Real-time Communications on the Internet

Jon Peterson

Stanford - June 2008

NEUSTAR®

The Origins of Telephony

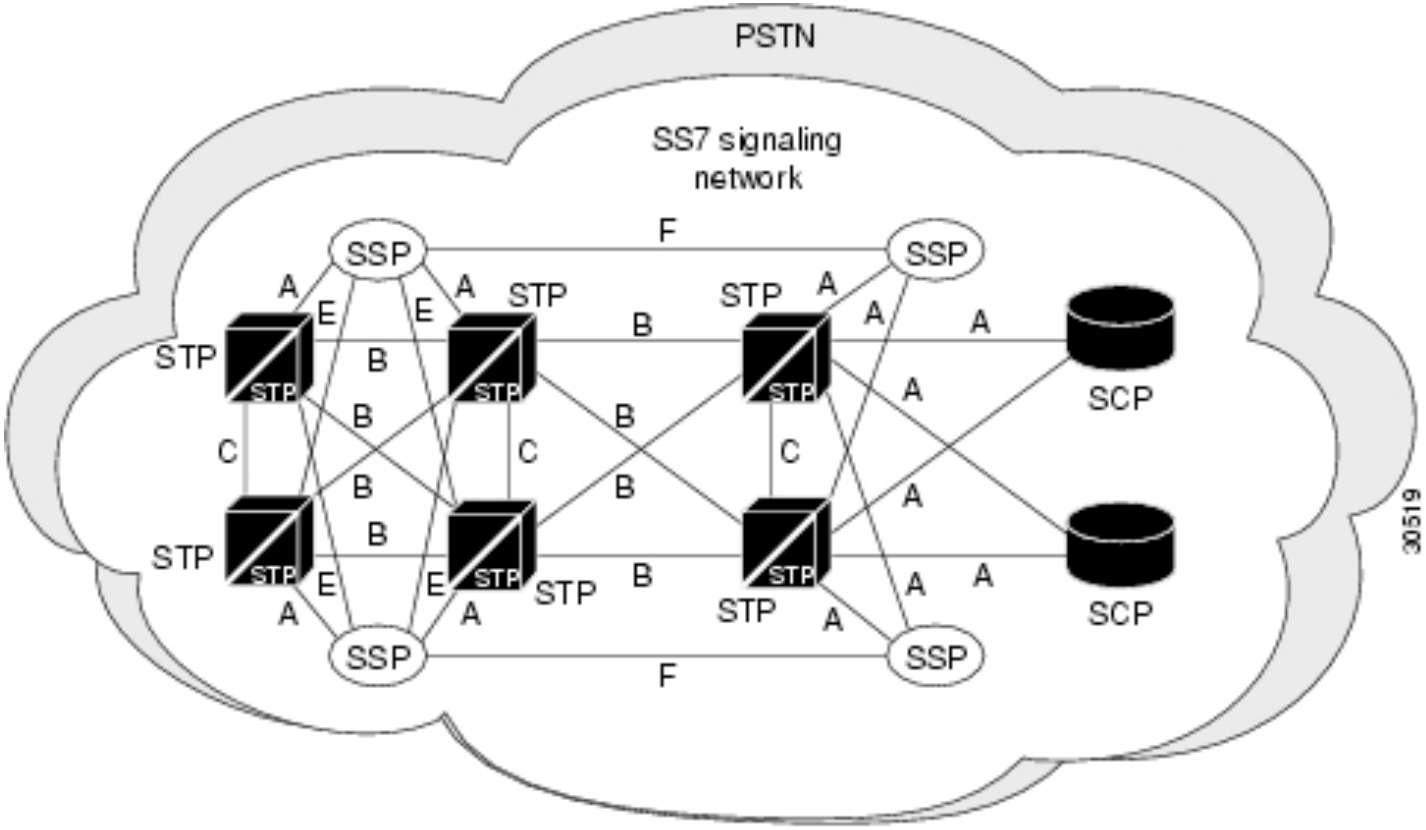
- Point-to-point audio over electricity
  - One step up from tin cans and string
- To enable contact with more than one party, "switches"
  - Something like rail yard switching initially
  - The first switches were at the endpoints themselves
    - An n-squared problem
  - Subsequently switches were centralized and manned by operators
- Trunks, audio channels between switches
  - Tandems, switches of switches

NEUSTAR®

Signaling and Media

- User instructions to the switch are a form of signaling
  - Signals are also sent between switches
- Rotary dialing
  - Rapid on-off hook to signal addresses to the switch
- Dual Tone Multi Frequency (DTMF)
  - Familiar touchtone keypad of phones today
  - Between switches, MF or CAS
    - “Fancy” tones like 2600Hz and 3700Hz
  - Downside: Using bearer to find out someone is busy
  - Downside: Some significant security problems
- Thus signaling and media were separated
  - Audio media on one network, signals on another

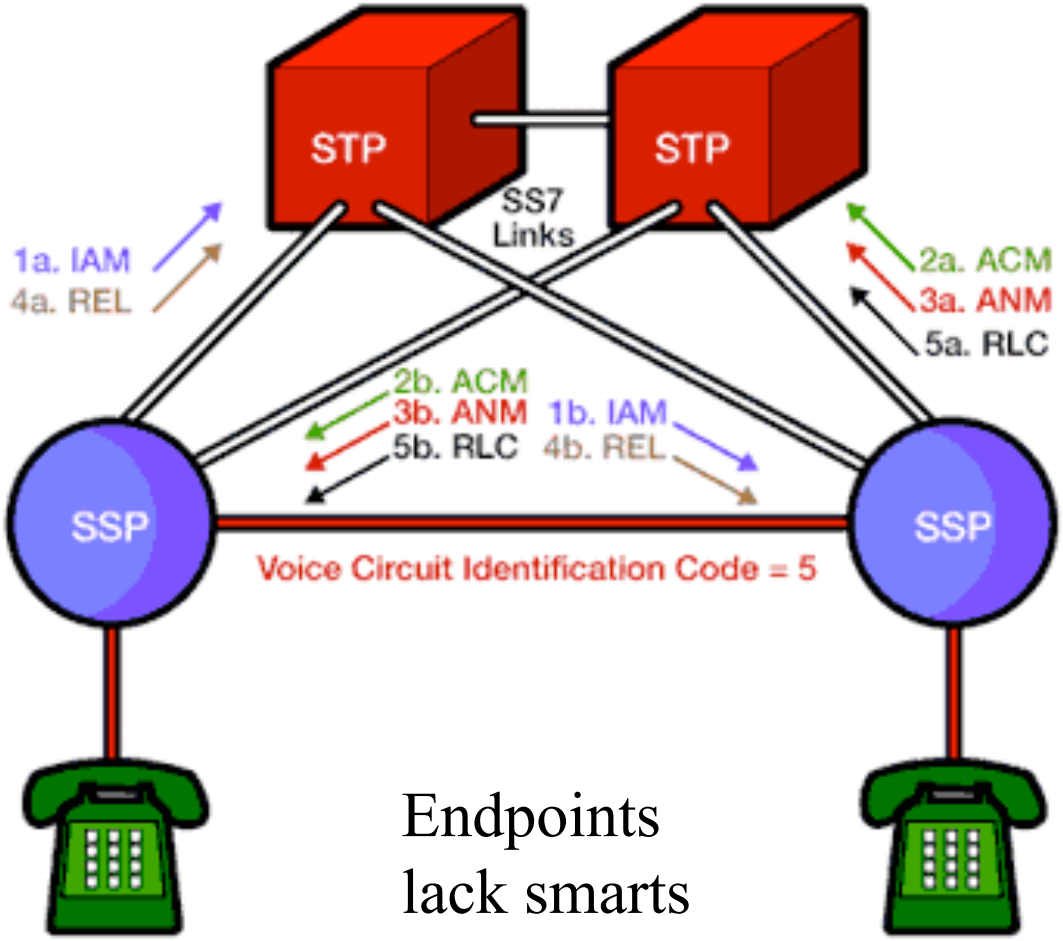
The SS7 Network



http://www.cisco.com/univercd/illus/3/19/30519.jpg

NEUSTAR®

Phones and the SS7 Call Flow

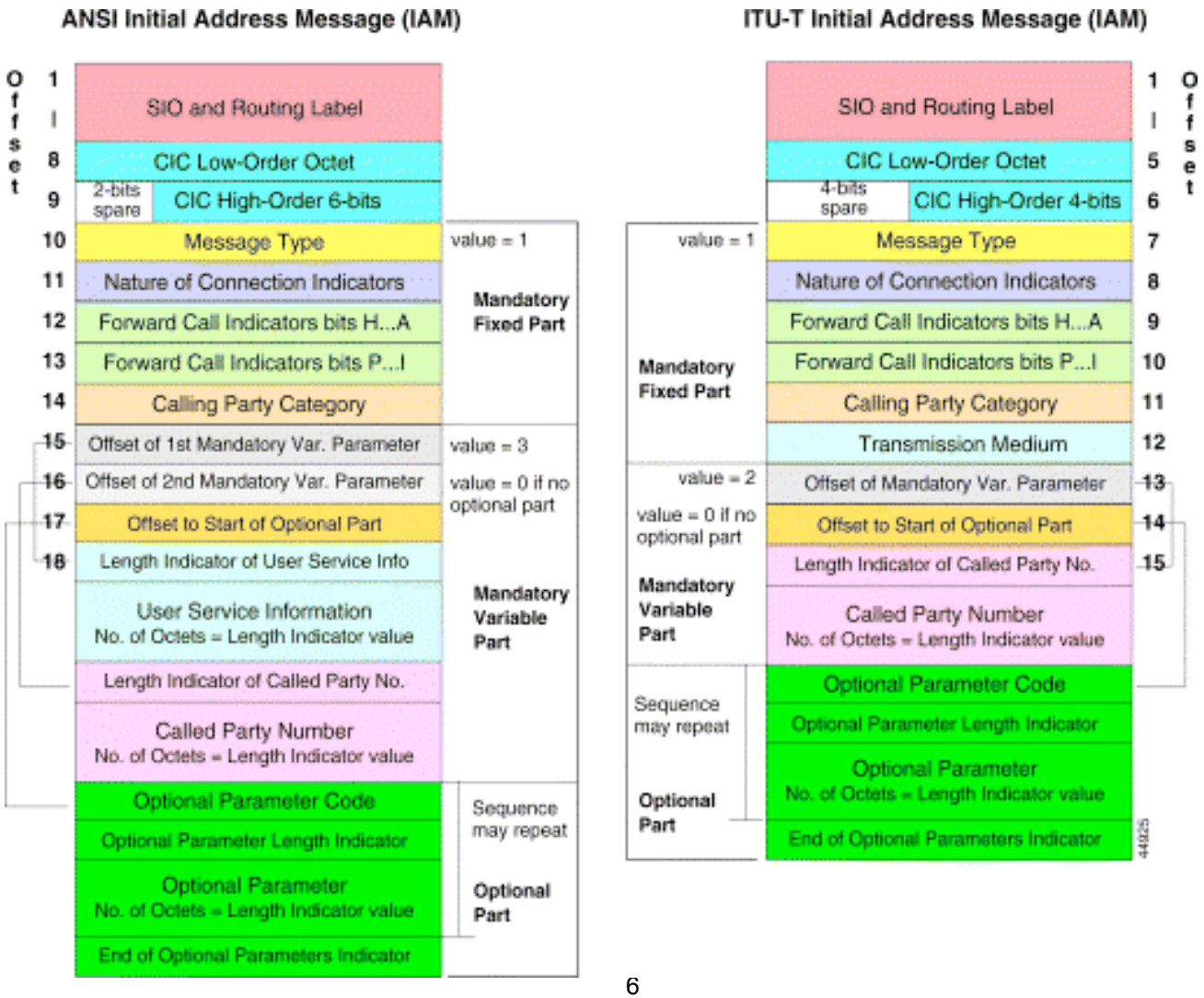


Endpoints
lack smarts

www.pt.com/tutorials/ss7/isup.html

NEUSTAR®

Inside the IAM



NEUSTAR®

The Limits of the PSTN

- Scope limited to audio
  - Lousy codec too
- Intelligence lives in the network, not the endpoints
  - Endpoints cannot innovate without network support
    - e.g., your phone cannot choose a new codec
- Per-call circuit allocation
  - Capacity management
- Service-based billing
  - Many phone companies are in the billing business rather than the communications business
- Over-engineering
  - 99.999% reliability and the resulting inflexibility

The Emergence of the Internet

- Telephony trunks (T1s, E1s, T3s) used for raw data rather than audio bearer channels
- Modems tunnel point-to-point data over audio
- Voice over IP (VoIP) begins to appear. Why?
  - Hobbyism and science (programs like vat)
  - Costs less than PSTN
    - Lack of regularity environment / taxation
  - Packetization uses less bandwidth
    - Silence suppression is a big factor
  - Integration with existing Internet applications
  - Promise of new and exciting services

The Internet Model

- A layer 3 overlay on a diverse set of layer 2 networks, with the property that:
  - Any node on the network can send arbitrary data to any node associated with a public Internet address

- Higher layers do not assume any properties of layer 3 or below other than that
- Every layer 2 network could, and should, and will change, but the Internet is unperturbed
  - An assumption about layer 1 and 2: if you want more of it, you can buy it

Building Services on the Internet

- Packet forwarding on the Internet lacks smarts
  - Endpoints have application intelligence
- The Internet is the service
  - Thus telephony on the Internet is an application
    - Anyone can just write a new VoIP program tomorrow
- Reuse existing Internet standards
  - Steal from success: email, web
  - Rely on existing support systems: DNS
  - Use human-readable encoding for applications used by humans

The Real-time Communications Suite

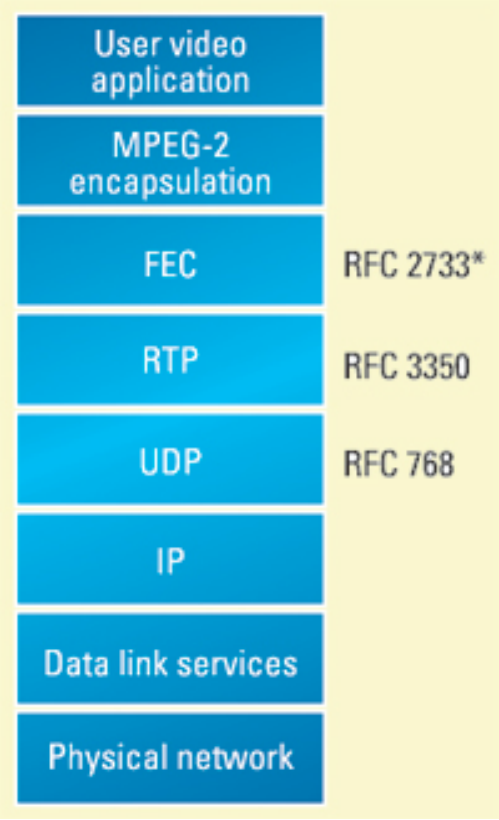
- Real-time Transport Protocol
  - RTP (RFC1889 in 1993, now RFC3550)
- Session Description Protocol
  - SDP (RFC2327, …)
- Session Initiation Protocol
  - SIP (RFC2543, now RFC3261)

- Related:
  - Real-time Streaming Protocol (RTSP)
  - Session Announcement Protocol (SAP)

RTP (Real-time Transport Protocol)

- (Not actually a transport protocol)
  - Runs over UDP, a transport
  - Thus, unreliable delivery
    - No head-of-line-blocking
    - No congestion control
- Carries a media payload

- Also, an adjunct protocol
  - RTCP: Real-time Control Protocol
  - Provides diagnostic information



NEUSTAR®

Session Description Protocol (SDP)

- Describes a session (IP, ports, codecs)

user name | session identifier | session version | network type | address type | unicast address

version	v=0
origin	o=alice 2890844526 2890844526 IN IP4 host.atlanta.example.com
session-name	s=
connection	c=IN IP4 host.atlanta.example.com
time	t=0 0
media	m=audio 49170 RTP/AVP 0 8 97
attribute	a=rtpmap:0 PCMU/8000
attribute	a=rtpmap:8 PCMA/8000
attribute	a=rtpmap:97 iLBC/8000
media	m=video 51372 RTP/AVP 31 32
attribute	a=rtpmap:31 H261/90000
attribute	a=rtpmap:32 MPV/90000

media description

media description

media type | transport port | transport protocol | media formats

NEUSTAR®

Internet Applications

- Delay Tolerant
  - Email
  - Newsgroups
  - File transfer
    - Including client/server and peer-to-peer
  - Web
    - (somewhat)
- Real-time
  - Instant messaging
  - VoIP
  - Videoconferencing
    - And streaming video
  - Internet gaming
  - Shell

SDP sets up these sorts of sessions

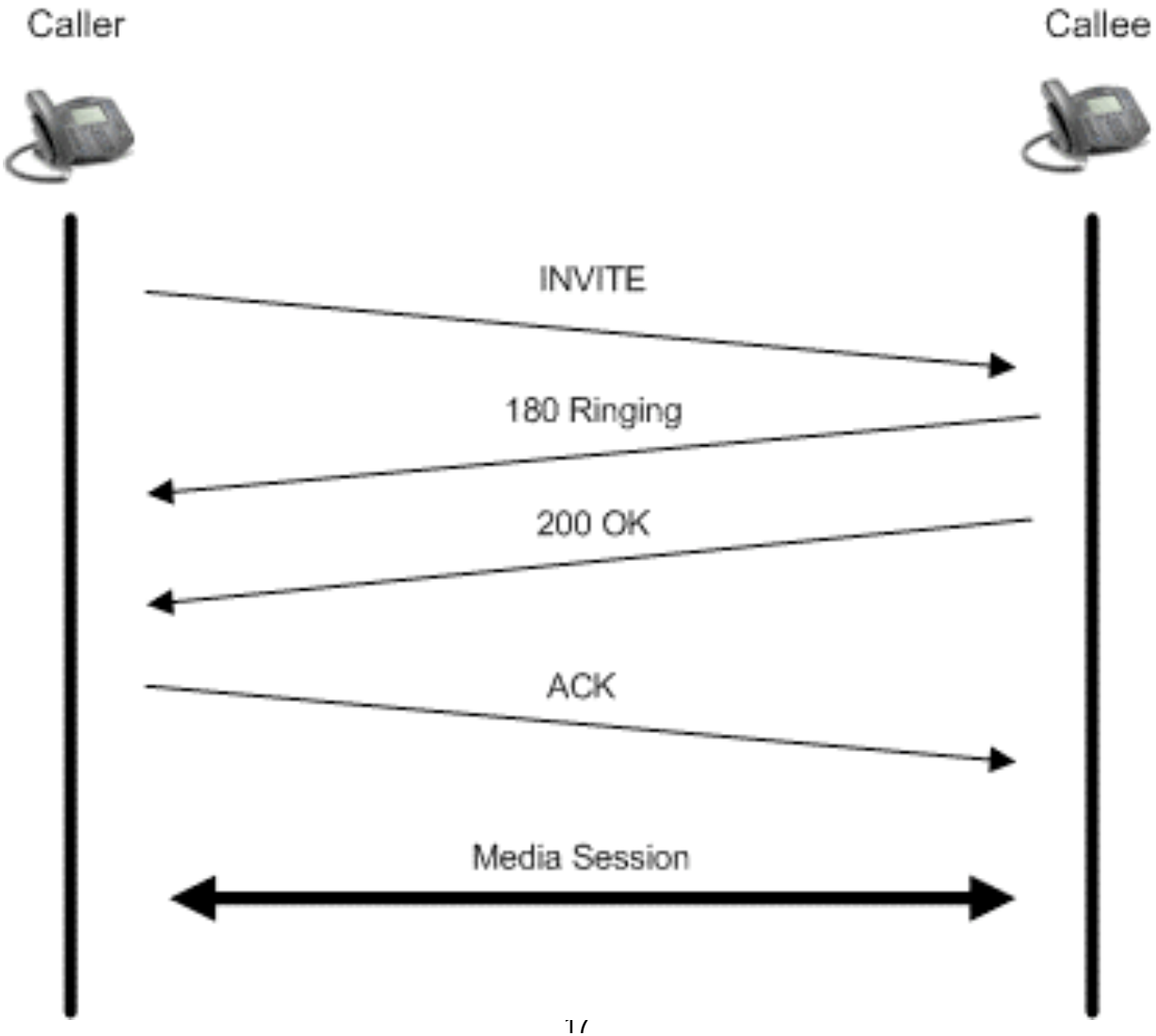
Session Initiation Protocol

- Provides a rendez-vous function
  - Discovers the endpoint associated with a user
  - Users have “address of record” (AoR) URIs
    - Acts as a persistent identifier, e.g. sip:jon@example.com
  - The rendez-vous function maps AoRs to a URI representing a device
    - Devices URIs are associated with AoRs through registration
- Exchanges SDP
  - Offer/answer model
  - Allows endpoints to negotiate session preferences and needs
  - Can also exchange other MIME types
- Offers additional services in support of session management

SIP Entities

- User Agent
  - Endpoint, acts on behalf of the user
  - Could be a physical phone, an application on a PC
- Proxy Server
  - Application-layer routing entity
  - Makes forwarding decisions and serves as a point of policy enforcement
- Registrar
  - Receives registrations as users become associated with devices

Basic SIP Session Establishment



NEUSTAR®

A Child of Email: Example INVITE

```
INVITE sip:bob@example.com SIP/2.0
Via: SIP/2.0/UDP 10.20.30.40:5060
From: Alice <sip:alice@example.com>;tag=589304
To: Bob <sip:bob@example.com>
Call-ID: 8204589102@example.com
Subject: Dinner tomorrow
CSeq: 1 INVITE
Contact: <sip:alice@10.20.30.40>
Content-Type: application/sdp
Content-Length: 141

v=0
o=UserA 2890844526 2890844526 IN IP4 10.20.30.40
s=Session SDP
c=IN IP4 10.20.30.40
t=3034423619 0
m=audio 49170 RTP/AVP 0
a=rtpmap:0 PCMU/8000
```


NEUSTAR®

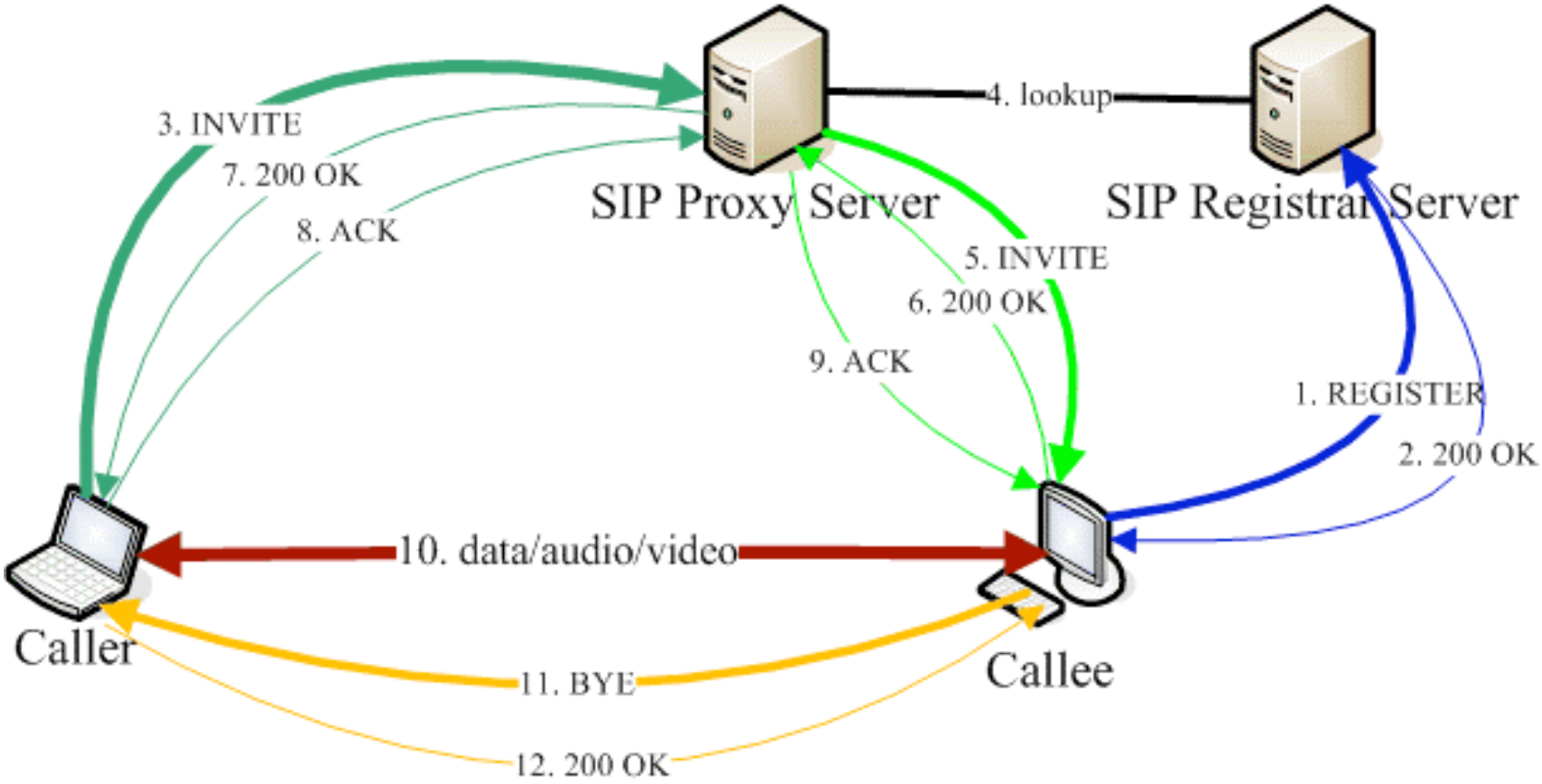
A Child of HTTP: Example Responses

```
SIP/2.0 404 Not Found
Via: SIP/2.0/TLS 10.20.30.40:5060
From: Alice <sip:alice@example.com>;tag=589304
To: Bob <sip:bob@example.com>;tag=a73kszlfl
Call-ID: 8204589102@example.com
CSeq: 1 INVITE
```

- Similar familiar HTTP responses
  - 200, 302, 401, 407, 503

NEUSTAR®

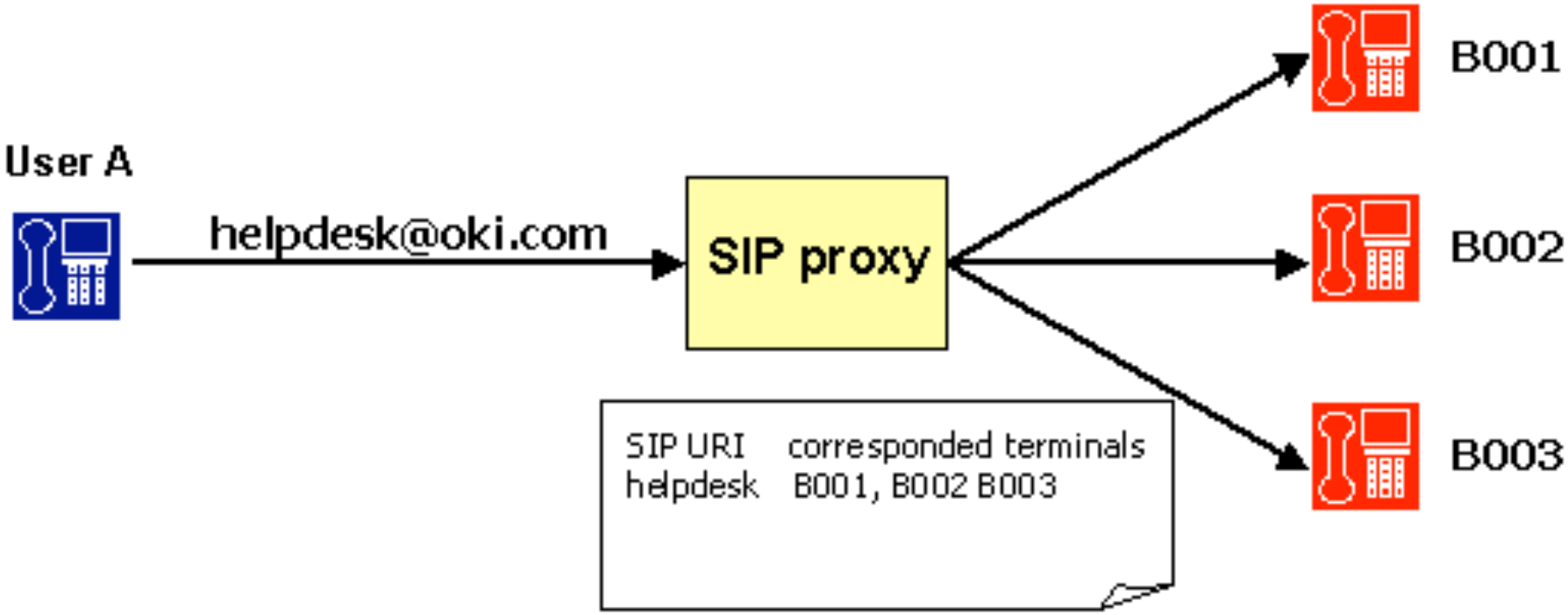
The Rendez-vous Function



http://www.javaworld.com/javaworld/jw-06-2006/images/jw-0619-sip2.gif

NEUSTAR®

Advanced Rendez-vous: Forking



http://edocs.bea.com/wlcp/wlss31/programming/wwimages/forking.gif

NEUSTAR®

Beyond Registration: Presence

- Originally emerged from the instant messaging world
  - IRC /notify, eventually the buddylist
- Presence begins with availability
  - Are you online or not? Registration
  - Exposed by the UI to potential callers
    - Know before you call if someone is there
- Also includes disposition
  - Away, busy, etc.
- Finally, capability
  - Can I reach you for voice, video, IM, etc.

- SIP supports event publication
  - PUBLISH, SUBSCRIBE, NOTIFY

NEUSTAR®

Well, that's not so bad

- The Internet was once an overlay of the telephone network
  - Inevitably, telephony operates as an application over the Internet
  - We are all witnesses to the death throes of telephony
- However, replacing something as complicated as the PSTN is quite difficult

After the Break

- Interworking
- Telephone numbers on the Internet
- NAT Traversal
- Security
- Emergency Services

- Current Events

NEUSTAR®

Interworking

- To replace the PSTN you must first connect to it
  - Global reachability is priority one
    - SIP phones must be reachable from the PSTN, and vice versa
- Interworking is achieved by gateways
  - Alignment between SIP and SS7 is close but inexact
  - Gateways must:
    - Translate signaling (as closely as possible)
    - Transcode media
- Interworking doesn't stop with the PSTN
  - Alternate VoIP protocols

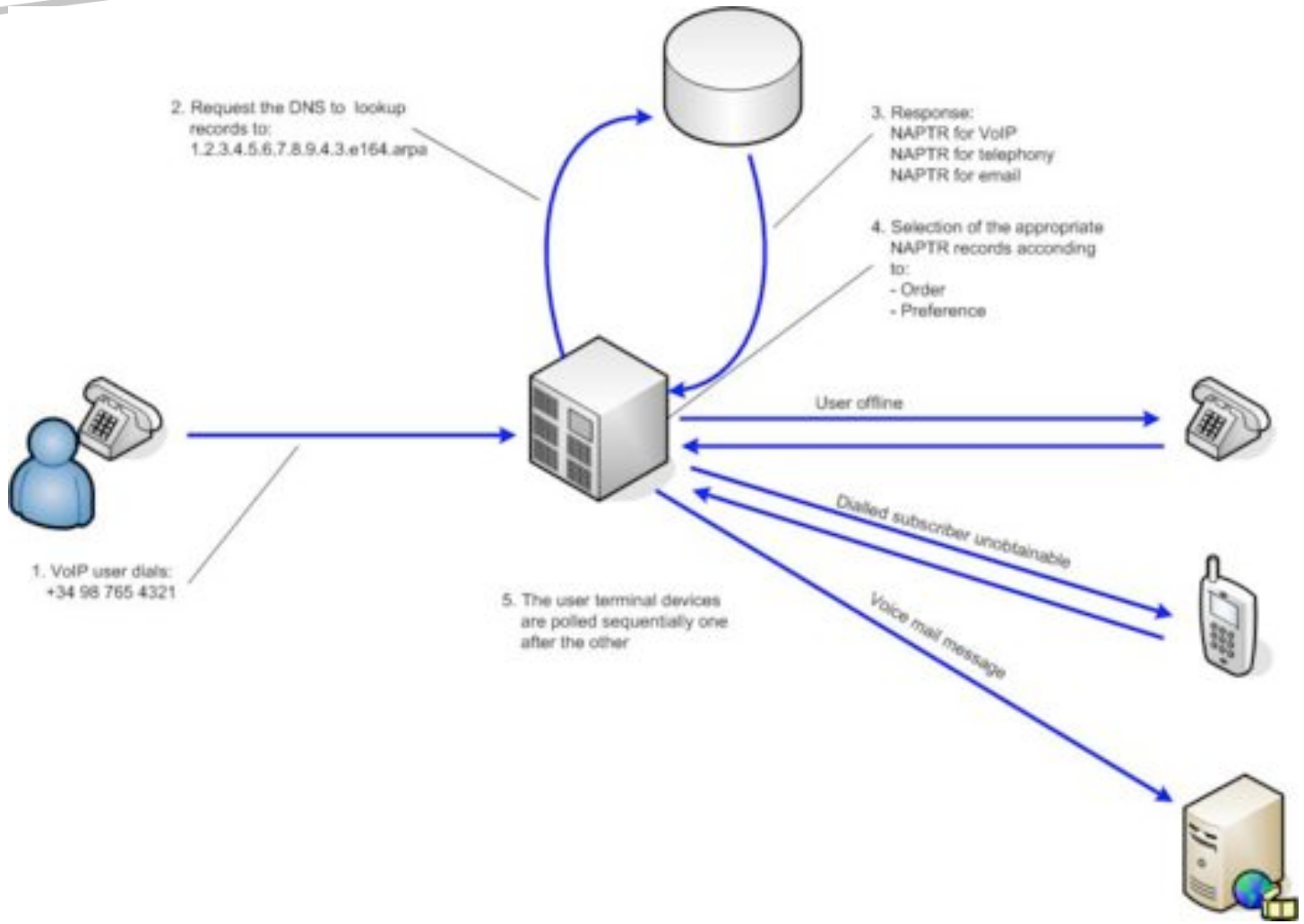
Detailed PSTN interworking

NEUSTAR®

Addressing

- Names and addresses
  - A name is a pointer
    - www.stanford.edu points to many IP addresses
- A telephone number is a name
  - Consider freephone numbers, or number portability
    - Find Me/Follow Me, voice mail
    - Translating it to an address in the PSTN is complicated
    - Even worse on the Internet
- Why not just use URIs?
  - Telephone numbers still have virtues
    - Non-linguistic and thus international
    - Necessary for interworking

ENUM: Mapping Telephone Numbers to URIs



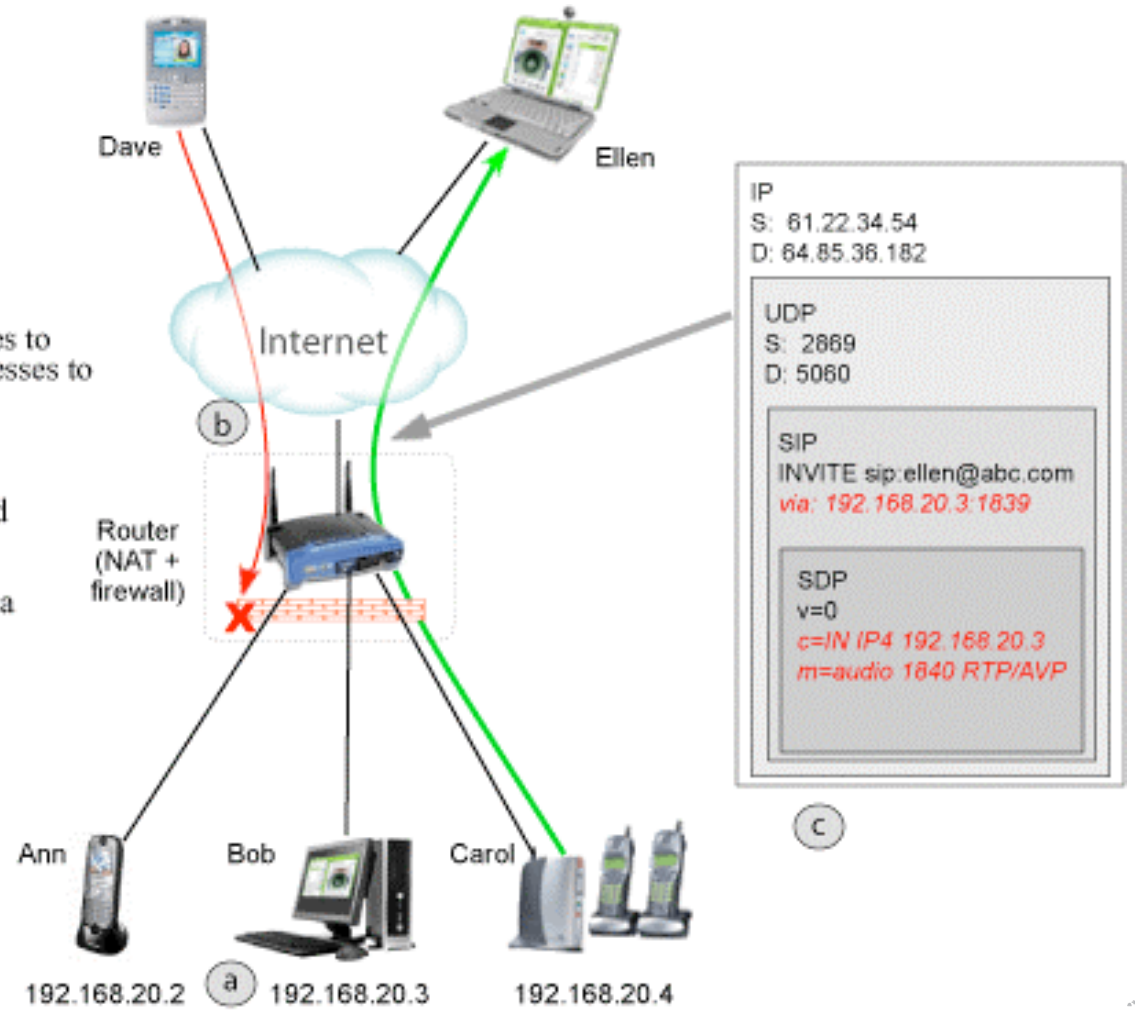
NEUSTAR®

NAT Traversal Woes

(a) NATs provide private IP addresses to local computers and map these addresses to public IP for Internet access.

(b) Firewalls protect devices from external threats by filtering uninvited IP packets. Carol can call Ellen, allowing Ellen's response to reach Carol. But user Dave cannot initiate a call to user Carol.

(c) Problem of making VoIP calls between users behind NATs



www.nattraversal.com/nattraversal2.htm

NEUSTAR®

Security

- SIP is difficult to secure
  - Proxies make end-to-end message confidentiality problematic
    - They need to read messages to route them
  - Also, proxies add and modify certain headers
    - End-to-end integrity also therefore a problem
- Caller-ID is a very desirable security service
  - In the PSTN, Caller-ID mostly works through transitive trust
    - (i.e., it doesn't work)
  - In the SIP world, user agents have credentials in their local realm
    - Much like HTTP username/password
    - Passwords, however, require pre-association
    - Telephony needs to work even when there is no pre-association
- The result: a hodgepodge
  - Digest authentication, TLS, possibly S/MIME, other situational mechanisms

Emergency Services

- What if someone dials 911?
- Initially, need to get the call to the PSTN
  - Need to reach right Public Safety Access Point (PSAP)
    - This requires coarse geographical location of the caller
  - Getting geolocation associated with Internet devices is tricky
    - VPNs and so on lead to trivial mistakes
    - Easier for mobile devices, or those with embedded GPS
  - SIP messages can now carry geolocation (RFC4119)
    - Mechanisms for mapping geolocation to PSAPs exist
- Future requirement to access IP-enabled PSAPs

Current Events

Peer-to-Peer File-sharing

&

Real-time Communications

NEUSTAR®

The Scene: A Typical Broadband ISP

- Operates access networks
  - May run metropolitan networks as well
  - All feeds into backbone
- Like all ISPs, network is optimized for download-balanced traffic
  - Web traffic, primarily
- Like all ISPs, uses some form of statistical multiplexing
  - Uplink capacity is less than aggregate bandwidth available to users
    - Exact location of chokepoint may vary
  - In the envisioned usage model, users actively instigate bursty traffic and sleep sometimes
- Internet traffic growing by >30% a year
  - Adding new capacity is habitual
- With respect to its customers, it is probably positioned as a monopoly broadband provider
  - At worst, half of a duopoly

Enter P2P File-sharing

- File-sharing virally spreads to a community of interest
  - May go from obscure to ubiquitous practically overnight
  - >30 million installs, conservatively
- Elastic bulk file transfer protocol
  - *Elastic* meaning that the application still works if the bit rate is fast or slow, if latency is high or low
  - Files may be segmented, but still bulk
  - Finds the best transfer rate for users by probing multiple potential sources for a file or segment
- As a content distribution network, also attracts corporate interest and sponsorship
  - Why pay for content distribution when network customer systems can provide the same function for free?

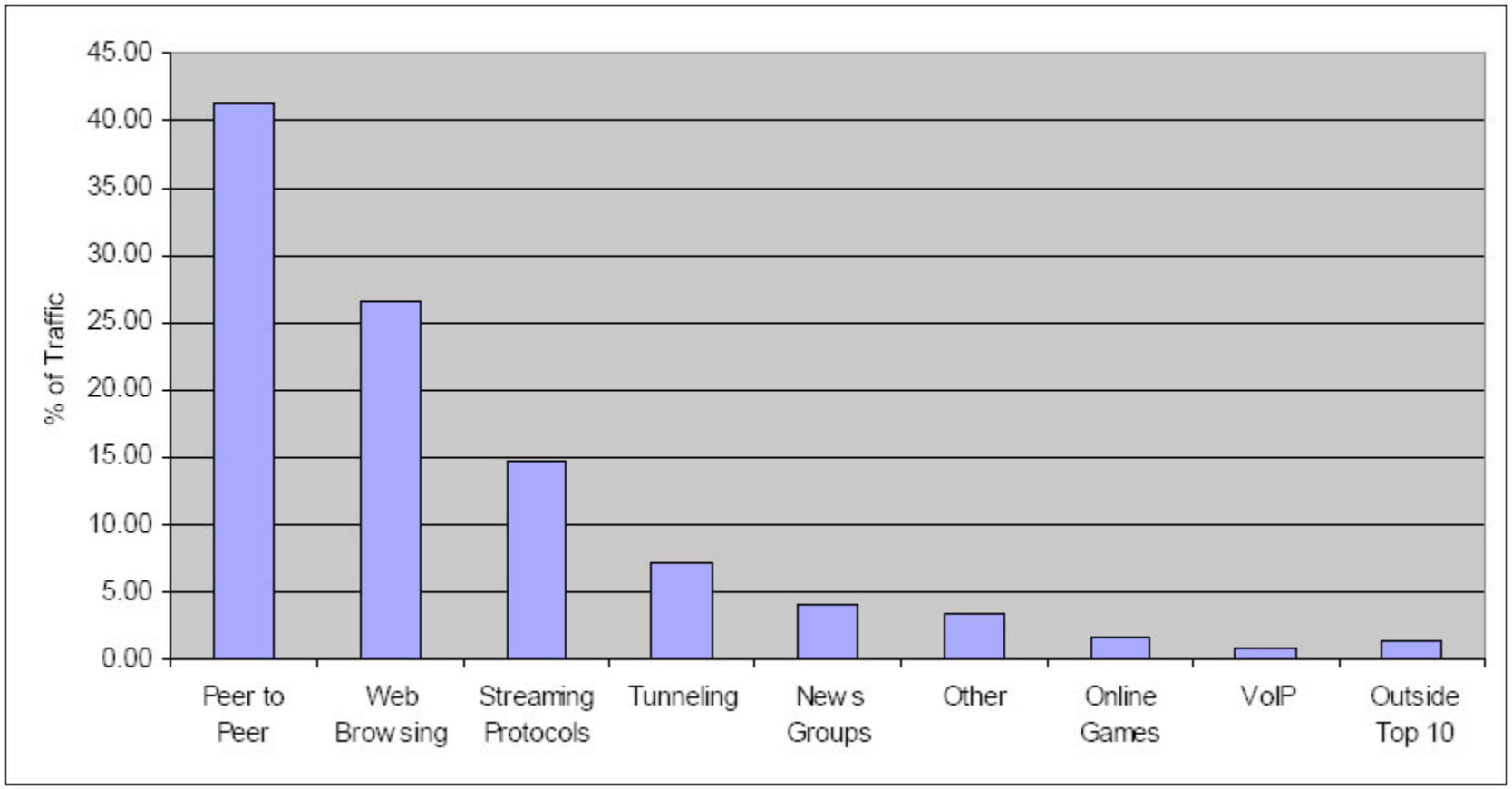
Trouble begins

- The ISP receives customer complaints
  - Network is sluggish
- File-sharing traffic does not behave as ISP expects
  - Defies traditional uplink/downlink ratio of web traffic
    - Uplink bandwidth consumption equal to downlink, if not greater
  - Always active, “unattended consequences”
    - Usage unchanged across time of day
  - Statistical multiplexing assumptions at provisioning time turn out to be false
    - Chokepoints thus emerge at various locations in the network
  - Opens multiple simultaneous TCP connections
    - May confound traditional transport-layer management of fairness

Estimates of P2P breadth vary widely, between 35% to 85% of Internet bandwidth

NEUSTAR®

Network Utilization: Averaged Up/Downlinks



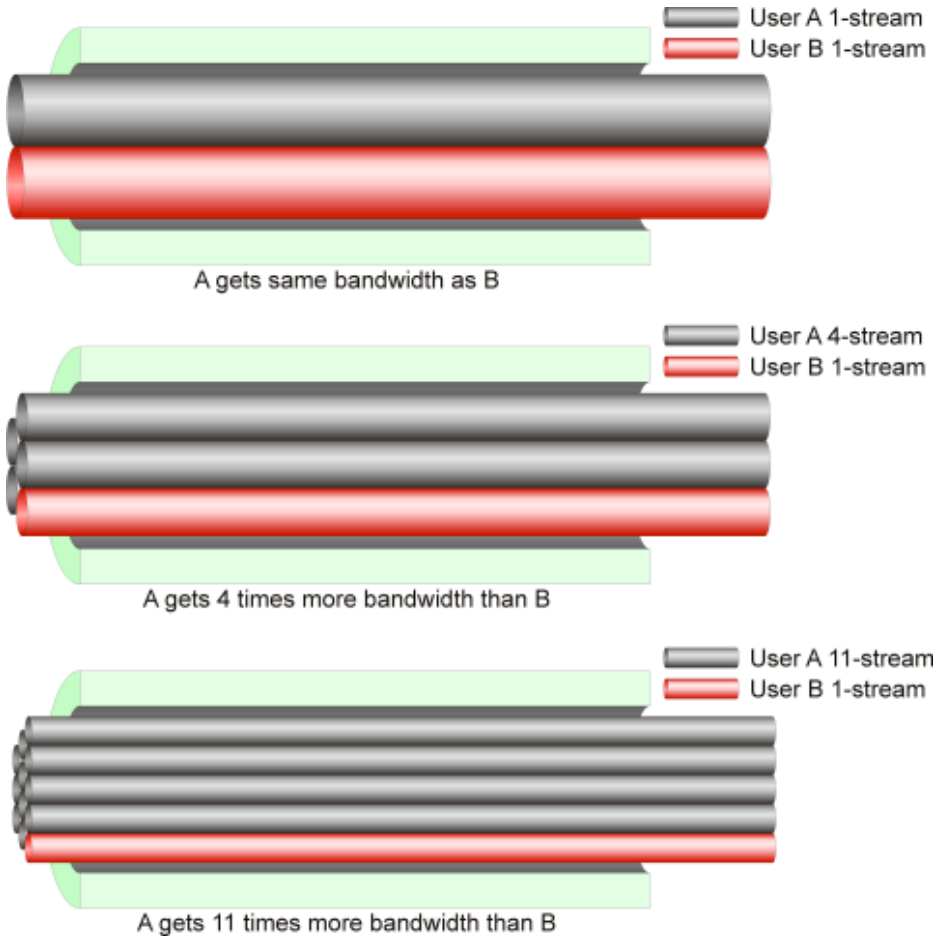
Courtesy of SandVine

NEUSTAR®

The Result: Dropped packets

- Congestion levels approach bandwidth limits
  - Especially on the uplinks
- Routers begin to discard packets
  - Tail drop queuing discipline
    - Problems possibly abetted by large buffers on residential devices
  - Depending on where the chokepoints are, Active Queue Management (AQM) is unlikely
- The greediest application wins
  - Bulk traffic wins over trickles of constant traffic
- This quickly starts to affect applications visible to end users
  - Especially real-time communications

The TCP Congestion Problem



http://blogs.zdnet.com/Ou/images/tcp-2.png

NEUSTAR®

Who complains to the ISP?

- Not bulk file transfer users
  - If it takes 15 minutes instead of 10 minutes, the application still works
- Complaints come from users of inelastic applications
  - Applications that do not work in the face of loss
  - Extremely delay-intolerant, need only the most recent data, retransmission does not help
- Instantaneous personal communications, like
  - Voice-over-IP (VoIP)
  - Instant Messaging
    - Downlink “streaming” is less of a concern
- Telnet or archaic terminal shells (geeks only)
- Even web browsing
  - Highly interactive, networks optimized to serve pages quickly

- What should the ISP do?

The P2PI Workshop

- May 28th at MIT
- Speakers included Comcast and BitTorrent
- Attempting to identify engineering efforts that
  - Add value incrementally
  - Are complimentary
  - Are achievable
- You are finding out about it a bit late
  - All one can do in a day is sketch future work
  - Numerous ongoing IETF efforts are planned from this

(p2pi@ietf.org)

NeuStar®