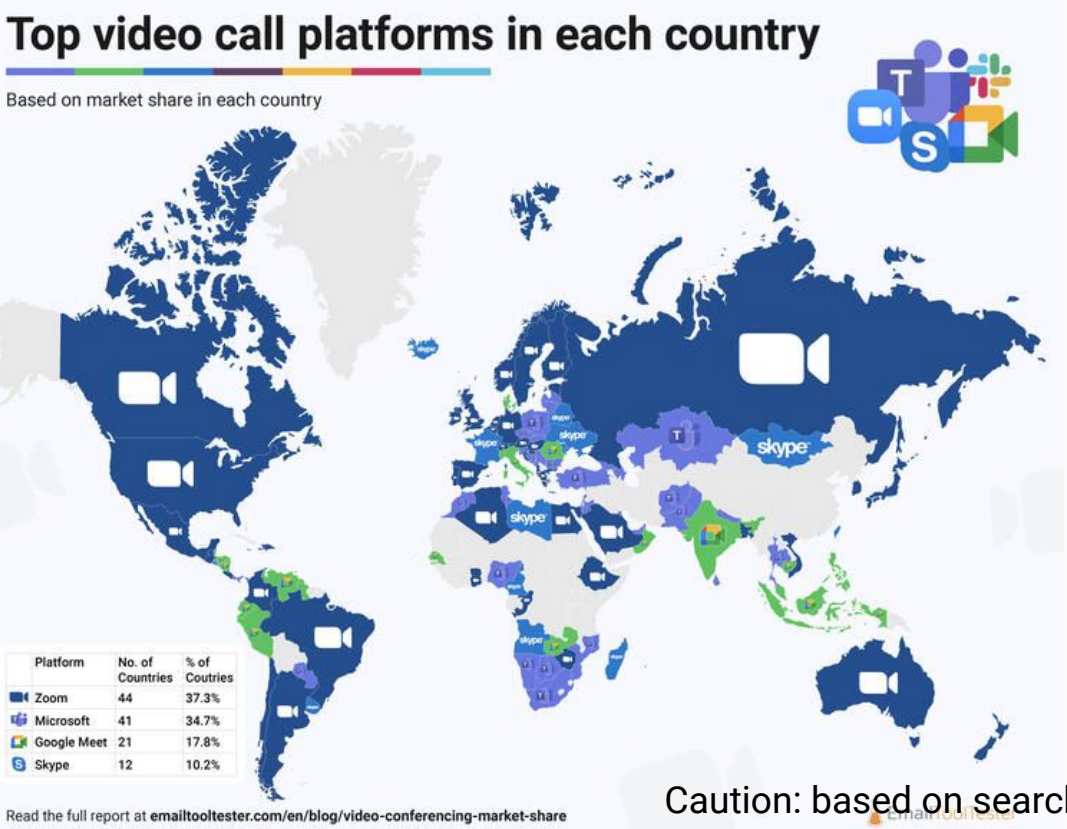
Why is Zoom so much more popular than standards-based videoconferencing?

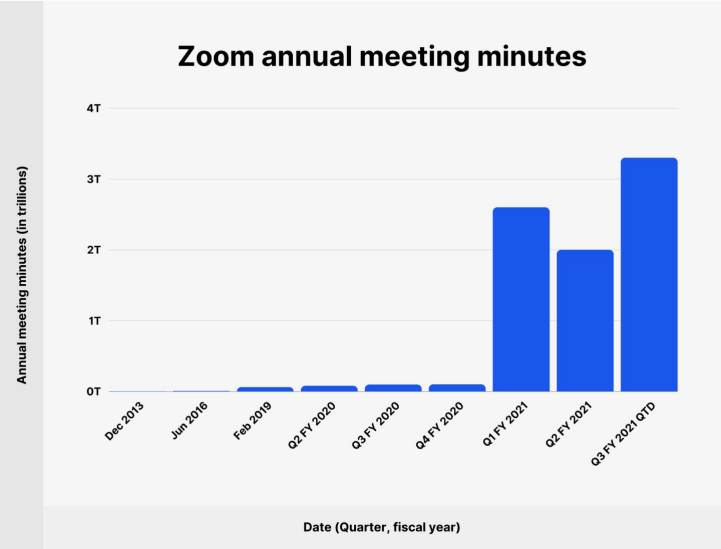
Henning Schulzrinne - Columbia University

Only very few systems matter



Caution: based on searches, not minutes or users

Not news: Lots of people spend lots of time on video



~ 5.7 M people on 24/7
300M participants per day

Microsoft Teams

115M daily users

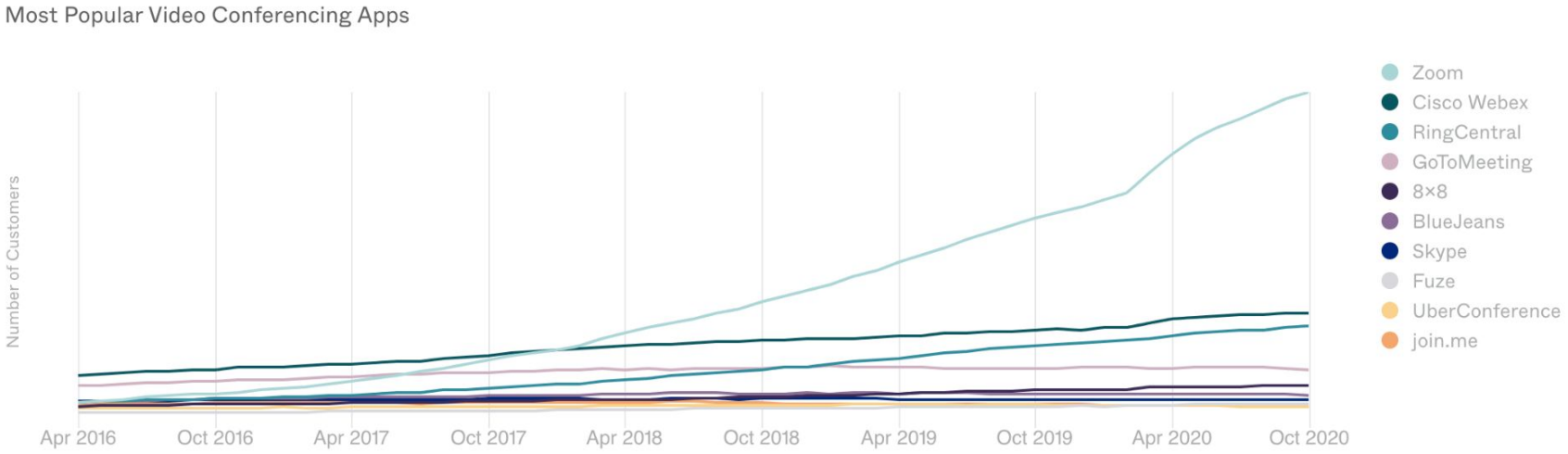
Google

100M daily users

webex by cisco

300 M users

No surprise, either: Video conferencing ≅ Zoom



AT&T videophone 1995 ($1,499 or $30/day)

1964

Video relay service: VP-100 (2000)

H.323 (TV)

Reach users by E.164 phone number

Now: SIP-based
Probably largest interoperable, public video network
(IETF RUM working group working on profile)

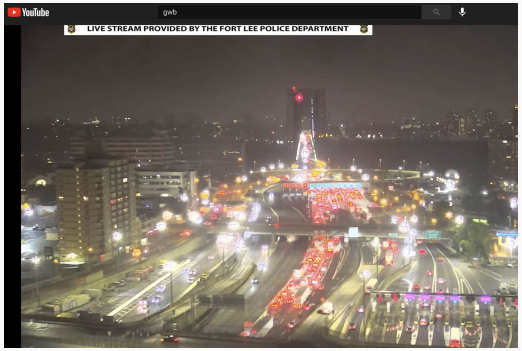
The landscape of IP video communications

Unified Communications (UC)

2-party phone call, spontaneous

CuSeeMe (1992)

- Differentiated roles (organizer, panelist, audience)
- Some audience participation
- Up to 50,000 participants

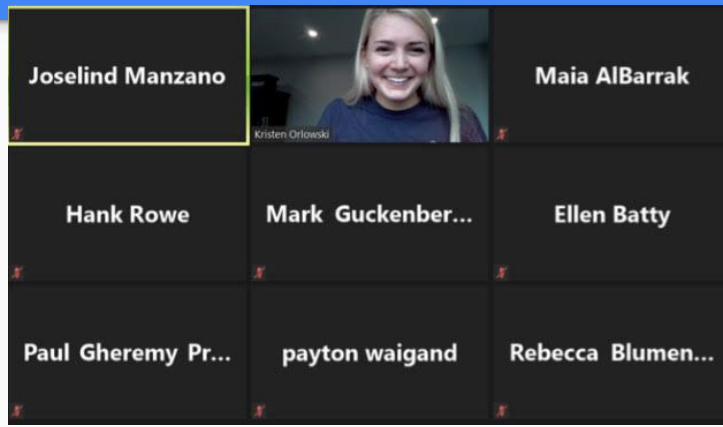


Multi-party streaming (Mbone, YouTube, FB Live, Livestream)

- One way, except chat & comments

Lessons learned since 1964

- Two-party video is rarely useful except for specialty applications (telemedicine & adult entertainment)
  - But popular for environment sharing (“let me show you my new apartment”)
- Most video “calls” are scheduled → call signaling by calendar and SMTP, not SIP
- Chat and screen sharing are the most useful Zoom features
- The most useful video conferencing accessory is a better *microphone* (and maybe a ring light)



Video calls as basic augmented reality

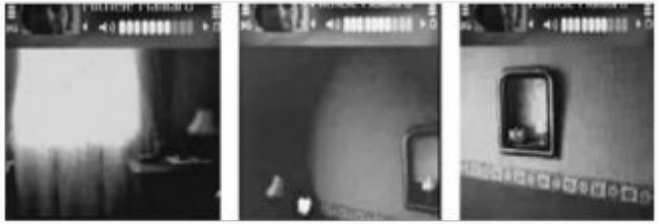
Mundane Video Directors in Interaction: Showing One's Environment in Skype and Mobile Video Calls

By CHRISTIAN LICOPPE, JULIEN MOREL

Book Studies of Video Practices

Edition 1st Edition

First Published 2014



Figures 14–16 The images produced by the call recipient during the caller's noticing turn (lines 22–23) as she pans the camera to the right from the window to the wall.

From Get To Know

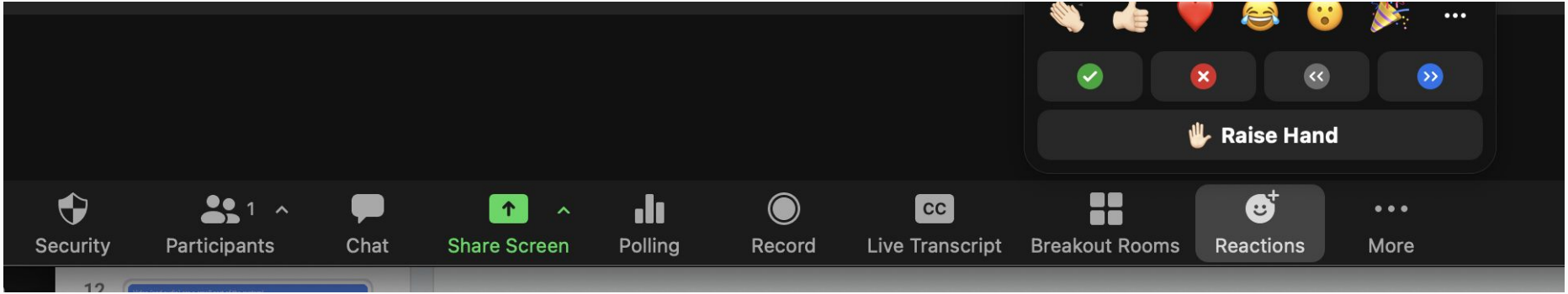
How to make the most of NYC apartment tours via FaceTime and Zoom

By Michelle Sinclair Colman
Tuesday, June 16, 2020

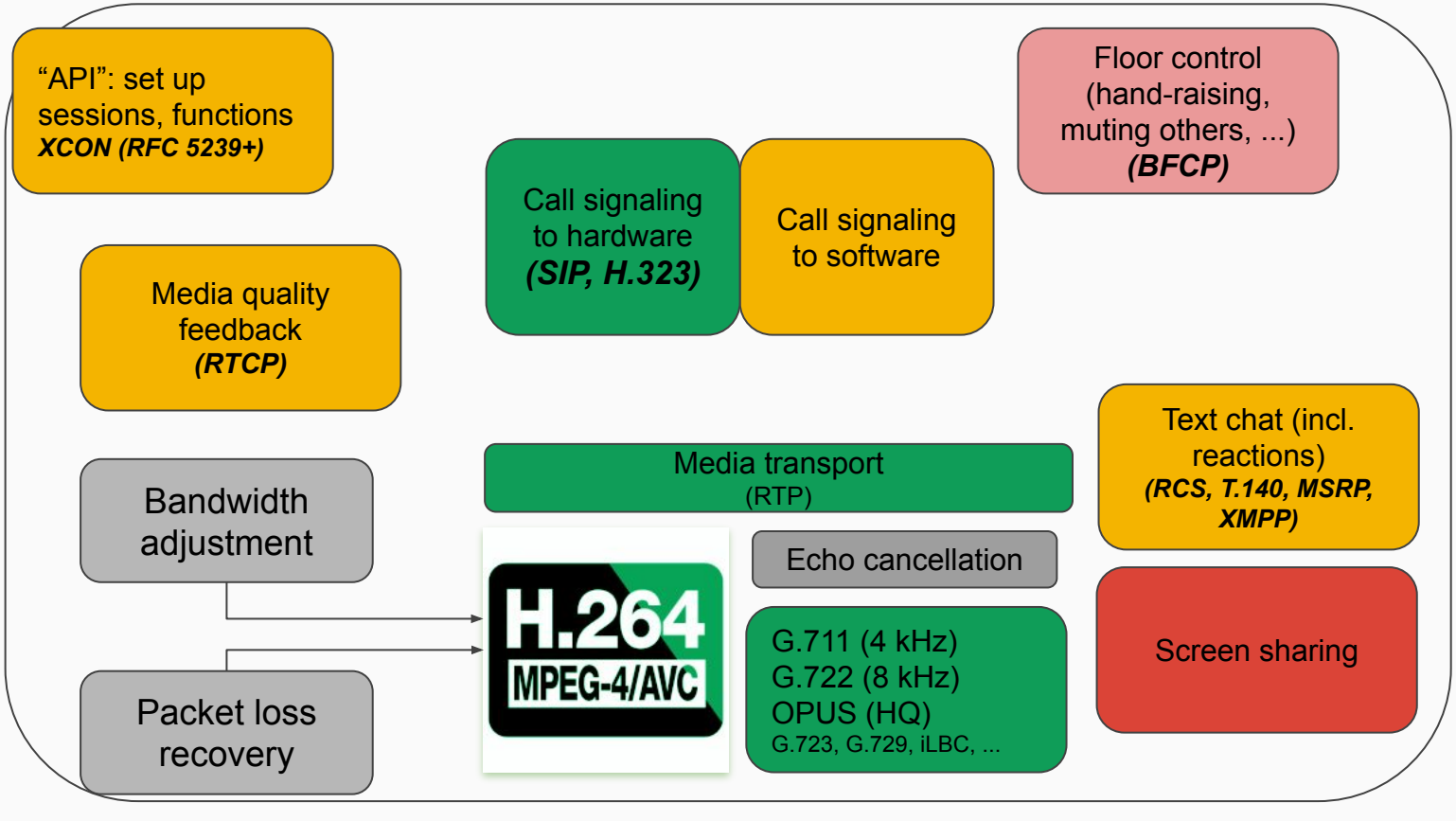
Such a mobility turn in video communication enables participants to show something to their interlocutor. Thirty percent of mobile video conversations seem to unfold around the intent of one of the participants to show something to the other, which is probably an underestimate because showing also occurs in video calls that do not have that as an initial goal. From what we observed in the Skype part of our own corpus, the numbers should be much in the same range also for Skype interactions. With the possibility of video communication technologies being able to show something during a call, these at last seem to fulfill their early and heretofore unkept promise that they would allow remote conversationalists to share their environments. A related line of research has looked at "video-as-data," that is, how some part of the ongoing activity could be recorded and made available in real time to provide a shared field of interaction in collaborative situations. In such a configuration, the participants work to articulate video and speech occurrences in a way that is relevant to the unfolding interaction.

What we think Zoom is...

The hard part for interoperable video interaction



Video (and audio) are a small part of the system!



Required, available and widely used

Available, but not widely used

Not available or limited functionality

No (interoperability) need for standards

Standards = technology translator

- Similar in some ways to textbooks
- “accepted technology”
  - lower/known risks (“vetted”)
  - infrastructure (“eco system”)
    - libraries, test tools, text books, certification, …
  - reduce cost of picking among roughly equal choices
  - sometimes reduce IPR risks (“patent pool”, RAND)
- requires expertise and broader training

  many CS standards don’t have either

  example: HTTP/1.0, HTML 1.0, 802.11 WEP

394 SIP (and related) RFCs (incomplete)

SIP Standards

Core SIP Documents

RFC	Document Title
RFC 2543	SIP: Session Initiation Protocol (obsolete)
RFC 3261	SIP: Session Initiation Protocol
RFC 3262	Reliability of Provisional Responses
RFC 3263	Locating SIP Servers
RFC 3265	SIP-Specific Event Notification
RFC 5954	Essential Correction for IPv6 ABNF and URI Comparison in RFC 3261

SDP-Related Documents

RFC	Document Title
RFC 2327	Session Description Protocol (SDP) (obsolete: see RFC 4566)
RFC 3264	An Offer/Answer Model with the Session Description Protocol (SDP)
RFC 3266	Support of IPv6 in SDP
RFC 3388	Grouping Media Lines in SDP (obsolete: see |RFC 5888|)
RFC 3407	Session Description Protocol (SDP) Simple Capability Declaration
RFC 3524	Mapping of Media Streams to Resource Reservation Flows
RFC 3556	SDP Bandwidth Modifiers for RTCP Bandwidth
RFC 3605	Real Time Control Protocol (RTCP) attribute in Session Description Protocol (SDP)
RFC 3890	A Transport Independent Bandwidth Modifier
RFC 4091	An Alternative NAT Semantics for SDP
RFC 4145	TCP-Based Media Transport in the SDP
RFC 4566	Session Description Protocol (SDP)
RFC 4567	Key Management Extensions for SDP and RTSP
RFC 4568	SDP Security Descriptions for Media Streams
RFC 4570	SDP Source Filters
RFC 4572	Connection-Oriented Media Transport over TLS in SDP
RFC 4574	SDP Label Attribute

roughly 300 with SIP in title (RFC editor)

IMS 23.228: 329 pg.
RCS 5.1: 482 pg.

Simple core protocols have acquired technical debts

RFC	Type	Status	Title	Bgnd	Prot	Names	Ops	RR	Proxy	Stub	Auth	Res			
882		Obsolete	Domain Names – Concepts and Facilities	x		x	x				x				
883		Obsolete	Domain Names – Implementation and Specification		x		x	x							
920			Domain Requirements				x								
973		Obsolete	Domain System Changes and Observations			x		x			x				
1032			Domain Administrators Guide				x								
1033			Domain Administrators Operations Guide				x								
1034	Standard		Domain Names – Concepts and Facilities	x		x	x			x	x	x			
1035	Standard		Domain Names – Implementation and Specification		x	x		x			x	x	x		
1101			DNS Encoding of Network Names and Other Types			x									
1123	Standard		Requirements for Internet Hosts – Application and Support	x							x	x			
1178	Informational		Choosing a Name for Your Computer				x								

DNS:
~143 active RFCs

Sidebars: XCON and CCMP

IETF attempt in 2008-2012 to standardize basic conference management

Data model for conference (XML)

e.g., user admission, sidebars (breakout rooms), floors

API (operations) on data model → CCMP

Left out polling, advanced breakout functions, waiting rooms, ...

Addressing - vision & reality

Original idea: SIP URLs (sip:user@domain) or tel URLs (tel:+1-201-555-0123)

still exists and useful for hardware

Current reality: web URLs via web page, email, calendar, Slack, IM, SMS, ...

Beyond protocols - what do users expect?

Video conferences:

- NAT traversal
- Cross-domain authentication and authorization
- Calendar interface
- Media routing
- Scalable capacity (tens to thousands per session)
- End-to-end security
- Media gateways (phone, room systems)
- Polling
- Recording and playback
- Transcription (accessibility, records)
- Language translation
- Managing abuse (“Zoom bombing”, criminal activity, extremism)

Webinars:

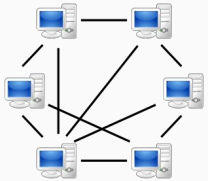
- Attendee management
- Connect to YouTube, Facebook Live, ...
- Monetization
- Polling and “engagement”

Operational models

Enterprise

PBX heritage
“Unified communications”
Hosted in corporate data center

Peer-to-peer



Early Skype architecture
Common elsewhere: SMTP, XMPP, IRC*, Usenet
but usually large user/server ratio

jami
a GNU package

Carrier

SIP-based: RCS (mostly messaging)
struggled with higher-quality audio (HD audio)

Cloud

Rooted in corporate heritage
Struggling with consumer use (and abuse)

Not quite peer-to-peer: “permissioned” networks

IRC

today	yesterday	network	users Ø	channels Ø	servers Ø
1.	1.	Libera.Chat	36564	18711	27
2.	2.	IRCnet	20115	10685	23
3.	3.	Undernet	14574	6065	34
4.	4.	EFnet	11765	6892	17
5.	5.	OFTC	11623	2327	11
6.	6.	Rizon	11511	8803	16
7.	7.	QuakeNet	9909	8780	26
8.	8.	DALnet	7839	3861	38
9.	9.	Snoonet	4262	5734	17
10.	10.	GIMPnet	3352	368	6
11.	11.	KampungChat	3197	459	13
12.	12.	hackint	3195	1753	9
13.	13.	GeekShed	3175	219	4
14.	14.	P2P-NET	2757	722	13
15.	15.	SimosNap	2631	522	10
16.	16.	OltreIrc	2596	30	14
17.	17.	ExplosionIRC	2591	61	9
18.	18.	EsperNet	2430	2533	11
19.	19.	GameSurge	2122	1639	12
20.	20.	synIRC	2092	1103	15
21.	21.	Abjects	2074	341	11
22.	22.	SceneP2P	1771	68	7
23.	23.	IRCHighWay	1445	661	17
24.	24.	EuropNet	1353	983	7
25.	25.	OpenJoke	1095	51	27
26.	26.	Geveze	1041	84	5
27.	27.	tilde.chat	1006	445	12

Freenode IRC staff resign en masse, unhappy about new management

Network boss Andrew Lee disputes claims made by those leaving the internet chat community

Thomas Claburn in San Francisco — Wed 19 May 2021 // 21:50 UTC

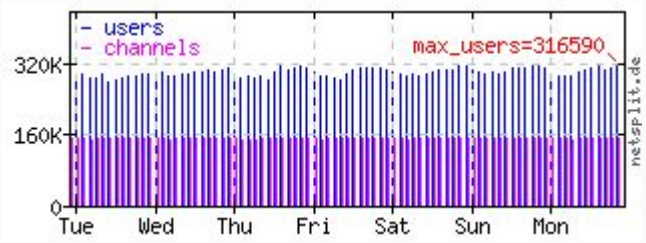
32

UPDATED Most of the volunteer staff of Freenode, an internet relay chat (IRC) network dating back to 1995, have resigned in protest over what they describe as a hostile takeover of the chat service.

And many have launched an alternative service, Libera Chat.

Freenode, which has focused on serving as a real-time communication channel for free and open source software projects, currently has about 76,000 users and 42,000 chat rooms.

In a resignation letter, a staffer called Christian, who is also known as Fuchs on Freenode, said after 10 years helping with the network, he is leaving because he disagrees with the direction being taken by Andrew Lee, founder of VPN firm Private Internet Access (PIA), who acquired a controlling interest [PDF] in Freenode's holding company in 2017.



What are the strengths of the operational models?

Feature	Enterprise hosted	Peer-to-peer	Carrier	“VCaaS”
Predictable features	Mostly	Difficult	Unlikely (Android!)	Mostly
Cross-domain AA	guests with passwords	sybils	“roaming”	added SSO, but still mostly secret strings
Media routing	rare	challenging	usually national only	As far as the cloud will stretch
Scalable capacity	rare	freeloader problem	struggling with cloud	natural
End-to-end security	easy	easy for 2-party, no mixing	wiretapping laws	challenging with media mixing
Media gateways	PBX dial-in	nobody ever tried*	“we are the phone company!”	outsourced
Recording & playback	with effort (rare)	nobody ever tried	struggling with cloud	easy
Transcription, translation	challenging	nobody ever tried	similar to VCaaS	in progress
Manage abuse	Challenging for smaller entities (schools, nonprofits)	lots of PhD theses were written	have fraud & security departments, but “common carrier” tradition	incompatible with no-touch model; unexpected role

But it’s really the business model that killed interoperability

Old models: Open source, enterprise software license or built into phone

Open source: who is going to run the server → open source companies get bought by operations (“cloud”) companies (e.g., Jitsi)

Enterprise: who wants to run and maintain a PBX server?

see: email outsourcing

Caller pays is back: Caller (= host) pays for meeting; participants are free

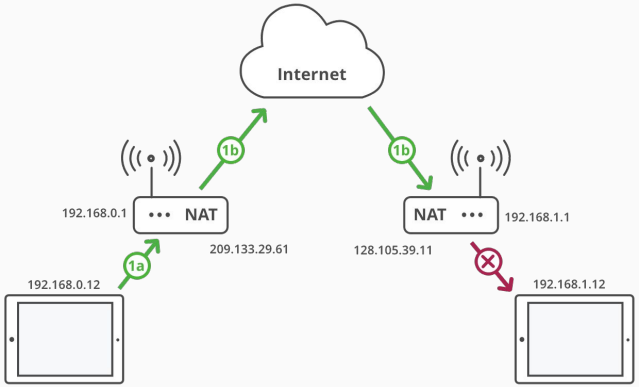
NATs killed the peer-to-peer model

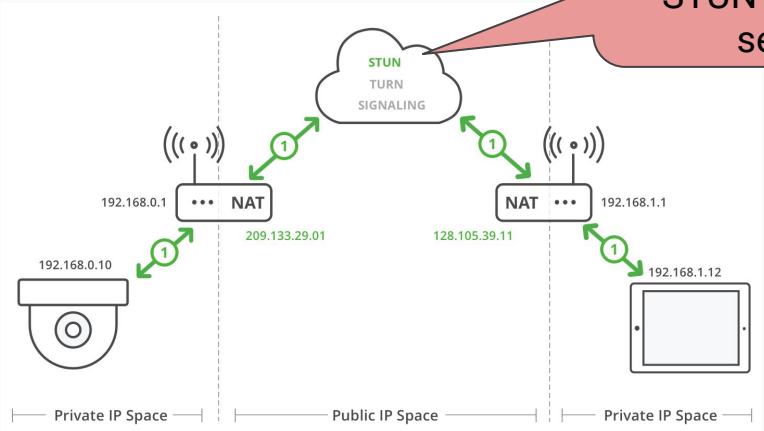
VoIP clients need inbound connections for call signaling and media

Video conference clients rely on participants to initiate sessions and participation - outbound only signaling — but still may need inbound media

Late 1990s: The only users with enough bandwidth didn't have NATs
Early 2000s: NATs are evil and IPv6 will kill them

somebody has to provide the STUN and TURN servers

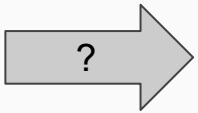




https://anyconnect.com/stun-turn-ice/

The versioning problem

XEP-0436	MUC presence versioning	Standards Track	Experimental	2020-05-10
XEP-0437	Room Activity Indicators	Standards Track	Experimental	2020-05-05
XEP-0438	Best practices for password hashing and storage	Informational	Experimental	2020-10-30
XEP-0439	Quick Response	Standards Track	Experimental	2020-05-05
XEP-0440	SASL Channel-Binding Type Capability	Standards Track	Experimental	2020-08-04
XEP-0441	Message Archive Management Preferences	Standards Track	Experimental	2020-08-25
XEP-0442	Pubsub Message Archive Management	Standards Track	Experimental	2020-08-25
XEP-0443	XMPP Compliance Suites 2021	Standards Track	Draft	2020-11-24
XEP-0444	Message Reactions	Standards Track	Experimental	2020-10-13
XEP-0445	Pre-Authenticated In-Band Registration	Standards Track	Experimental	2020-11-24
XEP-0446	File metadata element	Standards Track	Experimental	2020-11-24
XEP-0447	Stateless file sharing	Standards Track	Experimental	2020-12-30
XEP-0448	Encryption for stateless file sharing	Standards Track	Experimental	2020-11-24
XEP-0449	Stickers	Standards Track	Experimental	2020-11-24
XEP-0450	Automatic Trust Management (ATM)	Standards Track	Experimental	2021-06-27
XEP-0451	Stanza Multiplexing	Standards Track	Experimental	2021-01-19
XEP-0452	MUC Mention Notifications	Standards Track	Experimental	2021-02-12
XEP-0453	DOAP usage in XMPP	Informational	Experimental	2021-01-26
XEP-0454	OMEMO Media sharing	Historical	Experimental	2021-01-26
XEP-0455	Service Outage Status	Standards Track	Experimental	2021-02-09
XEP-0456	Content Rating Labels	Standards Track	Experimental	2021-03-28
XEP-0457	Message Fancying	Humorous	Active	2021-04-01
XEP-0458	Community Code of Conduct	Procedural	Experimental	2021-06-29
XEP-0459	XMPP Compliance Suites 2022	Standards Track	Experimental	2021-06-22



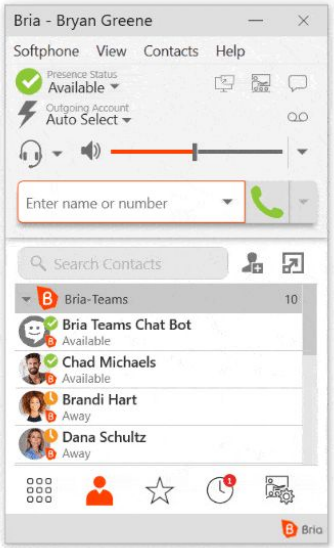
Project Name	Platforms
Aparté	BSD / Linux / macOS
AstraChat	Android / iOS / Linux / macOS / Windows
BeagleIM by Tigase, Inc.	macOS
blabber.im	Android
Bruno the Jabber™ Bear	Android
Conversations	Android
Converse	Browser
Dino	Linux
Gajim	Linux / Windows
Kaidan	Android / Linux / macOS / Other / Windows
Monal IM	iOS / macOS
Movim	Android / Browser / Linux / macOS / Windows
Poezio	Linux / macOS
Profanity	Linux / macOS / Windows
Psi	Linux / macOS / Windows
Psi+	Linux / macOS / Windows
Pàdé	Browser

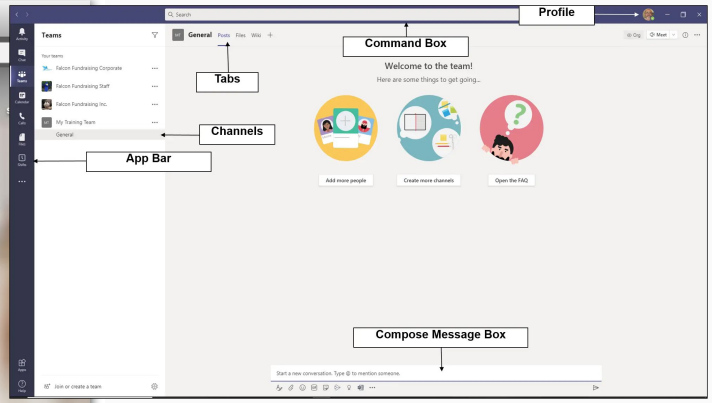
WebRTC as transition model

Standards-based client

WebRTC client

Application



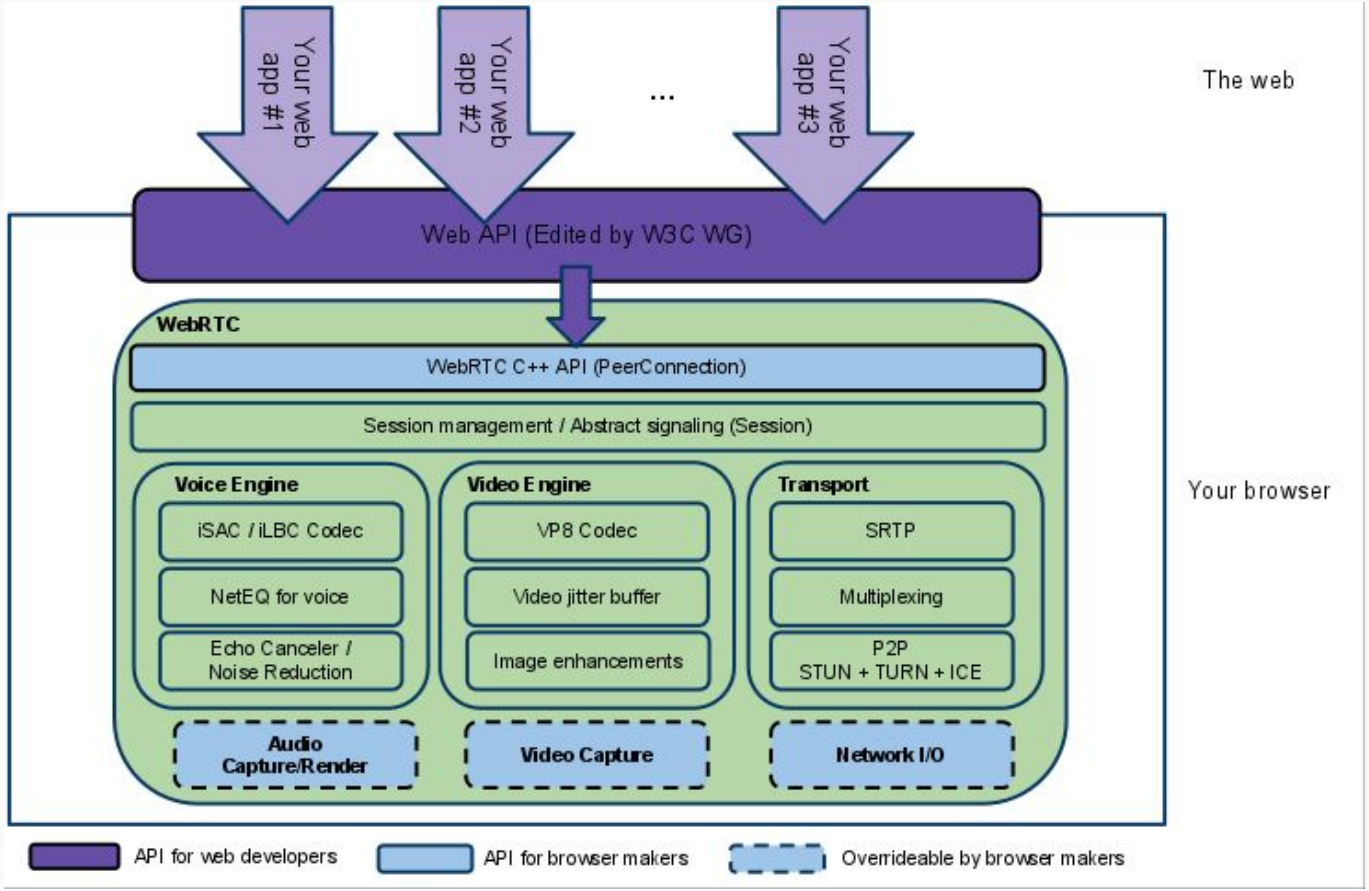


multiple services, one client

no installation - one “page” per service
switch browsers & maybe platforms
no interoperability between services

No interoperability between services

WebRTC architecture



Typical WebRTC architecture

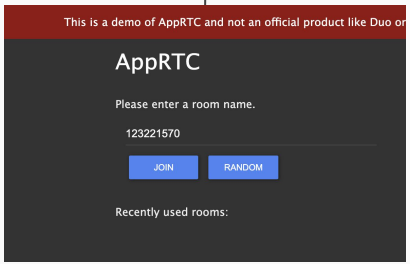
STUN
ICE

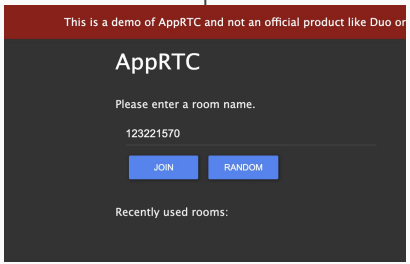
websocket
(bidirectional TCP)

JANVS
WEBRTC SERVER

SRTP (secure
media transport)

Apache or nginx serve JS and HTML

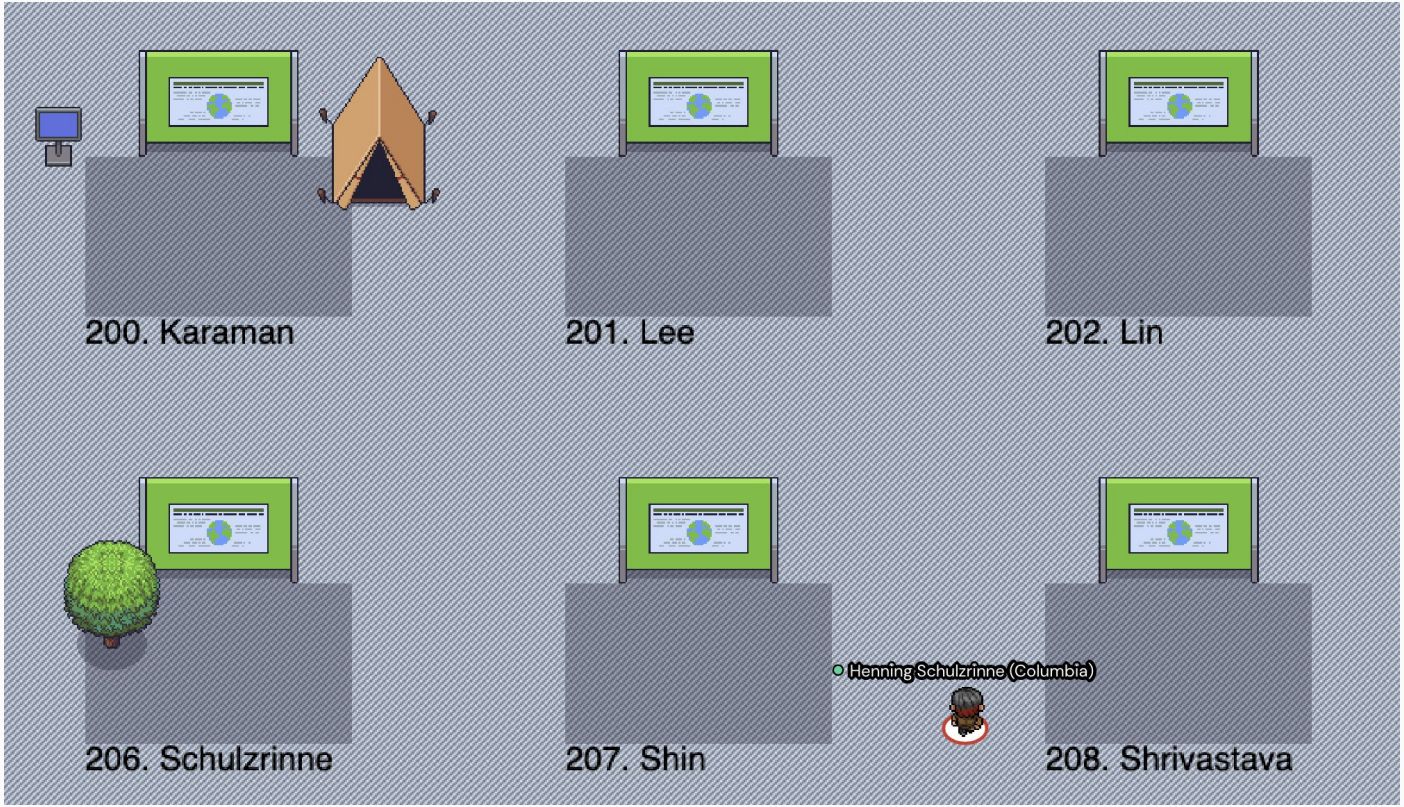




proprietary session signaling (can be SIP or XMPP)

Good for non-square UIs

gather.town



advantages to break-out rooms?

Or lower level still - browser as VM

WebAssembly SIMD: SIMD instructions, e.g., to replace video background

WebTransport: multiple cancellable streams: datagrams + bidirectional reliable streams

WebCodecs API: direct access to codecs

Bifurcation

Communication out front applications: collaboration, social interaction, telemedicine

- challenge: hybrid interactions → AR with remote participants?
- challenge: more structured meetings (e.g., recorded votes)
- challenge: unwanted communications -- robocalls and QAnon

Video in back applications: monitoring (traffic, agriculture, security, …) → consumers are ML applications

Conclusion

Video worked out quite differently than anticipated in the 1990s

- probably the component everybody would ditch first for Zoom and kin

Standards-based communications survived where communication without prior arrangement is valued → phone, email, SMS

We think codecs and protocols → systems and operations

Moving from protocol standards to browser as hardware abstraction layer

- happening with transport protocols, too (see QUIC)