The phone is ringing - now what?

Henning Schulzrinne
Dept. of Computer Science
Columbia University
New York, New York
(sip:)schulzrinne@cs.columbia.edu

Computer and Communications Societies Joint Chapter – IEEE NJ Coast Section

March 2, 2001

Overview

- A brief history
- VoIP service models
- The IETF VoIP architecture
- Peer-to-peer and master-slave architectures
- The Session Initiation Protocol (SIP)
- What makes VoIP difficult?
- Columbia CS prototype and trial
- The dangers of VoIP
- Instant messaging & presence ➠ generic event service
- Killer application ➠ programmable services

Historical perspective

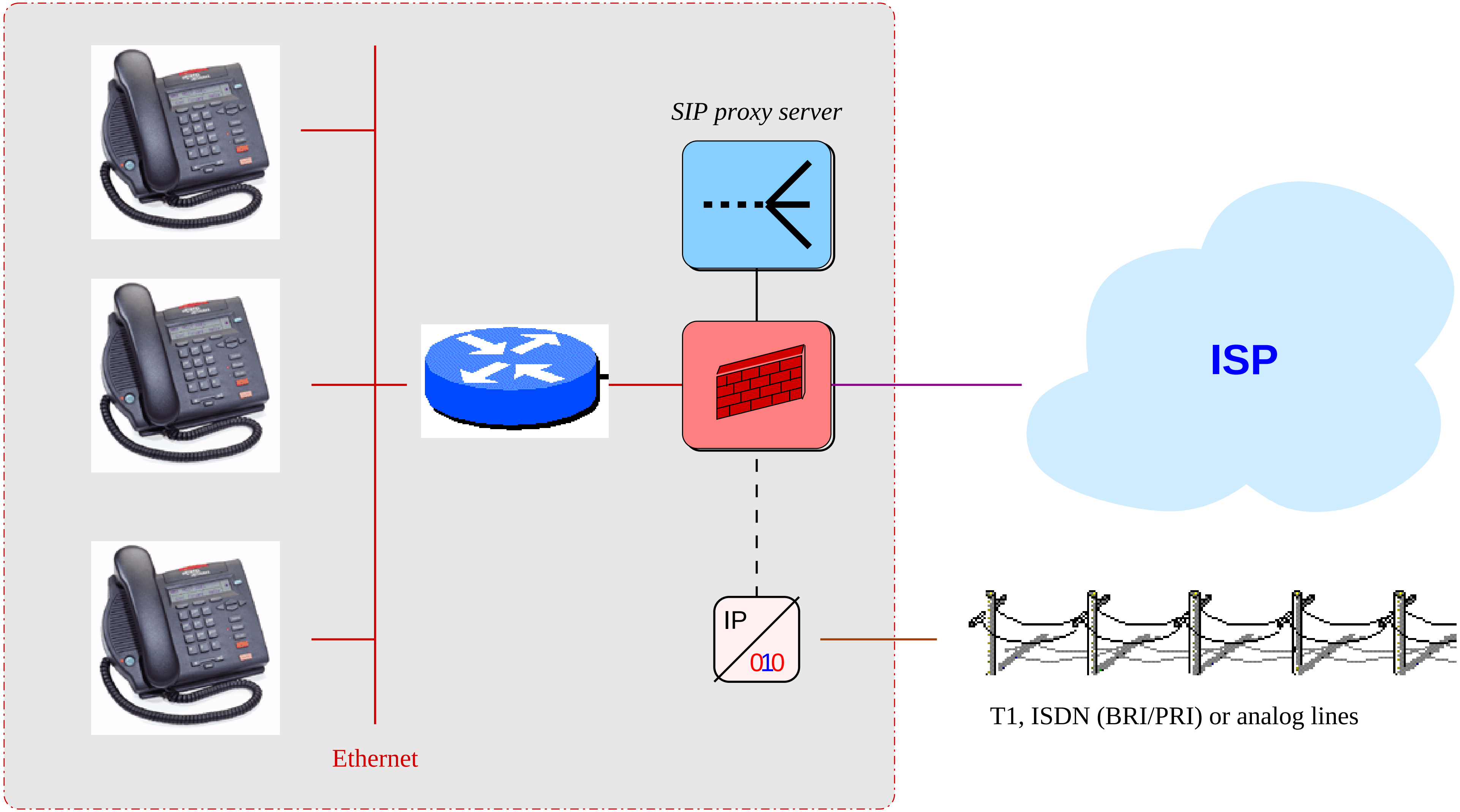
1876	invention of telephone
1915	first transcontinental telephone (NY–SF)
1920's	first automatic switches
1956	TAT-1 transatlantic cable (35 lines)
1962	digital transmission (T1)
1965	1ESS analog switch
1977	4ESS digital switch
1980s	Signaling System #7 (out-of-band)

Internet Telephony Service Models

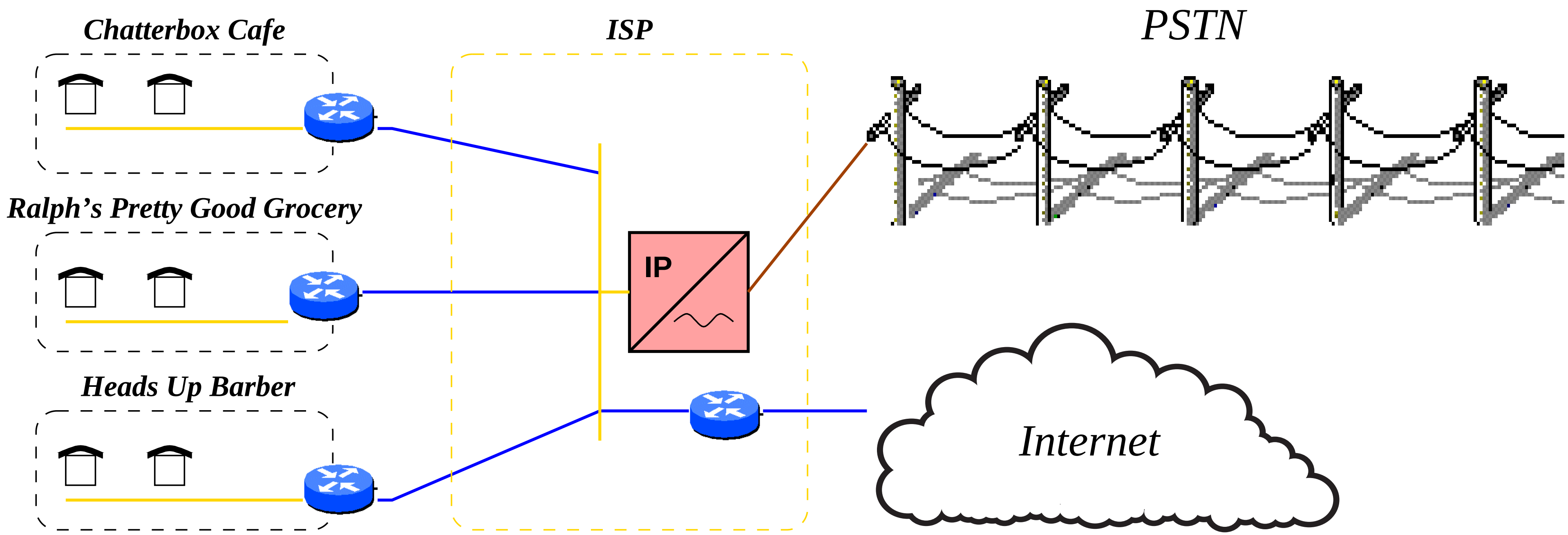
- Internet “PBX”
- Internet Centrex
- Internet Carrier

➠ same basic equipment, but size of gateway varies

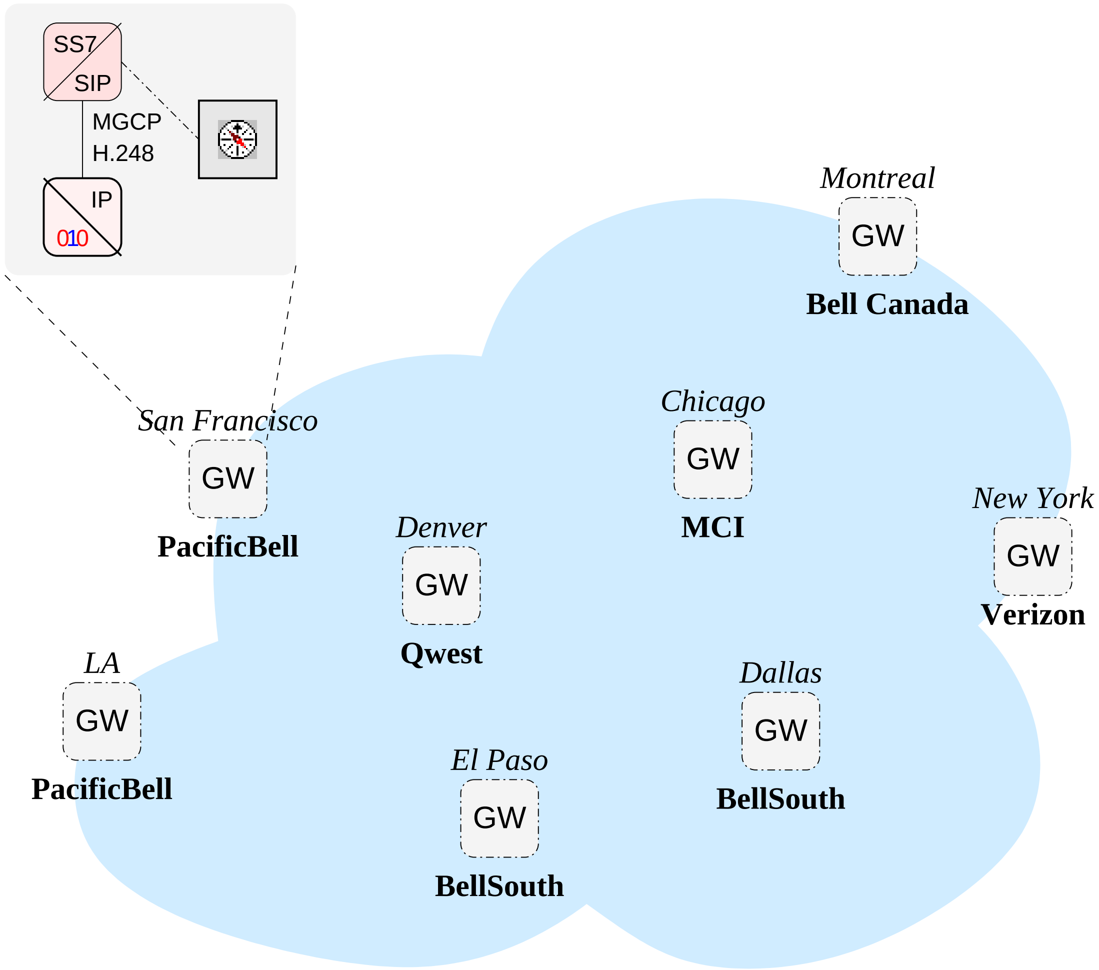
Internet PBX



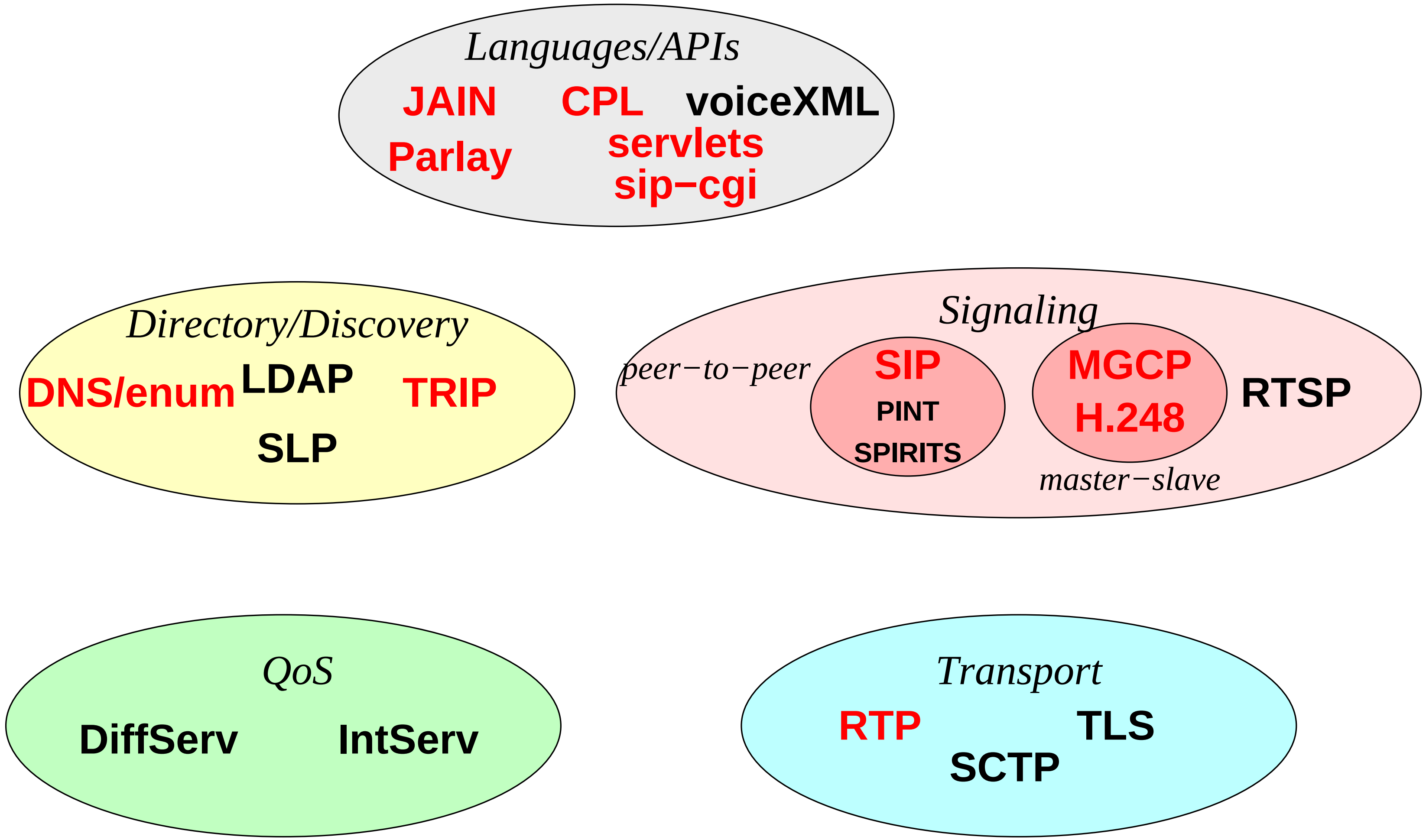
IP Centrex


Chatterbox Cafe
ISP
PSTN
Ralph's Pretty Good Grocery
IP
Heads Up Barber
Internet

IP Carrier


SS7
SIP
MGCP
H.248
IP
010
Montreal
GW
Bell Canada
San Francisco
GW
PacificBell
Chicago
GW
MCI
New York
GW
Verizon
Denver
GW
Qwest
LA
GW
PacificBell
Dallas
GW
BellSouth
El Paso
GW
BellSouth

IETF VoIP Protocol Architecture



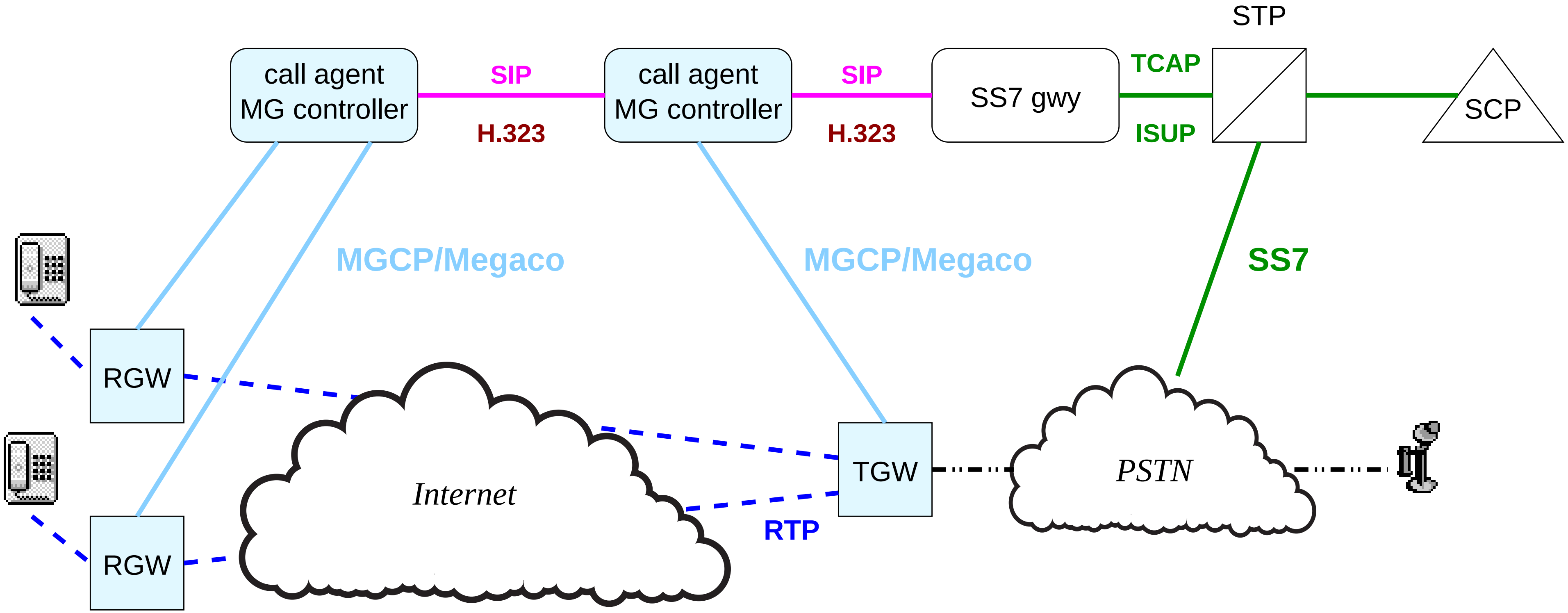
IETF VoIP Architecture Characteristics

- universal identifier *user@domain*: SIP URL = email = NAI
- separation of transport of services
- media-neutral, including beyond audio and video
- emphasis on user-programmable services
- web integration: content, mutual referral
- integration with IM and presence

Master-Slave Architecture

- master-slave: MGC controls one or more gateways
- allows splitting of signaling and media functionality
- "please send audio from circuit 42 to 10.1.2.3"
- uses MGCP (implemented) or Megaco/H.248 (standardized, but just beginning to be implemented)
- gateway can be residential
- basis of PacketCable NCS (network control system) architecture
- service creation similar to digital PBX or switch
- $\longrightarrow$ can charge for caller id, call waiting

MGCP Architecture



- for all but small system, need peer-to-peer!
- MGCP system can call SIP or H.323 end system
- all use RTP to transfer data

Peer-to-Peer Architecture

- “IP telephones”, gateways, PCs with software = IP hosts
- *may* use servers (H.323 gatekeepers, SIP proxy servers)
- end system fully state-aware
- protocols for call setup: H.323 or SIP
- more flexible user interface

SIP Overview

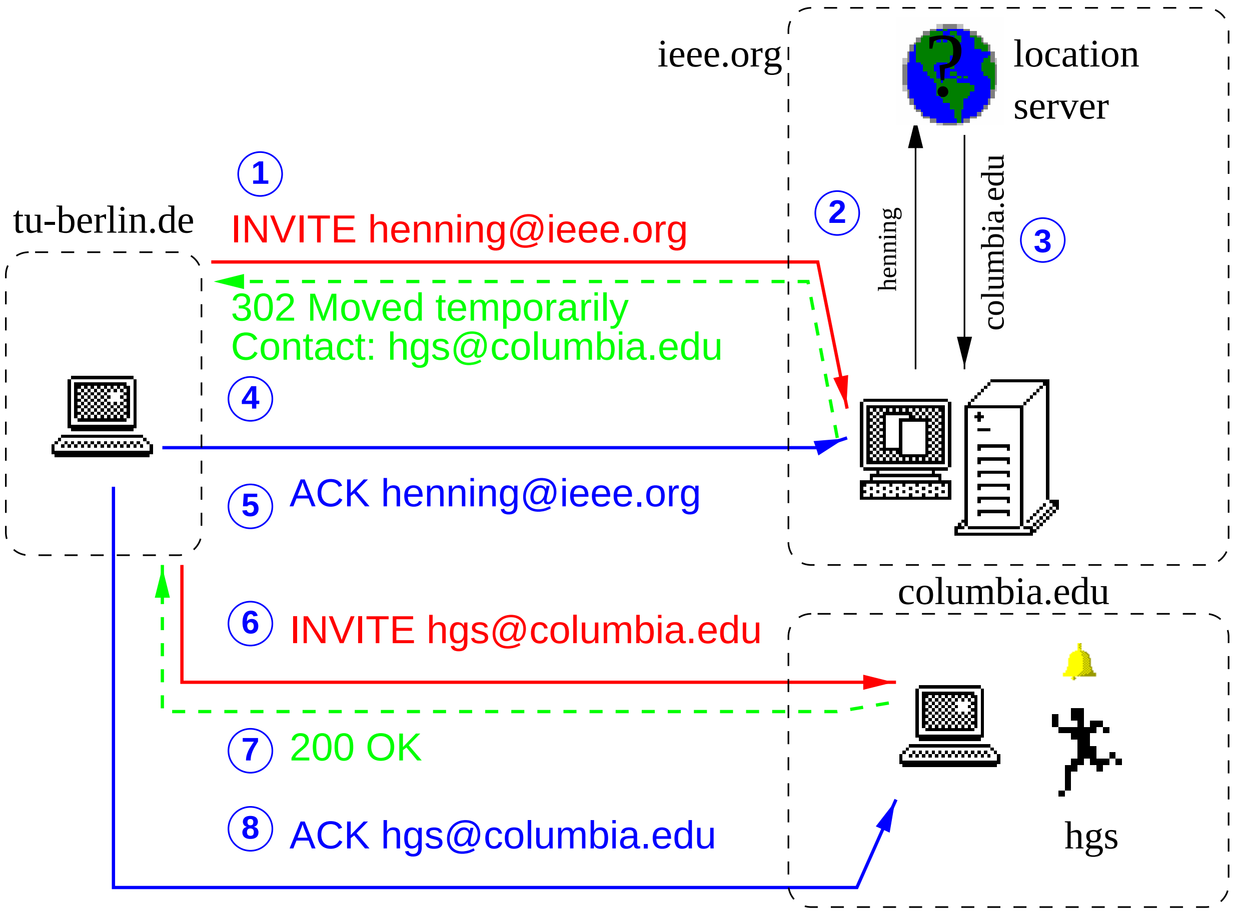
- protocol for establishing, modifying, tearing down (multimedia) sessions
- IETF Proposed Standard since March 1999
- multimedia = audio, video, shared applications, text, ...
- also used for “click-to-dial” (PINT wg) and possibly Internet call waiting (SPIRITS wg)
- to be used for PacketCable Distributed Call Signaling
- to be used for Third-Generation Wireless (3GPP, 3GPP2)

SIP Components

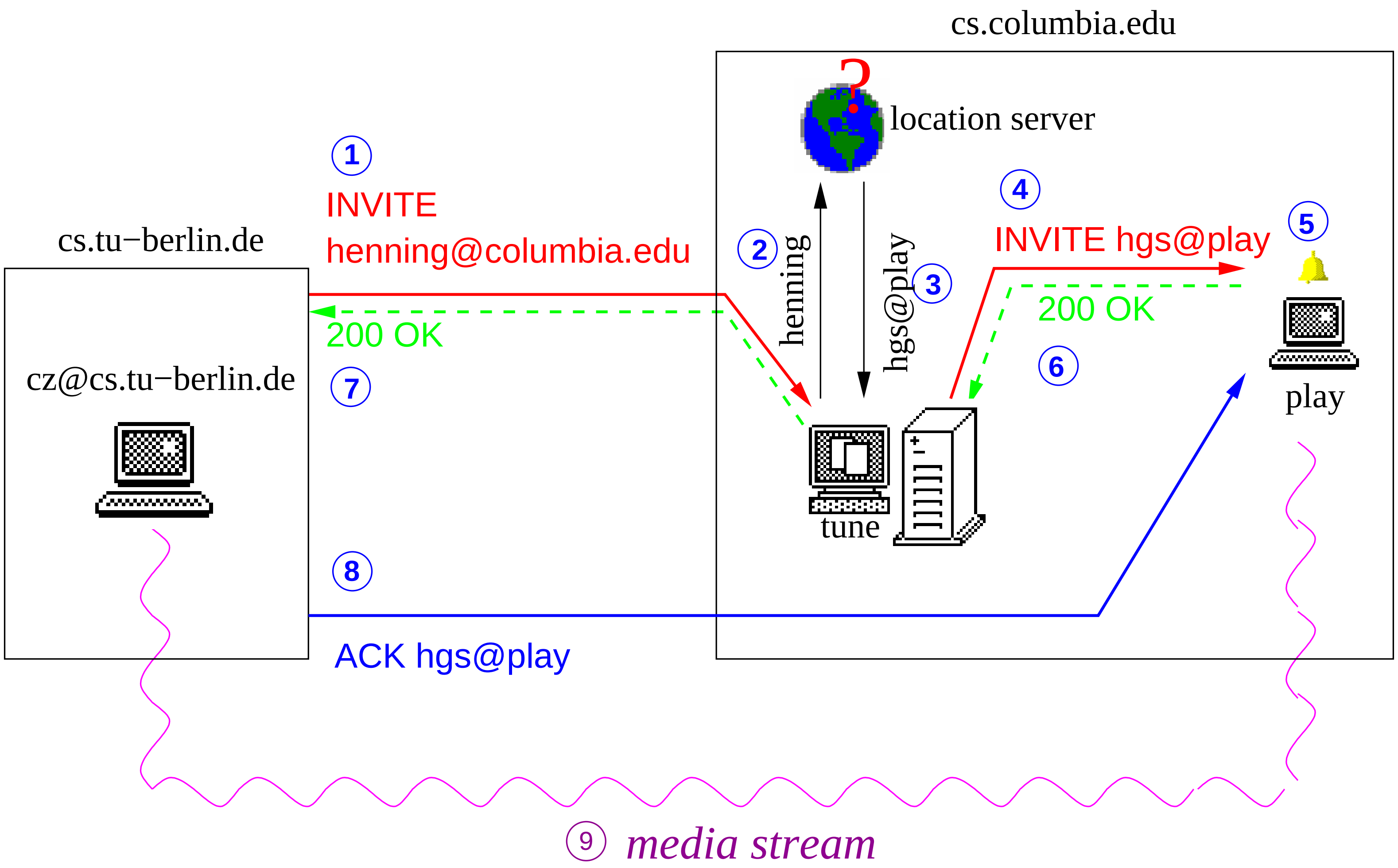
entity	does	examples
proxy server	forward calls	firewall controller, “call router”
redirect server		“application server”
user agent	end system	SIP phone, gateway, “softswitch”
registrar	location mgt.	mobility support

Roles are changeable, on a request-by-request basis

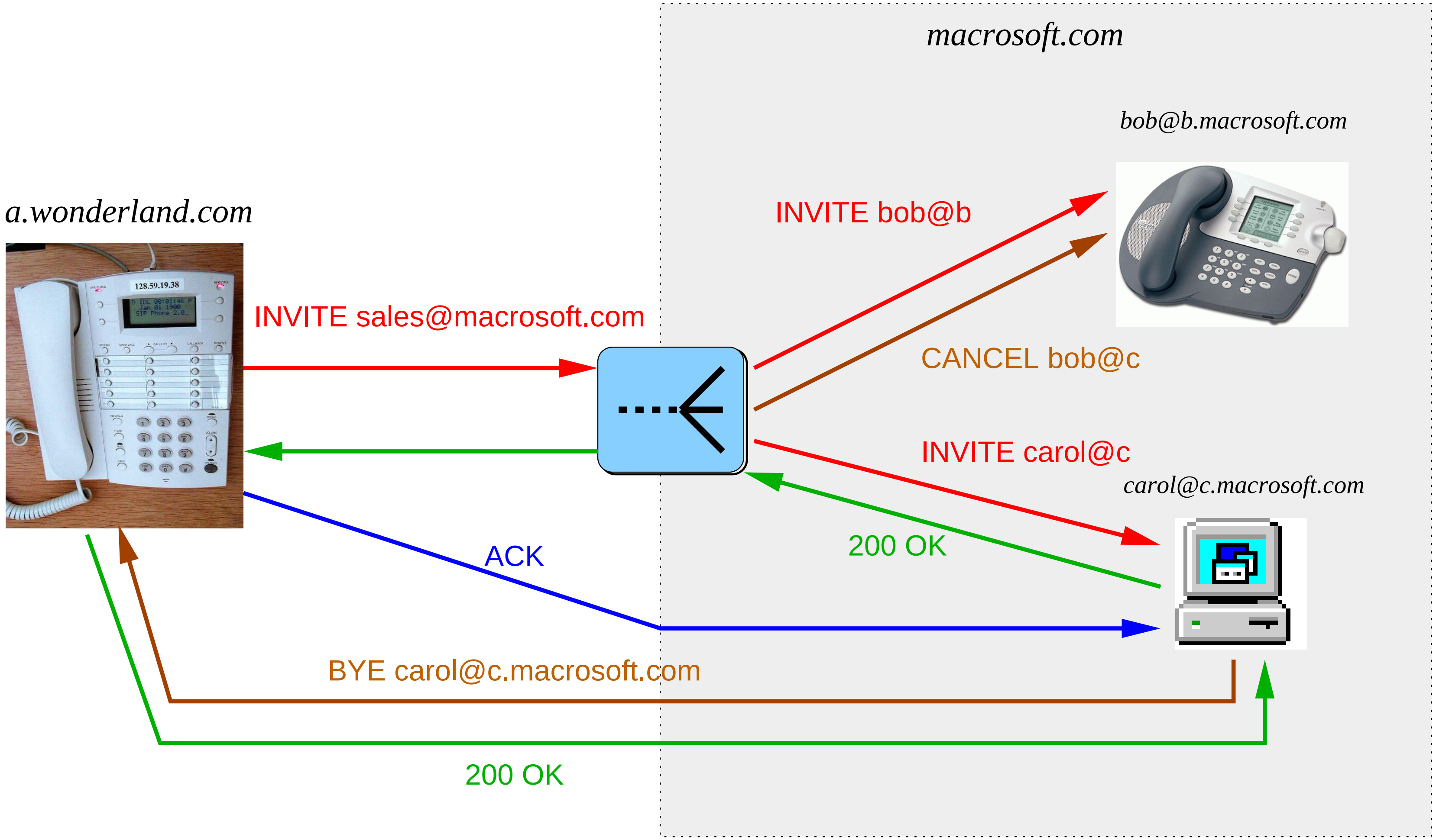
SIP Example: Redirection



SIP Example: Proxying



SIP Forking Proxies



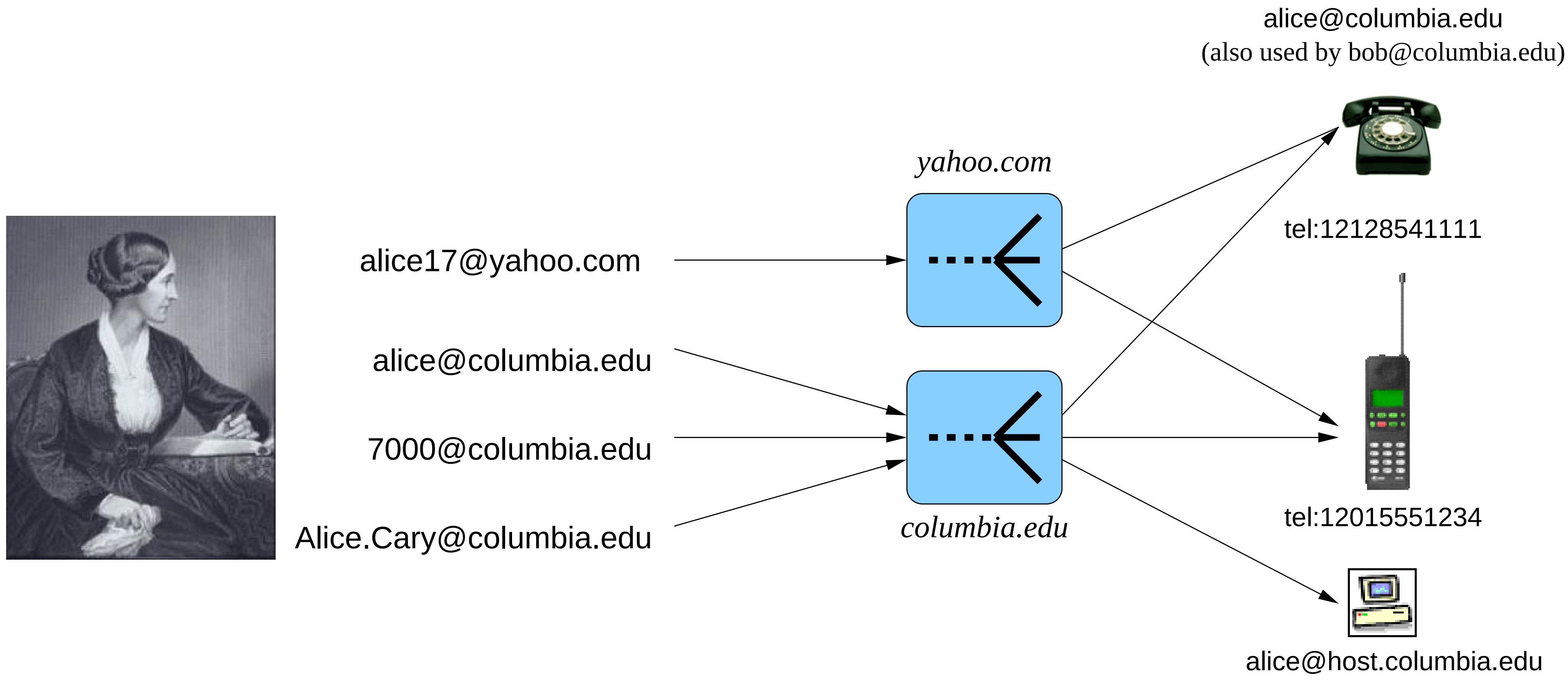
SIP Advanced Features

- forking
- extensibility: new headers, methods, bodies
- security: web-like, PPP/CHAP or PGP
- multicast-capable
- support for personal, session, terminal, service mobility
- caller preferences: direct calls based on properties

SIP Mobility

terminal	cross-provider	REGISTER, re-INVITE
personal	different terminals, same address	REGISTER
service	different terminals, same services	upload
session	move sessions across terminals	REFER

SIP Personal Mobility



SIP years

Year	development	trade rags
1996-1998	R&D	“academic exercise’, “distraction from H.323”
1999	standard & skunk works	“what does SIP stand for again?”
2000	product development	“SIP cures common cold!”
2001	pioneer deployment	“Where are the SIP URLs?”
2002	kmart.com/sip	SIP product comparisons

SIP Status Early 2001

- almost all telecom equipment vendors working on SIP products
- first general-availability SIP hardware (Ethernet phones, small gateways), but limited
- number of SIP proxy servers in customer trials
- ready for field trials and early-adopter “PBX-free” enterprises
- but can’t buy couple of SIP phones from web page

So I Want to Build a SIP Network...

Ready for trials, but probably not quite for shrink-wrap status:

- installation and operation still requires fair amount of expertise
- lots of web and email experts, few SIP experts
- needs some external infrastructure: DHCP and SRV, possibly AAA
- inconsistent configuration for Ethernet phones (being worked on)
- SIP phones still more expensive than analog phones ➠ hard to justify PBX replacement (incremental cost)
- no just-download or ship-with-OS “soft” clients

Why aren't we junking switches right now?

What made other services successful?

email: available within self-contained community (CS, EE)

web: initially used for local information

IM: instantly available for all of AOL

All of these ...

- work with bare-bones connectivity ($\geq$ 14.4 kb/s)
- had few problems with firewalls and NATs
- don't require a reliable network

Why aren’t we junking switches right now?

Telephone services are different:

- reliability expectation 99.9% $\nearrow$ 99.999%
- PC not well suited for making/receiving calls – most residential handsets are cordless or mobile
- business sets: price incentive minor for non-800 businesses
- services, multimedia limited by PSTN interconnection
- initial incentive of access charge bypass fading (0.5c/min.)
- international calls only outside Western Europe and U.S.

Prognosis

- much less cable telephony than predicted, mostly boring GR303
- greenfield PBX installations for net-savvy enterprises
- enhancements for maxed-out PBXs – need PBX Ethernet interfaces
- tie-line replacements for branch offices
- backbones for some carriers
- maybe DSL and cable modems, but lifeline? replacement of cordless phones?
- 3G deployment, assuming any 3G companies not bankrupted by license fees

Prognosis, cont'd.

- BICC & ISUP-carriage for legacy-burdened carriers
- H.323 for conferencing (until Microsoft ships Windows SIP client . . .)
- need T.120-equivalent for cross-platform screen sharing, e.g., VNC

Standardization

- interaction with resource reservation
- caller preferences (“no mobile phones, please”)
- interoperation with ISUP (“SIP-T”)
- call transfer and third-party control
- conferencing: central server, end system, full mesh
- server benchmarking and scaling
- requirements for deaf users
- call processing language: coordination with iCal

Status of SIP working group items

reliable provisional	IESG review
caller preferences	WG last call done
call flows	ready for last call
SIP guidelines	WG last call done
ISUP over MIME	ready for IESG
SIP MIB	needs update
server feature ann.	revisions based on IESG feedback
service examples	needs work
session timer	ready for WG last call
call transfer	in revision
state maintenance	ready for last call
DHCP	IESG revisions done

SIP Bake-Off

- takes place every four months, 7th at ETSI March 2001
- 45 organizations from 11 countries
- about 50-60 implementations:
  - IP telephones and PC apps
  - proxy, redirect, registrar servers
  - conference bridges
  - unified messaging
  - protocol analyzers
- first IM/presence interop test
- emphasis on advanced services (multi-stage proxying, tel URLs, call transfer, IVR, . . .)

PSTN legacies to avoid

- E.164 numbers – might as well wear bar codes
- tones and announcements
- in-band signaling for features (DTMF)
- systems with user-interface knowledge (12 keys, voice)
- voice-only orientation (BICC, MGCP/Megaco)
- integration of bit transport and services
- service-specific billing ➠ separate signaling & billing
- trusted networks without crypto

➠ confine PSTN knowledge to edge of network

The Dangers of VoIP

- focus on single service: voice, fax, ...
- PSTN: service orientation $\longleftrightarrow$ Internet: neutral transport
- APIs as least common denominator across POTS, ISDN, SS7 $\longrightarrow$ 100-year old functionality
- carbon-copy replication of existing services
- terminology overload

Replication of Existing Services

- “user is familiar with PSTN services”
- but how many users actually know how to use call transfer or directed pick-up?
- user interface is often just legacy of key systems or other ancient technology
- avoid binding of identifiers to devices – call person or group of people, regardless of location
- instead, model desired behavior
- single-server features don’t need standardization
- find general mechanisms (e.g., REFER for three-party calls and various call transfers)

Terminology Overload

Invasion of the meaningless technical-sounding terms, attempting to familiar mimic PSTN boxes:

- CO switch $\longrightarrow$ soft switches = gateway + SIP UA + ?
- SCP $\longrightarrow$ application servers = proxy? web server? media server?
- PBX $\longrightarrow$ Internet PBX = proxy? + gateway?
- ...

Temptation: new name $\longrightarrow$ new protocols, APIs, ... – the old box boundaries don't necessarily make sense!

It’s That Simple. . .

We really only have a few basic components:

- PSTN gateway, with some combination of FXO/FXS
- SIP proxy/redirect/registrar servers (or H.323 gatekeepers)
- SIP user agents (or H.323 terminals): PCs, phones
- media storage servers
- DNS, directory, web, email, news, . . . servers

Invisible Internet telephony

“VoIP” technology will appear in

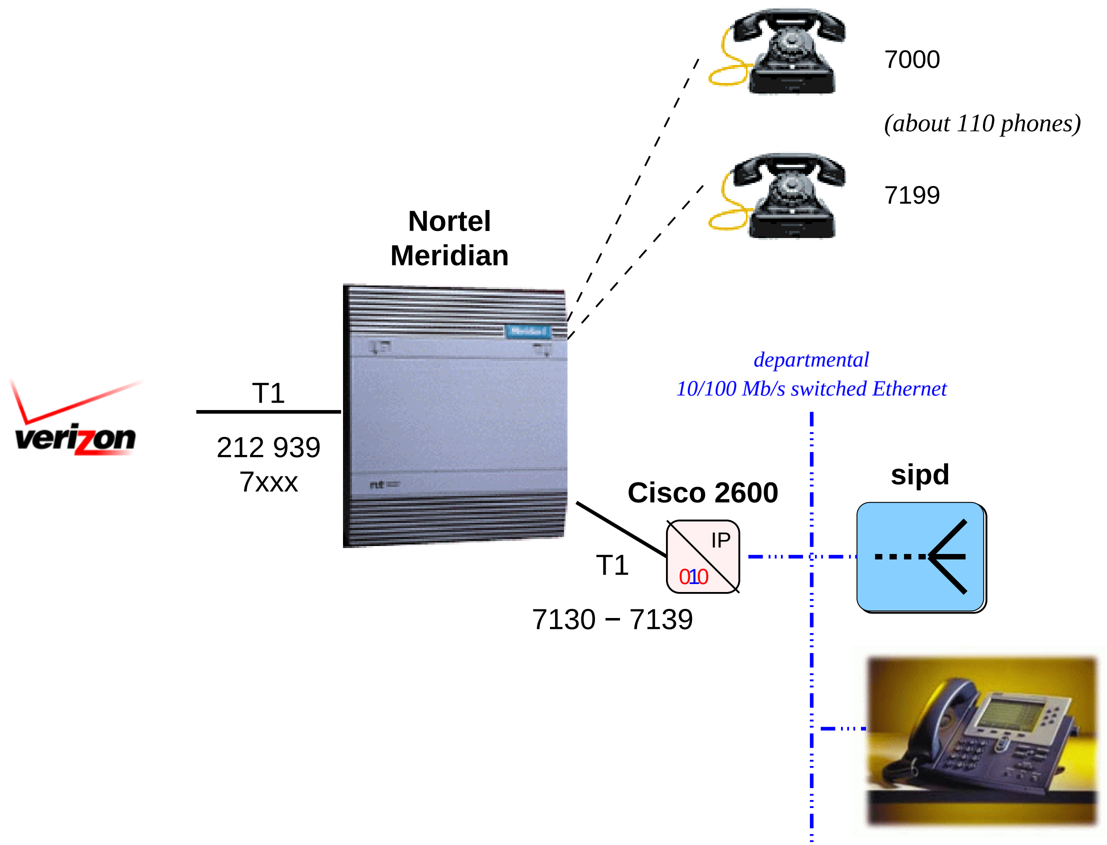
- Internet appliances
- home security cameras, web cams
- 3G mobile terminals
- fire alarms and building sensors
- chat/IM tools
- interactive multiplayer games
- 3D worlds: proximity triggers call

Internet Telephony – as Part of Internet

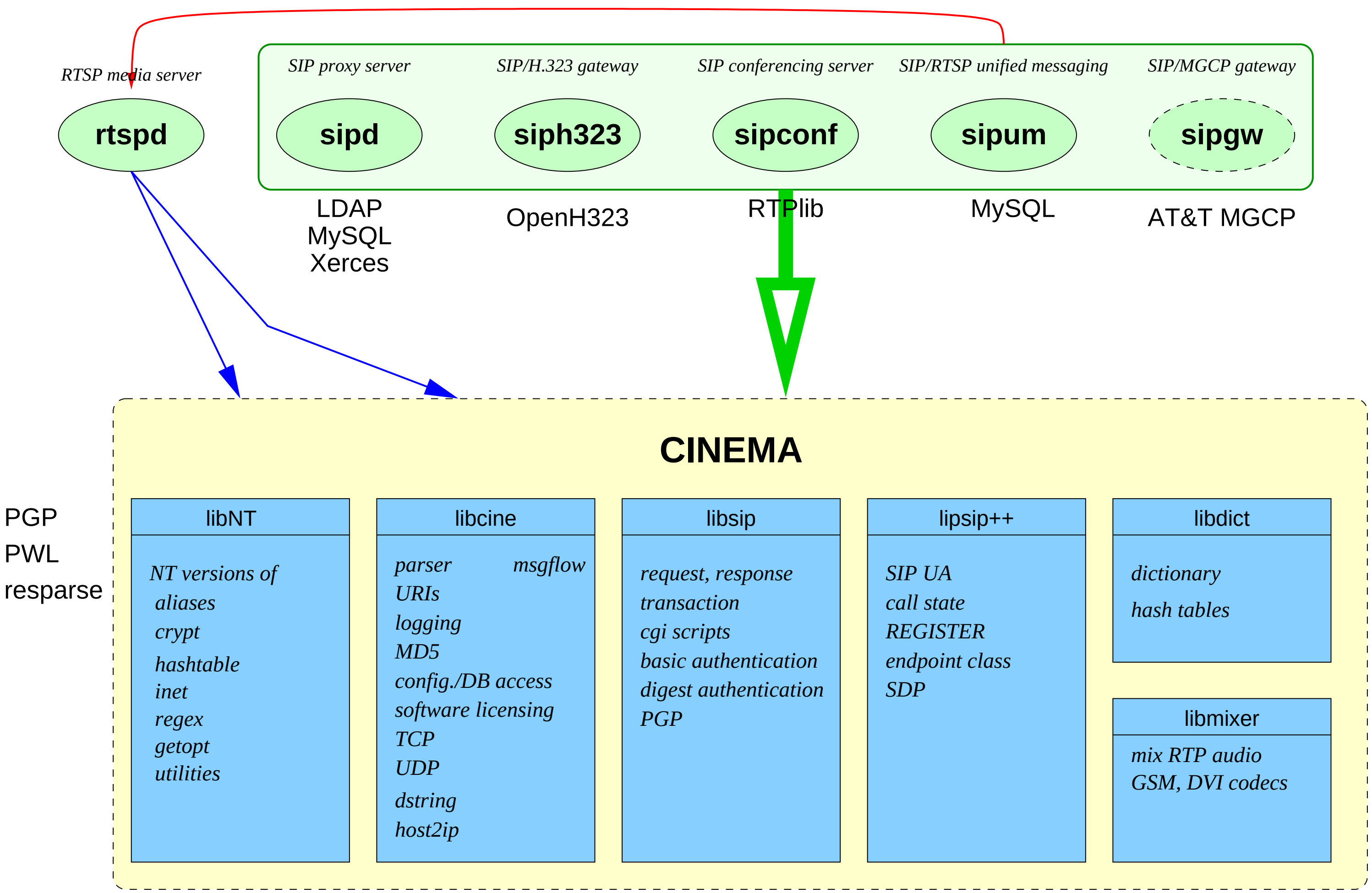
- universal identifier: email address = SIP address = IM address
- SIP URLs in web pages
- forward to email, web page, chat session, . . .
- include web page in invitation response (“web IVR”)
- third-party control of calls via scripts,
- include vCard, photo URL in invitation
- user-programmable services: CGI (RFC 3050), CPL, servlets

Example: Columbia CS phone system

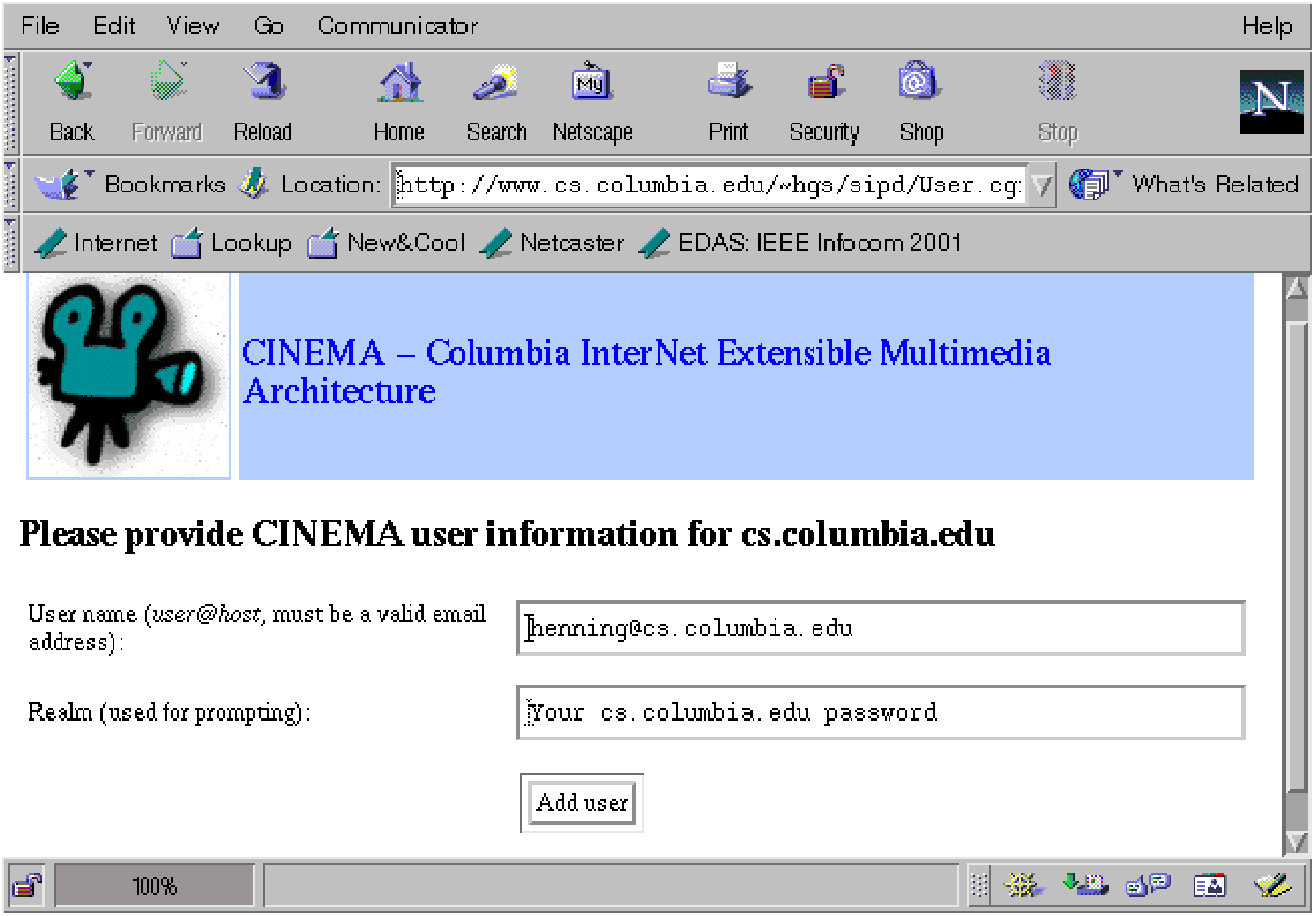
Expand existing PBX via IP phones, with transparent connectivity



CINEMA


RTSP media server
rtspd
SIP proxy server
sipd
SIP/H.323 gateway
siph323
SIP conferencing server
sipconf
SIP/RTSP unified messaging
sipum
SIP/MGCP gateway
sipgw
LDAP
MySQL
Xerces
OpenH323
RTPlib
MySQL
AT&T MGCP
PGP
PWL
resparse
CINEMA
libNT
NT versions of
aliases
crypt
hashtable
inet
regex
getopt
utilities
libcine
parser
msgflow
URIs
logging
MD5
config./DB access
software licensing
TCP
UDP
dstring
host2ip
libsip
request, response
transaction
cgi scripts
basic authentication
digest authentication
PGP
lipsip++
SIP UA
call state
REGISTER
endpoint class
SDP
libdict
dictionary
hash tables
libmixer
mix RTP audio
GSM, DVI codecs

sipd single sign-on for account creation and modification



sipd contact management

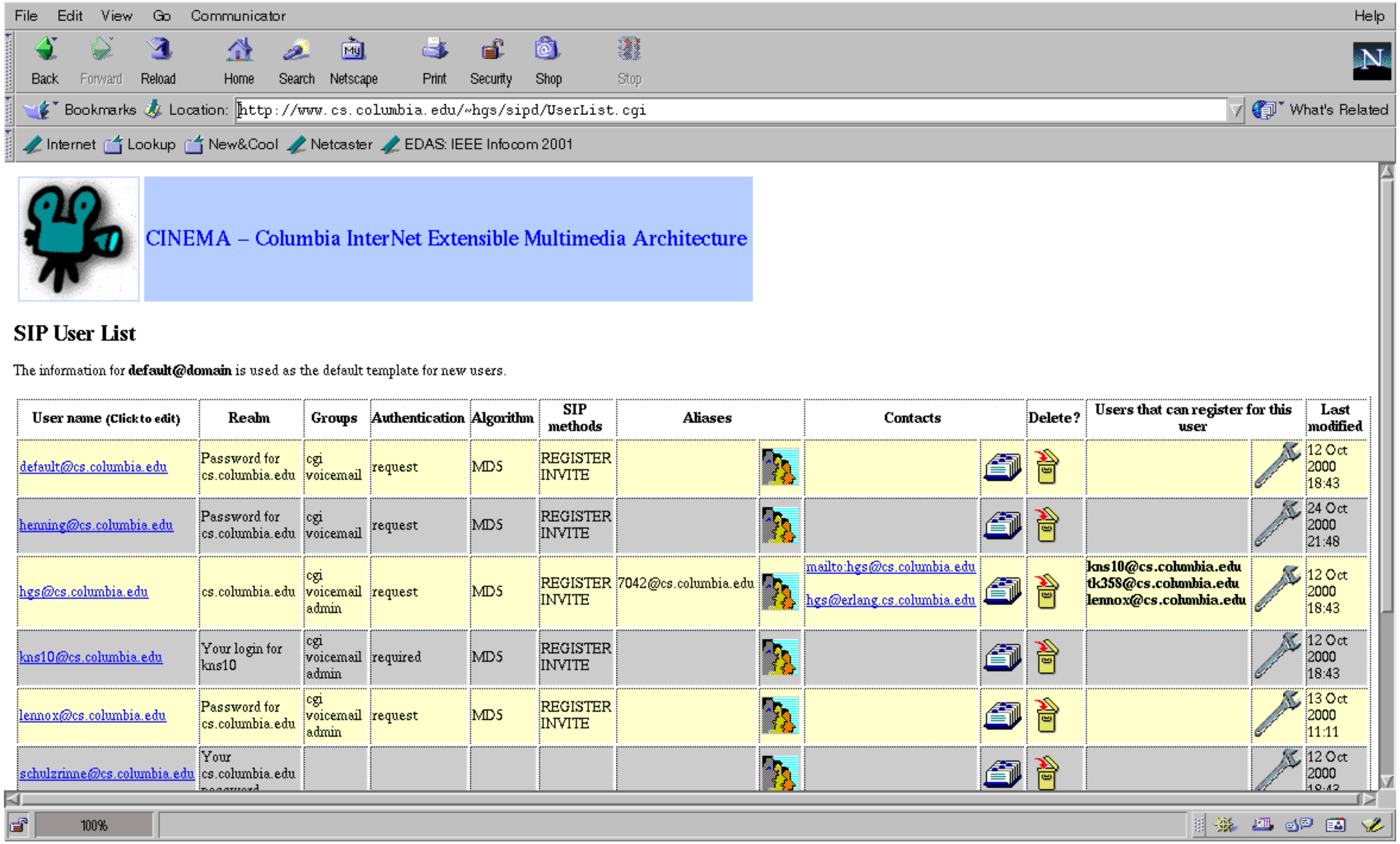
CINEMA – Columbia InterNet Extensible Multimedia Architecture

Contacts for User hgs@cs.columbia.edu

Deleted contact **hgs@erlang.cs.columbia.edu**.

Contact		Preference	Expires	Action	Last modified	
<mailto:hgs@cs.columbia.edu>		0.1	01 Dec 2001 00:00	Proxy	12 Oct 2000 18:41	Change
<sip:hgs@muni.cs.columbia.edu>				Proxy	15 Dec 2000 18:06	Change
<sip:hgs@128.59.19.205>		1.0	05 Jan 2001 00:49	Proxy	04 Jan 2001 18:49	Change
<sip:hgs@128.59.19.216:5060>		1.0	05 Jan 2001 00:53	Proxy	04 Jan 2001 18:53	Change

sipd user configuration



SIP User List

The information for **default@domain** is used as the default template for new users.

User name (Click to edit)	Realm	Groups	Authentication	Algorithm	SIP methods	Aliases	Contacts	Delete?	Users that can register for this user	Last modified
default@cs.columbia.edu	Password for cs.columbia.edu	cgi voicemail	request	MD5	REGISTER INVITE					12 Oct 2000 18:43
henning@cs.columbia.edu	Password for cs.columbia.edu	cgi voicemail	request	MD5	REGISTER INVITE					24 Oct 2000 21:48
hgs@cs.columbia.edu	cs.columbia.edu	cgi voicemail admin	request	MD5	REGISTER INVITE	7042@cs.columbia.edu	mailto:hgs@cs.columbia.edu hgs@erlang.cs.columbia.edu		**kns10@cs.columbia.edu tk358@cs.columbia.edu lennox@cs.columbia.edu**	12 Oct 2000 18:43
kns10@cs.columbia.edu	Your login for kns10	cgi voicemail admin	required	MD5	REGISTER INVITE					12 Oct 2000 18:43
lennox@cs.columbia.edu	Password for cs.columbia.edu	cgi voicemail admin	request	MD5	REGISTER INVITE					13 Oct 2000 11:11
schulzrinne@cs.columbia.edu	Your cs.columbia.edu password									12 Oct 2000 18:43

The largest signaling network does not run SS7

- AT&T: 280 million calls a day
- AOL: 110 million emails/day, total about 18 billion/day
- total $>$ 1 billion instant messages/day (AOL: 500 million)
- telephony signaling $\approx$ IM, presence

Commonalities between signaling and events

- presence is just a special case of an event: “Alice logged in” $\approx$ “house temperature dropped below 50 deg.”
- need to locate mobile end points (for notifications)
- may need to find several different destinations
- same addressing for users
- presence often precursor to calls
- may replace call back, call waiting and voice mail tag
- likely to be found in same devices
- events already in VoIP: message alert, call events, conf. joins

SIP as a presence & event platform

- minimal SIP extension: SUBSCRIBE to request notifcations, NOTIFY when event occurs
- also, MESSAGE for IM, sessions for multi-party chats
- transition to true “chat” (and video)
- services such as reaching mobile phone while in meeting

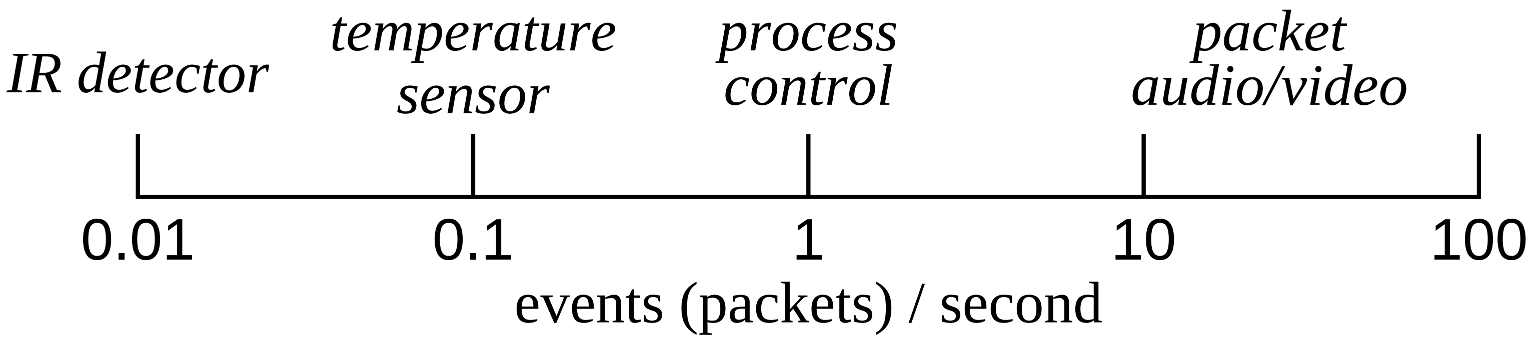
Architecture proposal

- integrate into common architecture for Internet-wide notification and messaging
  ➠ new basic internet service:

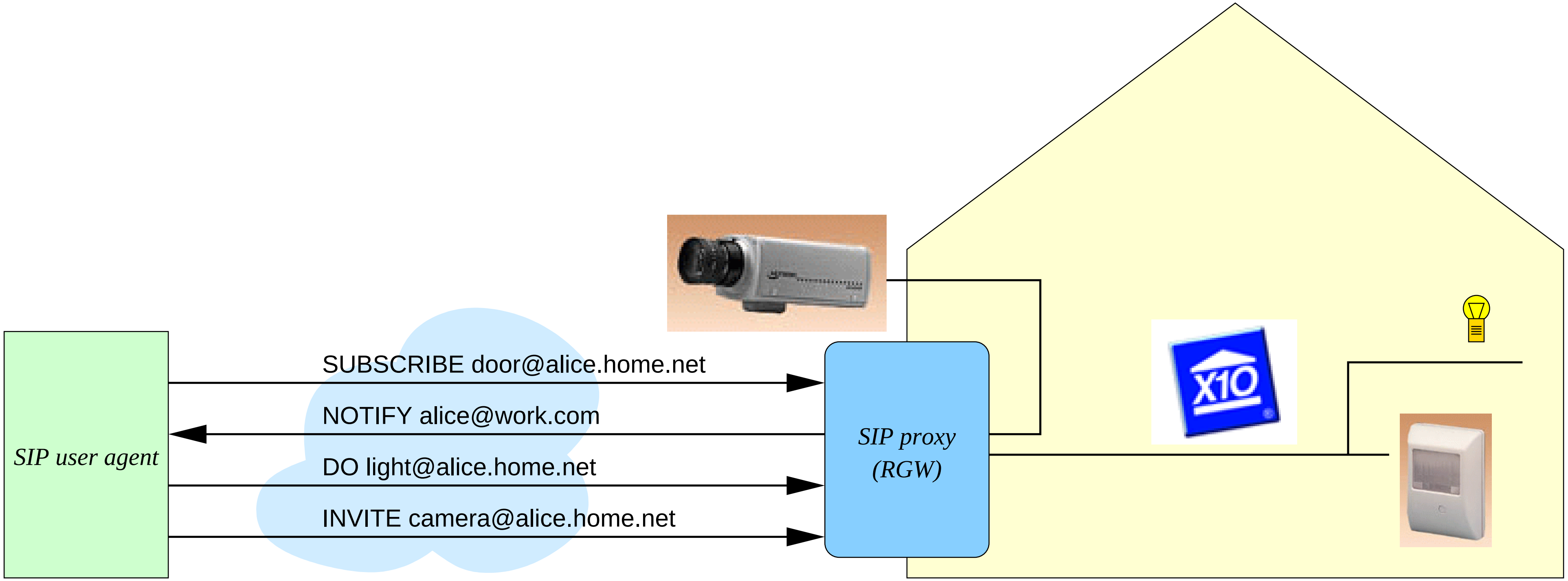
Asynchronous messaging with pickup	SMTP + POP/IMAP
Data retrieval	HTTP, ftp, tftp
Export computer UI	telnet, ssh, X11, vnc
Synchronous messaging	**SIP**

Observations

- single-valued (light-switch) to complex (CD changer) to multi-valued (temperature samples)
- both built-in and mediated (X10)
- often combined with audio/video in same system: security, industrial control, home entertainment
- notification rates vary ➠ gradual transition to continuous media



Events: SIP for appliances



Programming Internet multimedia services

Primarily, creation, forwarding, proxying, rejection of calls

APIs (Parlay, JAIN): protocol-neutral (SIP, H.323, ISUP), but may be least common denominator

SIP CGI: use Perl and other scripting languages; easy to learn

Servlets: Java only; faster than cgi; limited functionality

CPL: = XML-based language for *user* service creation; portable across providers, but not all services

- Protocol-neutral: Parlay, JAIN, CPL
- Call creation: Parlay, JAIN
- VoiceXML is for voice-service creation *after* call setup

Example Call Processing Language script

```
<?xml version="1.0" ?>
<!DOCTYPE cpl SYSTEM "cpl.dtd">

<cpl>
 <subaction id="voicemail">
  <location url=
    "sip:kns10@vm.cs.columbia.edu">
   <redirect />
  </location>
 </subaction>

 <incoming>
  <address-switch field="origin"
     subfield="host">
   <address
     subdomain-of="cs.columbia.edu">
    <location url=
      "sip:kns10@cbb.cs.columbia.edu">
     <proxy>
      <busy>
        <sub ref="voicemail" />
      </busy>
      <noanswer>
        <sub ref="voicemail" />
      </noanswer>
      <failure>
        <sub ref="voicemail" />
      </failure>
     </proxy>
    </location>
   </address>
   <otherwise>
    <sub ref="voicemail" />
   </otherwise>
  </address-switch>
 </incoming>
</cpl>
```

Conclusion

- basic IETF-based architecture in place
- SIP as foundation for services
- extensions for mobility, emergency services, . . . in progress
- first (and last?) chance to recover from 120 years of legacy
- avoid replication of PSTN on packets
- most VoIP applications won't look like telephones
- range of engagement and asynchronicity, from call to IM to email
- challenge of mobile services

For more information. . .

SIP: http://www.cs.columbia.edu/sip

RTP: http://www.cs.columbia.edu/˜hgs/rtp

Papers: http://www.cs.columbia.edu/IRT