory
- Intelligent routing
- Security
- Super-simple UI

Login

- During login process SC:
  - Authenticates its user name and password with login server
  - Advertises its presence to other peers and its buddies
  - Determines type of NAT and firewall it is behind
  - Discovers online Skype nodes with public IP addresses
- Login server is the only central component in Skype network

Skype Login Algorithm



Skype Login Process

- After installation and first time startup, HC was observed empty
- Bootstrap supper nodes:
  - After login for the first time after installation, HC was initialized with seven (IP,port) pairs
- Bootstrap (IP,port) information either
  - Hard coded in SC
  - Encrypted and not directly visible in Skype Windows registry, or
  - One-time process to contact bootstrap node

Skype Login Process

- First time Login Process
  - SC sends UDP packets to some bootstrap SNs
  - SC establishes TCP connection with bootstrap SNs that respond
  - SC perhaps acquires address of login server from SNs
  - SC establishes TCP connection with login server, exchanges authentication information
- Subsequent Login Process
  - Similar to first-time login process
  - SC uses login algorithm to determine at least one available peer and establishes TCP connection
  - HC was periodically updated with new peers’ (IP,port)

Skype Login Process

- Comparison of three network setups
  - Exp:A both Skype users with public IP address
  - Exp B: one Skype user behind port-restricted NAT
  - Exp C: Both Skype users behind port-restricted NAT and UDP-restricted firewall
- Message flows for first time login process
  - Exp A and Exp B are roughly the same;
  - Exp C only exchange info over TCP

	Total Data Exchanged	Login Process Time
Exp A	About 9 KB	3~7 seconds
Exp B	About 10 KB	3~7 seconds
Exp C	About 8.5 KB	About 34 seconds

User Search

- Skype uses Global Index technology to search for a user
- Skype claims that search is distributed and is guaranteed to find a user if it exists and has logged in during last 72 hours
- Search results are observed to be cached at intermediate nodes

Call Establishment and Teardown

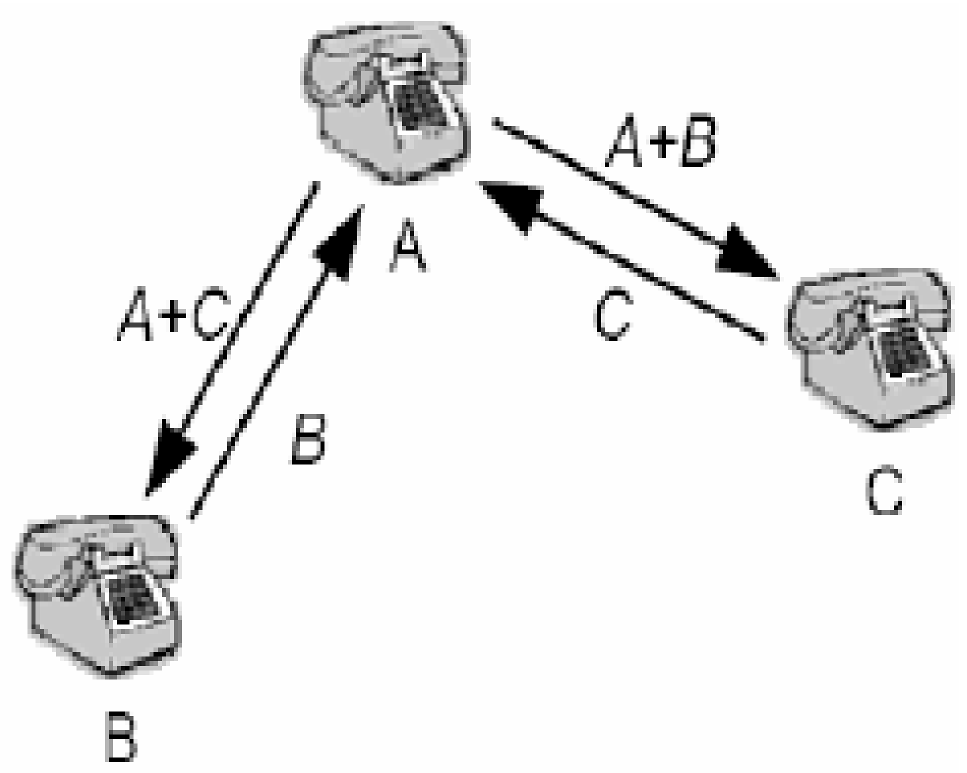
- Call signaling is always carried over TCP
- For user not present in buddy list, call placement is equal to user search plus call signaling
- If caller is behind port-restricted NAT and callee is on public IP, signaling and media flow through an online Skype node which forwards signaling to callee over TCP and routes media over UDP
- If both users are behind port-restricted NAT and UDP-restricted firewall, both caller and callee SCs exchange signaling over TCP with another online Skype node, which also forwards media between caller and calllee over TCP

Media Transfer and Codec

- Bandwidth usage: 3-16 Kbytes/s
- Skype allows peers to hold a call. To ensure UDP binding, SC sends three UDP packets per second to the call peer on average
- No silence suppression is supported in Skype
- The min. and max. audible frequencies Skype codecs allow to pass through are 50 Hz and 8000 Hz.
- Uplink and downlink bandwidth of 2KB/s each is necessary for reasonable call quality

Conferencing

- Node A acts as mixer, mixing its own packets with those of node B and sending to C and vice versa
- For three party conference, Skype does not do full mesh conferencing
- Most powerful machine will be elected as conference host and mixer
- Two-way call: 36kb/s
- Three-way call: 54kb/s



Impact of Skype

- Impact on fixed-line operator
  - Skype will introduce SkypIN
- Impact on mobile phone operator
  - Skype will be embedded in Wi-Fi/mobile phone
  - WLAN is now limited by
    - Not many Wi-Fi phone models
    - Wi-Fi phone's high price
    - Batter life
    - Not enough hot-spots

Impact of Skype

- Skype has shown, at least has suggested, the following
  - Signaling, the most unique property of traditional phone systems, can now be accomplished effortlessly with self-organizing P2P networks
  - P2P overlay networks can scale up to handle large-scale connection-oriented real-time services such as voice

Conclusion

- More than 2 million on-line subscribers per day
- More than 2.7 billion minutes served (minutes of free Skype-to-Skype calles)
- More than 38 million of software download
- More than 7 million of registered subscribers
- More than 1 million concurrently on-line subscribers