What can we learn from 20+ years of SIP and VoIP?

Henning Schulzrinne - Columbia University

STIR/SHAKEN Enterprise Summit - October 17, 2022

Exploring past and future

What landscape did SIP emerge from?

What (likely) made it successful?

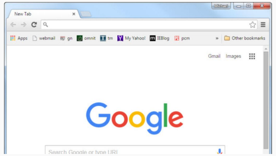
Why aren't Zoom/Teams/WebEx/… using SIP & RTP (mostly)?

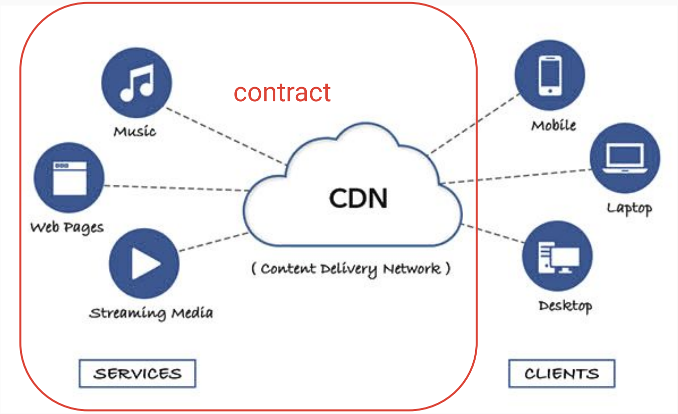
What have we learned about video calls & conferencing since 1964?

What could be next?

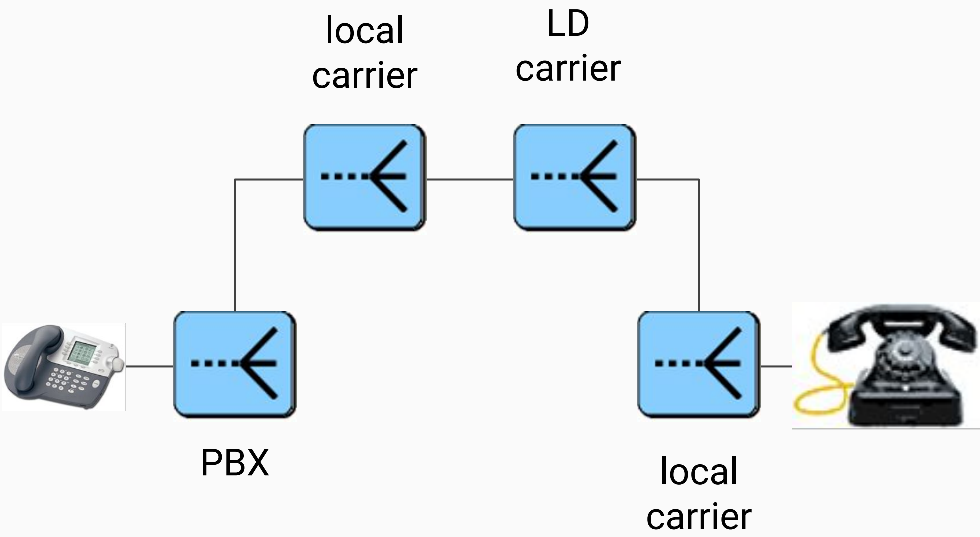
Where did we start?

Web vs. VoIP





request routing static, initiated by client
mostly DNS + some HTTP redirect



fairly sophisticated price-based routing
largely provider-driven

December 1996

[Search] [txt|ps|pdfized|bibtex] [Tracker] [WG] [Email] [Diff1] [Diff2] [Nits
Versions: 01 02 03 04 05 06 07 08 09 10 11 12 IPR declarations
rfc2543

```
Internet Engineering Task Force                          MMUSIC WG
INTERNET-DRAFT                  M. Handley, H. Schulzrinne, E. Schooler
draft-ietf-mmusic-sip-01.txt                   ISI,Columbia,Caltech
                                                       2nd Dec 1996
                                           Expires: 2nd June 1997
```

SIP: Session Initiation Protocol

Abstract

Many styles of multimedia conferencing are likely to co-exist on the Internet, and many of them share the need to invite users to participate. The Session Initiation Protocol (SIP) is a simple protocol designed to enable the invitation of users to participate in such multimedia sessions. It is not tied to any specific conference control scheme, providing support for either loosely or tightly controlled sessions. In particular, it aims to enable user mobility by relaying and redirecting invitations to a user's current location.

This document is a product of the Multiparty Multimedia Session Control (MMUSIC) working group of the Internet Engineering Task Force. Comments are solicited and should be addressed to the working group's mailing list at confctrl@isi.edu and/or the authors.

INTERNATIONAL TELECOMMUNICATION UNION

ITU-T **H.323**

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

(11/96)

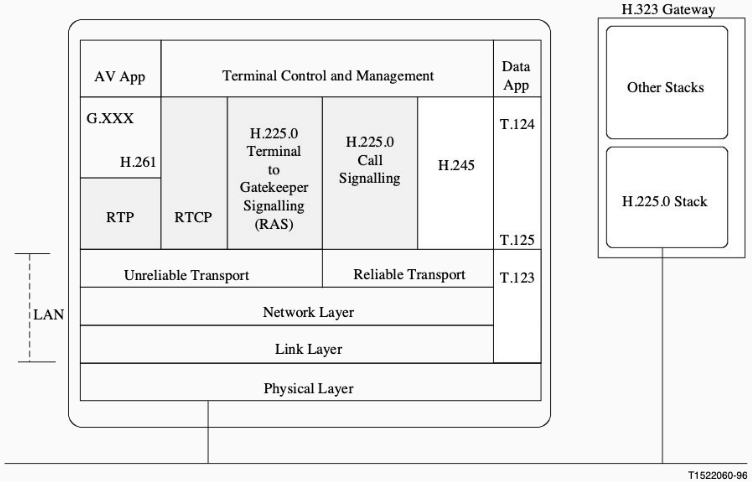
H.323 [70pg.]: RTP + Q.931 [349] + H.225.0 [104] + ...

ISDN *1988, † 2010-2018

ISDN was introduced by CCITT (ITU-T) in 1988 and [310318398_1602629443530781_2838172747381382600_n.mp4] ng deployed with varying success in countries around the world such as Japan, Australia, India and the United States. The biggest impact was in Europe, however, in countries like Norway, Denmark, Switzerland and above all Germany, which had 25 million channels (29% penetration) and one in five lines installed worldwide.

SIP could be explained on a (small) napkin

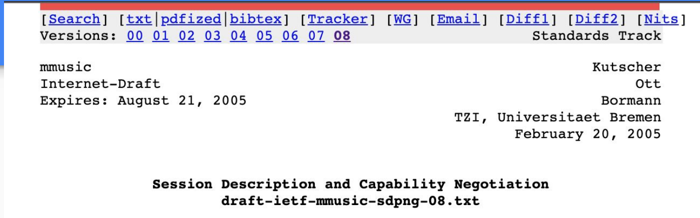
```
C->S: INVITE sip:watson@boston.bell-tel.com SIP/2.0
      Via: SIP/2.0/UDP kton.bell-tel.com
      From: A. Bell <sip:a.g.bell@bell-tel.com>
      To: T. Watson <sip:watson@bell-tel.com>
      Call-ID: 3298420296@kton.bell-tel.com
      CSeq: 1 INVITE
      Subject: Mr. Watson, come here.
      Content-Type: application/sdp
      Content-Length: ...

      v=0
      o=bell 53655765 2353687637 IN IP4 128.3.4.5
      s=Mr. Watson, come here.
      c=IN IP4 kton.bell-tel.com
      m=audio 3456 RTP/AVP 0 3 4 5
```



What (might have) been reasons for success

- **Timing**: early enough before proprietary or ITU-T solutions could catch up
  - ISDN compatibility was never incentive enough
  - but SIP was close enough to feature parity to digital PBX and analog phones (parallel forking!)
- **Scope**: Competitor H.323 was focused on conference rooms, not calls
  - remained niche market
- **Familiarity**: HTTP-like syntax and re-use
  - could be stateless (until SBCs took over)
- **Low barrier to entry**: text-based, UDP, copy-paste examples
  - pass the “assign as homework” test
  - H.323 had mix of Q.931 bit-based TLV & ASN.1 (H.225.0 & H.245)

But these also proved to be troublesome


[Search] [txt|pdfized|bibtex] [Tracker] [WG] [Email] [Diff1] [Diff2] [Nits]
Versions: 00 01 02 03 04 05 06 07 08 Standards Track

mmusic Kutscher
Internet-Draft Ott
Expires: August 21, 2005 Bormann
TZI, Universitaet Bremen
February 20, 2005

Session Description and Capability Negotiation
draft-ietf-mmusic-sdpng-08.txt

- **UDP transport:** significant edge-case complexity
  - embedding retransmission adds complexity (multi-hop)
  - mixed transport protocols add failure modes
  - lots of SIP headers + larger bodies + TLS ⇒ bad idea
- **SDP for media**: offer-answer has been trouble
  - hard to add structured alternatives and parameters
  - SDPng never made it (2nd system syndrome…)
- **Protocol encoding**: interoperability issues (code to example, not spec)
  - angry fruit salad of SMTP, HTTP/1.1, base-64 JWTs, SDP, MIME multipart, …
  - relatively few libraries → HTTP/3 binary mode affects few
  - “*While these exchanges are human readable, using whitespace for message formatting leads to parsing complexity and excessive tolerance of variant behavior*” (RFC 9114)

Many (most?) SIP vulnerabilities are parser-related

There are **467** CVE Records that match your search.

Name	Description
CVE-2022-31031	PJSIP is a free and open source multimedia communication library written in C language implementing standard based protocols such as SIP, SDP, RTP, STUN, TURN, and ICE. In versions prior to and including 2.12.1 a stack buffer overflow vulnerability affects PJSIP users that use STUN in their applications, either by: setting a STUN server in their account/media config in PJSUA/PJSUA2 level, or directly using `pjlib-util/stun_simple` API. A patch is available in commit 450baca which should be included in the next release. There are no known workarounds for this issue.
CVE-2022-31003	Sofia-SIP is an open-source Session Initiation Protocol (SIP) User-Agent library. Prior to version 1.13.8, when parsing each line of a sdp message, `rest = record + 2` will access the memory behind `\0` and cause an out-of-bounds write. An attacker can send a message with evil sdp to FreeSWITCH, causing a crash or more serious consequence, such as remote code execution. Version 1.13.8 contains a patch for this issue.
CVE-2022-31002	Sofia-SIP is an open-source Session Initiation Protocol (SIP) User-Agent library. Prior to version 1.13.8, an attacker can send a message with evil sdp to FreeSWITCH, which may cause a crash. This type of crash may be caused by a URL ending with `%`. Version 1.13.8 contains a patch for this issue.
CVE-2022-31001	Sofia-SIP is an open-source Session Initiation Protocol (SIP) User-Agent library. Prior to version 1.13.8, an attacker can send a message with evil sdp to FreeSWITCH, which may cause crash. This type of crash may be caused by `#define MATCH(s, m) (strncmp(s, m, n = sizeof(m) - 1) == 0)`, which will make `n` bigger and trigger out-of-bound access when `IS_NON_WS(s[n])`. Version 1.13.8 contains a patch for this issue.
CVE-2022-29855	Mitel 6800 and 6900 Series SIP phone devices through 2022-04-27 have "undocumented functionality." A vulnerability in Mitel 6800 Series and 6900 Series SIP phones excluding 6970, versions 5.1 SP8 (5.1.0.8016) and earlier, and 6.0 (6.0.0.368) through 6.1 HF4 (6.1.0.165), could allow a unauthenticated attacker with physical access to the phone to gain root access due to insufficient access control for test functionality during system startup. A successful exploit could allow access to sensitive information and code execution.
CVE-2022-29330	Missing access control in the backup system of Telesoft VitalPBX before 3.2.1 allows attackers to access the PJSIP and SIP extension credentials, cryptographic keys and voicemails files via unspecified vectors.
CVE-2022-27255	In Realtek eCos RSDK 1.5.7p1 and MSDK 4.9.4p1, the SIP ALG function that rewrites SDP data has a stack-based buffer overflow. This allows an attacker to remotely execute code without authentication via a crafted SIP packet that contains malicious SDP data.
CVE-2022-26370	On F5 BIG-IP 16.1.x versions prior to 16.1.2.2, 15.1.x versions prior to 15.1.5, and 14.1.x versions prior to 14.1.4.6, when a Session Initiation Protocol (SIP) message routing framework (MRF) application layer gateway (ALG) profile is configured on a Message Routing virtual server, undisclosed requests can cause the Traffic Management Microkernel (TMM) to terminate. Note: Software versions which have reached End of Technical Support (EoTS) are not evaluated
CVE-2022-23608	PJSIP is a free and open source multimedia communication library written in C language implementing standard based protocols such as SIP, SDP, RTP, STUN, TURN, and ICE. In versions up to and including 2.11.1 when in a dialog set (or forking) scenario, a hash key shared by multiple UAC dialogs can potentially be prematurely freed when one of the dialogs is destroyed . The issue may cause a dialog set to be registered in the hash table multiple times (with different hash keys) leading to undefined behavior such as dialog list collision which eventually leading to endless loop. A patch is available in commit db3235953baa56d2fb0e276ca510fefca751643f which will be included in the next release. There are no known workarounds for this issue.
CVE-2022-23025	On BIG-IP version 16.1.x before 16.1.1, 15.1.x before 15.1.4, 14.1.x before 14.1.4.4, and all versions of 13.1.x, when a SIP ALG profile is configured on a virtual server, undisclosed requests can cause the Traffic Management Microkernel (TMM) to terminate. Note: Software versions which have reached End of Technical Support (EoTS) are not evaluated.
CVE-2022-22204	An Improper Release of Memory Before Removing Last Reference vulnerability in the Session Initiation Protocol (SIP) Application Layer Gateway (ALG) of Juniper Networks Junos OS allows unauthenticated network-based attacker to cause a partial Denial of Service (DoS). On all MX and SRX platforms, if the SIP ALG is enabled, receipt of a specific SIP packet will create a stale SIP entry. Sustained receipt of such packets will cause the SIP call table to eventually fill up and cause a DoS for all SIP traffic. The SIP call usage can be monitored by "show security alg sip calls". To be affected the SIP ALG needs to be enabled, either implicitly / by default or by way of configuration. Please verify on SRX with: user@host> show security alg status | match sip SIP : Enabled Please verify on MX whether the following is configured: [services ... rule <rule-name> (term <term-name>) from/match application/application-set <name>] where either a. name = junos-sip or an application or application-set refers to SIP: b. [applications application <name> application-protocol sip] or c. [applications application-set <name> application junos-sip] This issue affects Juniper Networks Junos OS on SRX Series and MX Series: 20.4 versions prior to 20.4R3-S2; 21.1 versions prior to 21.1R3-S2; 21.2 versions prior to 21.2R2-S2; 21.2 versions prior to 21.2R3; 21.3 versions prior to 21.3R2; 21.4 versions prior to 21.4R2. This issue does not affect Juniper Networks Junos OS versions prior to 20.4R1. Juniper SIRT is not aware of any malicious exploitation of this vulnerability.
CVE-2022-22198	An Access of Uninitialized Pointer vulnerability in the SIP ALG of Juniper Networks Junos OS allows an unauthenticated network-based attacker to cause a Denial of Service (DoS). Continued receipt of these specific packets will cause a sustained Denial of Service condition. On all MX and SRX platforms, if the SIP ALG is enabled, an MS-MPC or MS-MIC, or SPC will crash if it receives a SIP message with a specific contact header format. This issue affects Juniper Networks Junos OS on MX Series and SRX Series: 20.4 versions prior to 20.4R3; 21.1 versions prior to 21.1R2-S1, 21.1R3; 21.2 versions prior to 21.2R2. This issue does not affect versions prior to 20.4R1.
CVE-2022-22178	A Stack-based Buffer Overflow vulnerability in the flow processing daemon (flowd) of Juniper Networks Junos OS on MX Series and SRX series allows an unauthenticated networked attacker to cause a flowd crash and thereby a Denial of Service (DoS). Continued receipt of these specific packets will cause a sustained Denial of Service condition. This issue can be triggered by a specific Session Initiation Protocol (SIP) invite packet if the SIP ALG is enabled. Due to this, the PIC will be rebooted and all traffic that traverses the PIC will be dropped. This issue affects: Juniper Networks Junos OS 20.4 versions prior to 20.4R3-S2; 21.1 versions prior to 21.1R2-S1, 21.1R3; 21.2 versions prior to 21.2R2; 21.3 versions prior to 21.3R2. This issue does not affect Juniper Networks Junos OS versions prior to 20.4R1.
CVE-2022-22175	An Improper Locking vulnerability in the SIP ALG of Juniper Networks Junos OS on MX Series and SRX Series allows an unauthenticated networked attacker to cause a flowprocessing daemon (flowd) crash and thereby a Denial of Service (DoS). Continued receipt of these specific packets will cause a sustained Denial of Service condition. This issue can occur in a scenario where the SIP ALG is enabled and specific SIP messages are being processed simultaneously. This issue affects: Juniper Networks Junos OS on MX Series and SRX Series 20.4 versions prior to 20.4R3-S1; 21.1 versions prior to 21.1R2-S2, 21.1R3; 21.2 versions prior to 21.2R1-S2, 21.2R2; 21.3 versions prior to 21.3R1-S1, 21.3R2. This issue does not affect Juniper Networks Junos OS versions prior to 20.4R1.

https://cve.mitre.org/cgi-bin/cvekey.cgi?keyword=SIP

SIP should have anticipated NAT

- SIP and IPv6 evolved at roughly the same time
  - assumption: NATs = nuisance all temporary
- High-speed home access for VoIP didn't exist (~ 3-4% of US households)

```
Network Working Group                          S. Deering, Xerox PARC
Request for Comments: 1883                 R.  Hinden, Ipsilon Networks
Category: Standards Track                              December 1995
```

Internet Protocol, Version 6 (IPv6)
Specification

(Over 200 Kbps in at Least One Direction)

Types of Technology*	December 1999	June 2000	% Change
ADSL	369,792	950,590	157%
Other Wireline	609,909	747,028	22
Coaxial Cable	1,414,183	2,248,981	59
Fiber	312,204	307,151	n.m.
Satellite & Fixed Wireless	50,404	65,615	n.m.
Total Lines	2,756,492	4,319,365	57%

High-Speed Services for Internet Access: Subscribership as of June 30, 2000 (FCC)

SIP design: NATs continue to constrain

Lots of streams

- audio, video, screen sharing
- signaling, BFCP, XCON, …

Good for

- modularity (choose different protocols)
- QoS (apply different treatments)
- selective forwarding vs. mixing

NATs

- need to get public IP for each
- hard to re-use inbound connections

Firewalls

- hard to have port rules

Multiplexed media over HTTP (or TCP)

- ugly, but that's what Zoom does

What would a “SIP” version 3 look like?

HTTP/3-based → RIPT (*)

- different trade-off between standards and local software
- asymmetric (client-server)

Unclear whether current common SIP use cases would be significantly improved

- e.g., unlikely to achieve UE interoperability for complex video scenarios

(*) draft-rosenbergjennings-dispatch-ript-00

Shouldn't STIR/SHAKEN been done in 2002?

Yes, but RFC 4474 was published in 2006!

Just like for SSL/TLS (1995)

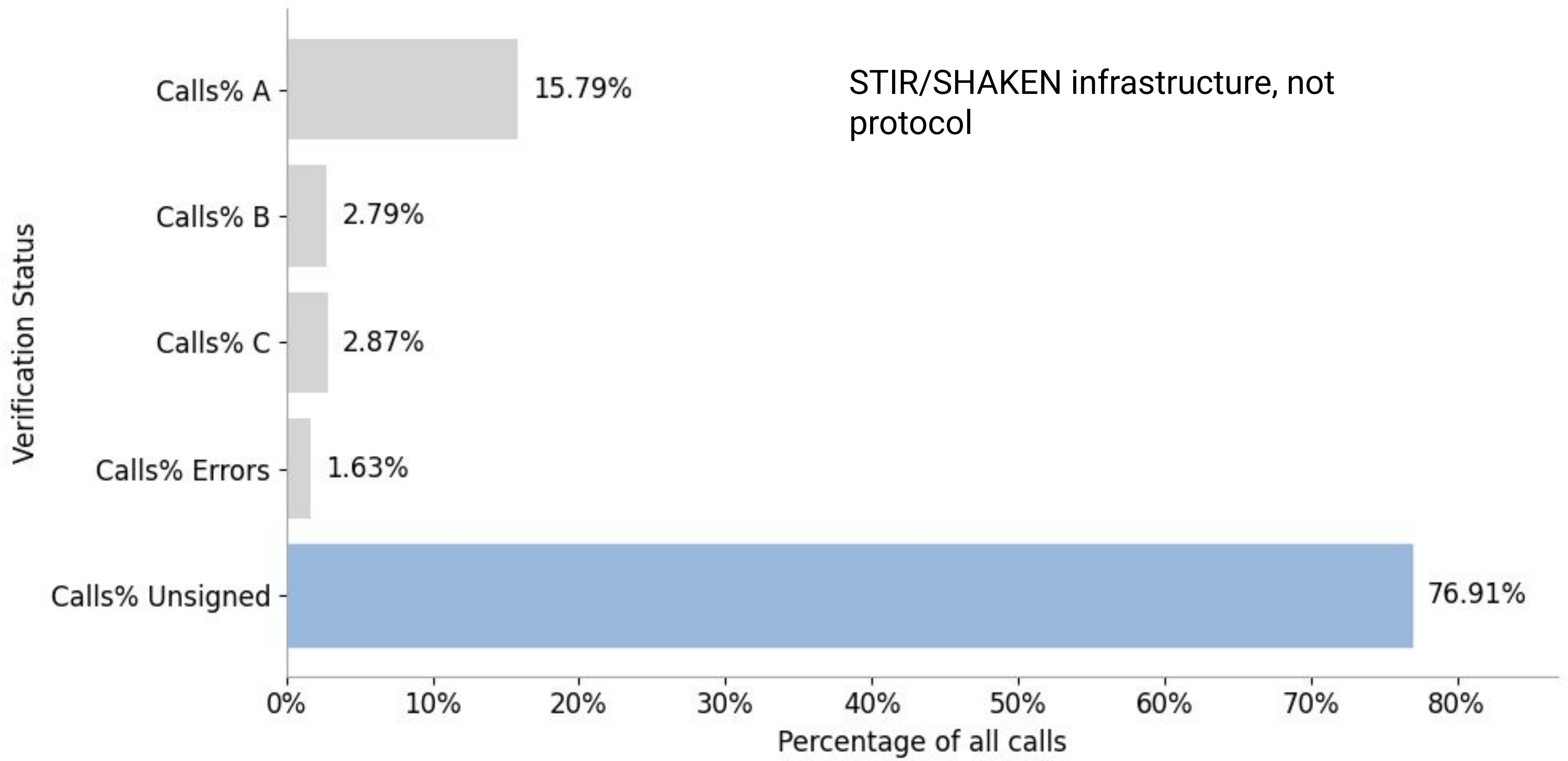
- in 2013, only 23% of European websites had encryption
- Let's Encrypt, AWS ACM, … probably mattered more than protocols

```
Network Working Group                                       J. Peterson
Request for Comments: 4474                                      NeuStar
Category: Standards Track                                   C. Jennings
                                                          Cisco Systems
                                                            August 2006


   Enhancements for Authenticated Identity Management in the
              Session Initiation Protocol (SIP)
```

```
INVITE sip:bob@biloxi.example.org SIP/2.0
Via: SIP/2.0/TLS pc33.atlanta.example.com;branch=z9hG4bKnashds8
To: Bob <sip:bob@biloxi.example.org>
From: Alice <sip:alice@atlanta.example.com>;tag=1928301774
Call-ID: a84b4c76e66710
CSeq: 314159 INVITE
Max-Forwards: 70
Date: Thu, 21 Feb 2002 13:02:03 GMT
Contact: <sip:alice@pc33.atlanta.example.com>
Identity:
  "ZYNBbHC00VMZr2kZt6VmCvPonWJMGvQTBDqghoWeLxJfzB2a1pxAr3VgrB0SsSAa
   ifsRdiOPoQZYOy2wrVghuhcsMbHWUSFxI6p6q5TOQXHMmz6uEo3svJsSH49thyGn
   FVcnyaZ++yRlBYYQTLqWzJ+KVhPKbfU/pryhVn9Yc6U="
Identity-Info: <https://atlanta.example.com/atlanta.cer>;alg=rsa-sha1
Content-Type: application/sdp
Content-Length: 147
```

https://doi.org/10.1145/2872518.2888605

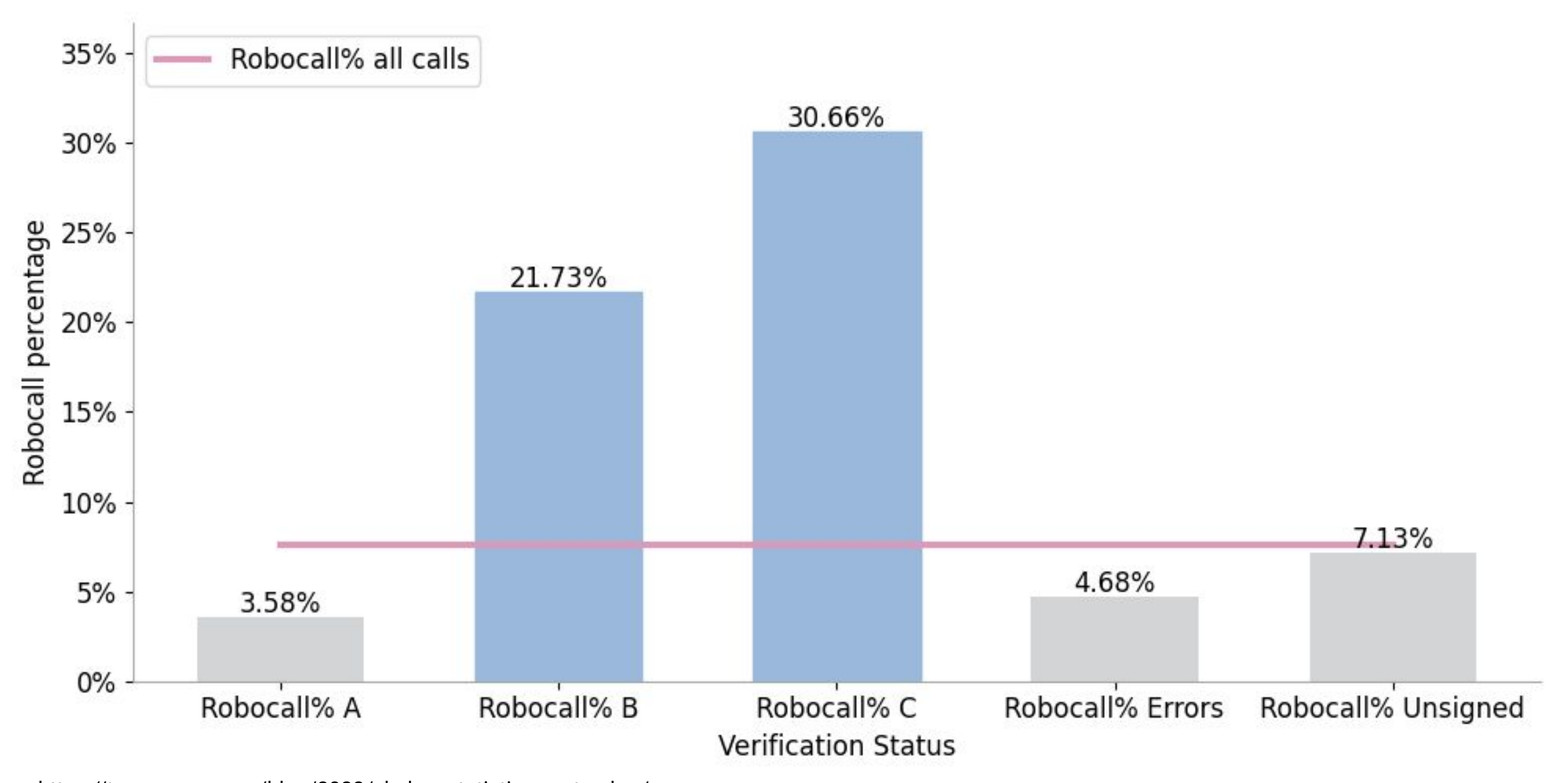
Regulatory push & pull likely matter more than technology



STIR/SHAKEN infrastructure, not protocol

https://transnexus.com/blog/2022/shaken-statistics-september/

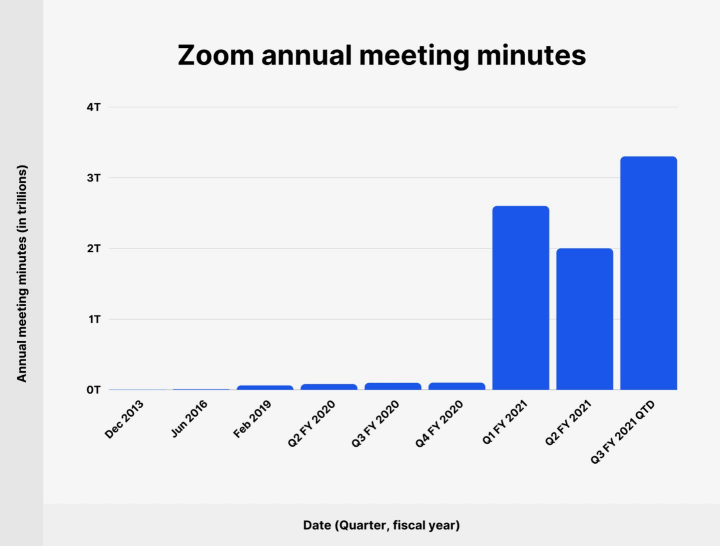
Signing is a very good predictor of — robocalls



https://transnexus.com/blog/2022/shaken-statistics-september/

Why did standards fail for video conferencing (mostly)?

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

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)

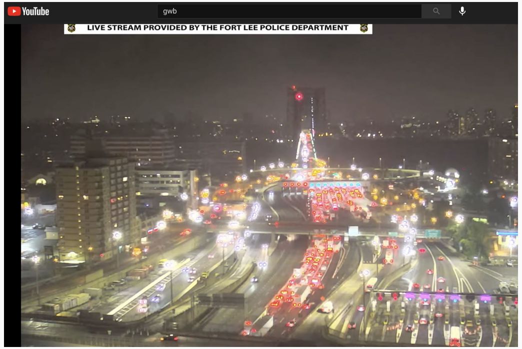
The landscape of IP video communications

Unified Communications (UC)

2-party phone call, spontaneous

CuSeeMe (1992)

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

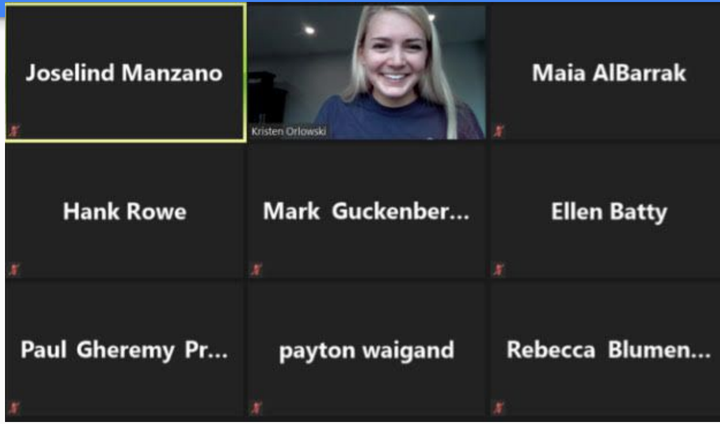


Multi-party streaming (Mbone, YouTube, FB Live, Livestream)

- 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 invite 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

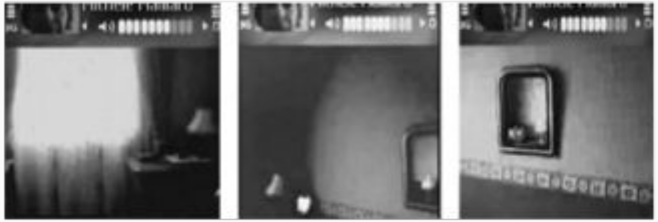
Mundane Video Directors in Interaction: Showing One's Environment in Skype and Mobile Video Calls

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

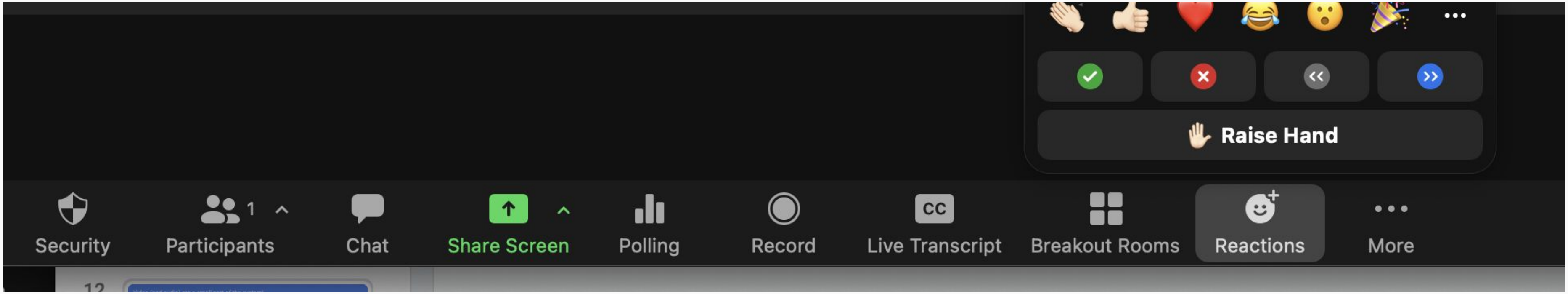
How to make the most of NYC apartment tours via FaceTime and Zoom

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



but also a reason people do audio-only Zoom calls!

Aside: What's wrong with the Zoom video model?

See Jeremy N. Bailenson (Feb. 2021):

- Eye gaze at a close distance (cf. elevator gaze aversion) - no zoom on Zoom!
  - "long stretches of direct eye gaze and faces seen close up" (~50 cm)
  - for mid-sized meetings, everybody looks at every other non-speaker
- Constant self-monitoring
  - "centering oneself in the camera's field of view, nodding in an exaggerated way for a few extra seconds to signal agreement, or looking directly into the camera")
  - side glances are misconstrued
- All day mirror
- Reduced mobility
  - moving out of camera view is seen as sign of non-attention
- VR may make this worse, e.g., by confusing positional cues
  - who "sits" where? How do I see the person's face if covered by VR goggles?

https://tmb.apaopen.org/pub/nonverbal-overload/release/2

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

"API": set up sessions, functions
XCON (RFC 5239+)

Media quality feedback
(RTCP)

Bandwidth adjustment

Packet loss recovery

Call signaling to hardware
(SIP, H.323)

Call signaling to software

Floor control (hand-raising, muting others, ...)
(BFCP)

Media transport
(RTP)

H.264
MPEG-4/AVC

Echo cancellation

G.711 (4 kHz)
G.722 (8 kHz)
OPUS (HQ)
G.723, G.729, iLBC, ...

Text chat (incl. reactions)
(RCS, T.140, MSRP, XMPP)

Screen sharing

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

Interoperability: indifferent, cooperative, competitive

[Doctorow, CACM 10/2021]

- Indifferent interoperability
  - company A does not care that B makes a complementary product
- Cooperative interoperability
  - typically via standards
  - but may play favorites
- Competitive (or adversarial) interoperability
  - “third-party inkjet ink, DVRs that record anything”
  - see copyright-for-API (Google vs. Oracle)

When do we get standards

Condition	VoIP
Connect different industries	PBX + carriers; mobile + landline; device + carrier
Industries with different emphasis	Hardware (incl. niches) vs. software vs. operations
Non-dominance of single vendor or operator	lots of local, niche & national carriers (unlike browser)
Minimize interconnection preparation	don’t want to install new software (with new UI) for each call
Interoperability with legacy technology	150 years: analog, SS7, ISDN

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

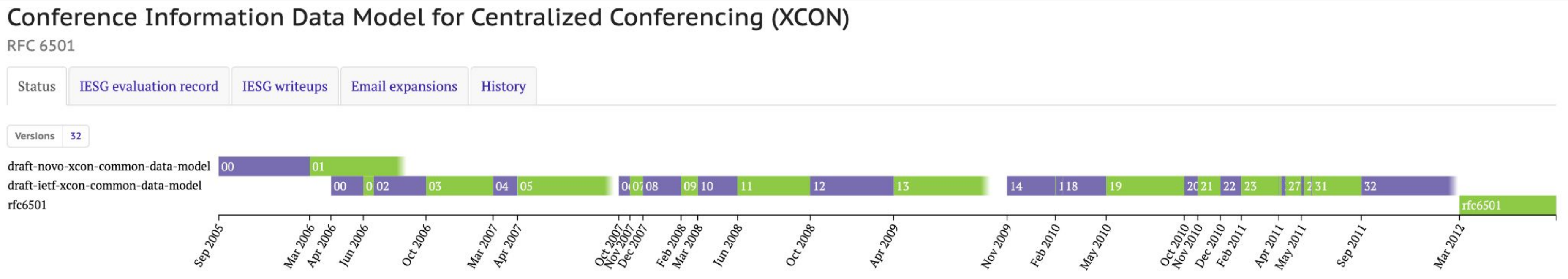
SDP-Related Documents

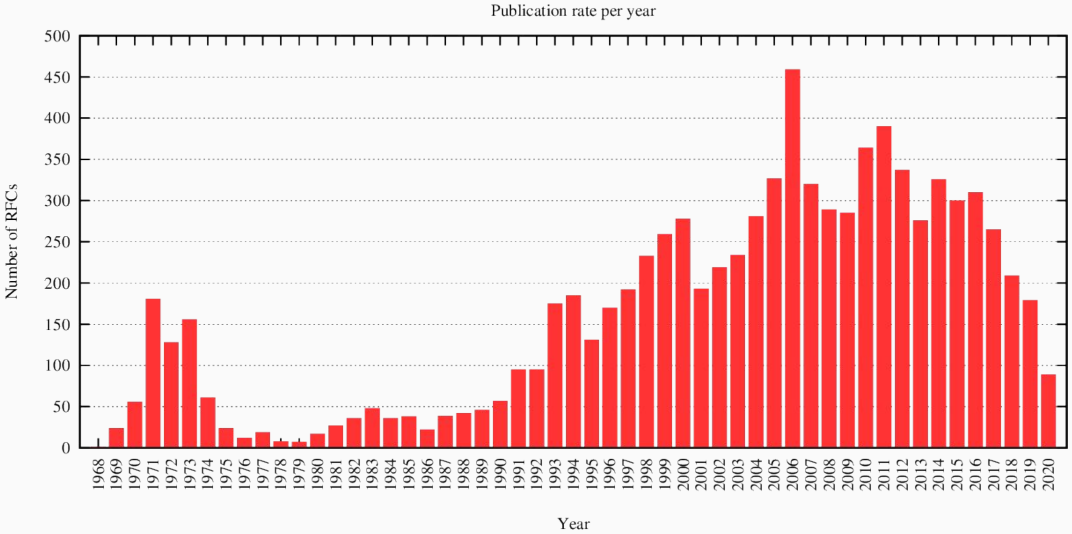
RFC	Document Title
RFC 2327	Session Description Protocol (SDP) (obsolete: see RFC 4566)
RFC 3264	An Offer/Answer Model with the Session Description Protocol (SDP)
RFC 3266	Support of IPv6 in SDP
RFC 3388	Grouping Media Lines in SDP (obsolete: see |RFC 5888|)
RFC 3407	Session Description Protocol (SDP) 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)

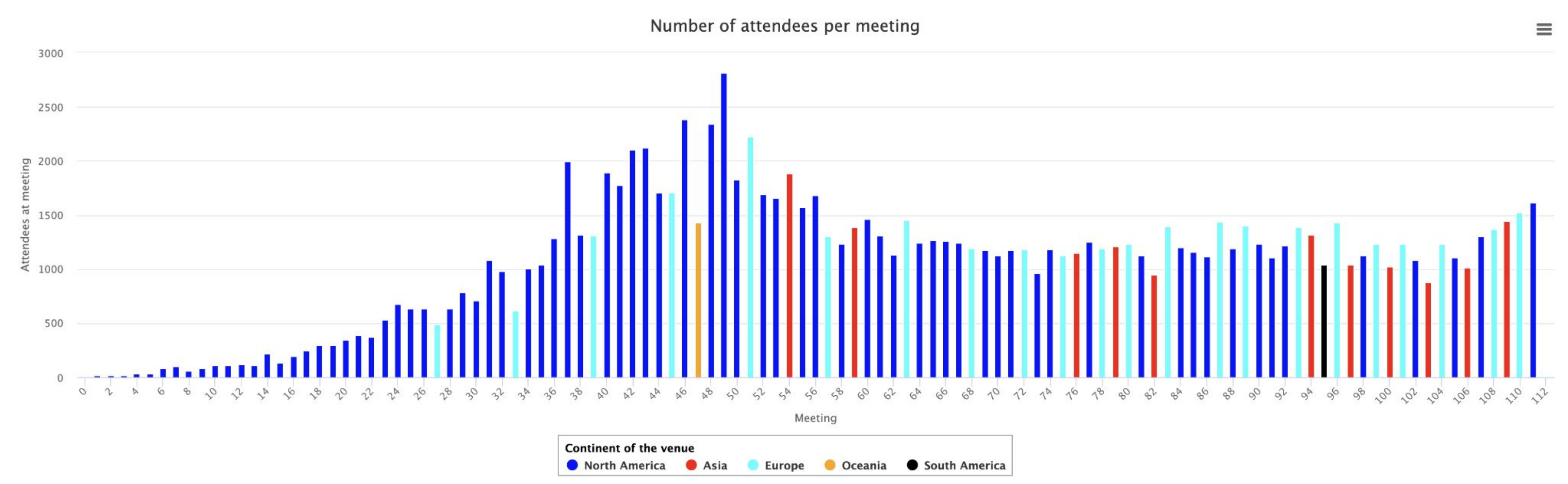
IMS 23.228: 329 pg.
RCS 5.1: 482 pg.

Standards can take a looong time (and RFCs are decreasing)





It takes a lot of people to do the work



Simple core protocols have acquired technical debts

RFC	Type	Status	Title	Bgnd	Prot	Names	Ops	RR	Proxy	Stub	Auth	Res	X		SSEC
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:

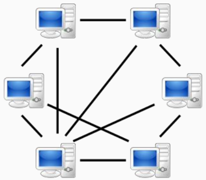
- Attendee management
- Connect to YouTube, Facebook Live, ...
- Monetization
- Polling and "engagement"

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

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

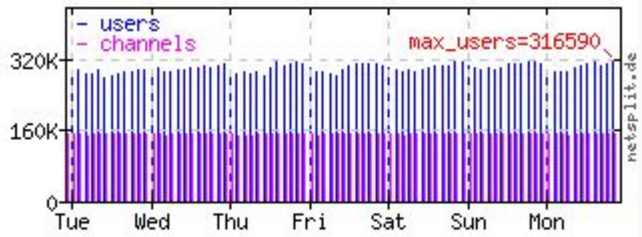
32

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

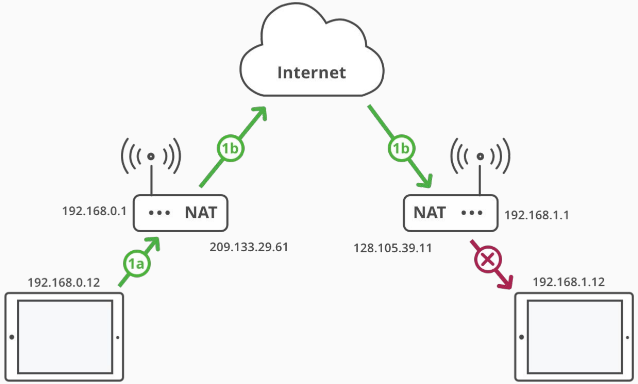
NATs killed the peer-to-peer model

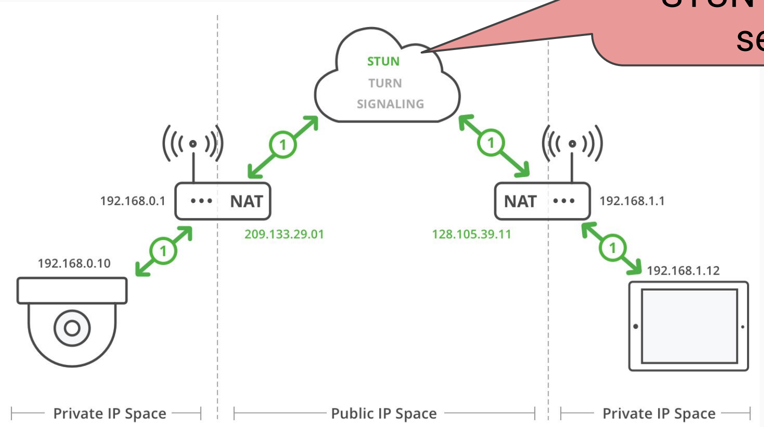
VoIP clients need inbound connections for call signaling and media

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

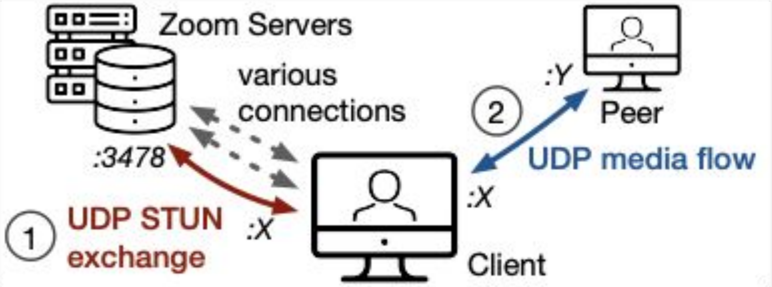
somebody has to provide the STUN and TURN servers





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

But not quite - Zoom uses P2P for two-party calls



Oliver Michel, Satadal Sengupta, Hyojoon Kim, Ravi Netravali, and Jennifer Rexford. 2022. Enabling Passive Measurement of Zoom Performance in Production Networks. In Proceedings of the 22nd ACM Internet Measurement Conference (IMC ’22)

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

?

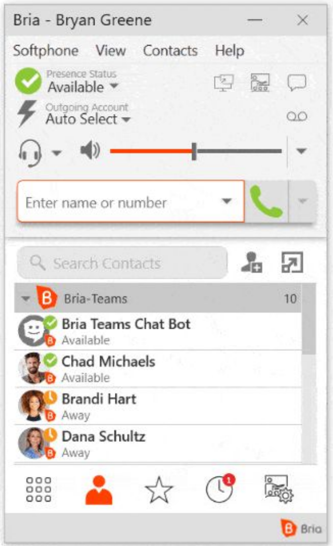
Project Name	Platforms
Aparté	BSD / Linux / macOS
AstraChat	Android / iOS / Linux / macOS / Windows
BeagleIM by Tigase, Inc.	macOS
blabber.im	Android
Bruno the Jabber™ Bear	Android
Conversations	Android
Converse	Browser
Dino	Linux
Gajim	Linux / Windows
Kaidan	Android / Linux / macOS / Other / Windows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

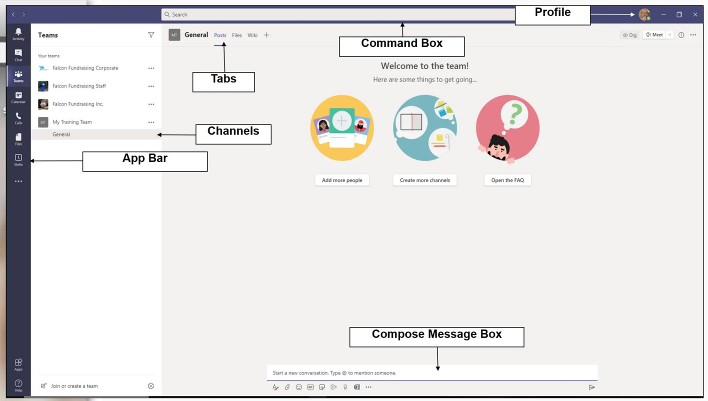
WebRTC as transition model

Standards-based client

WebRTC client

Application



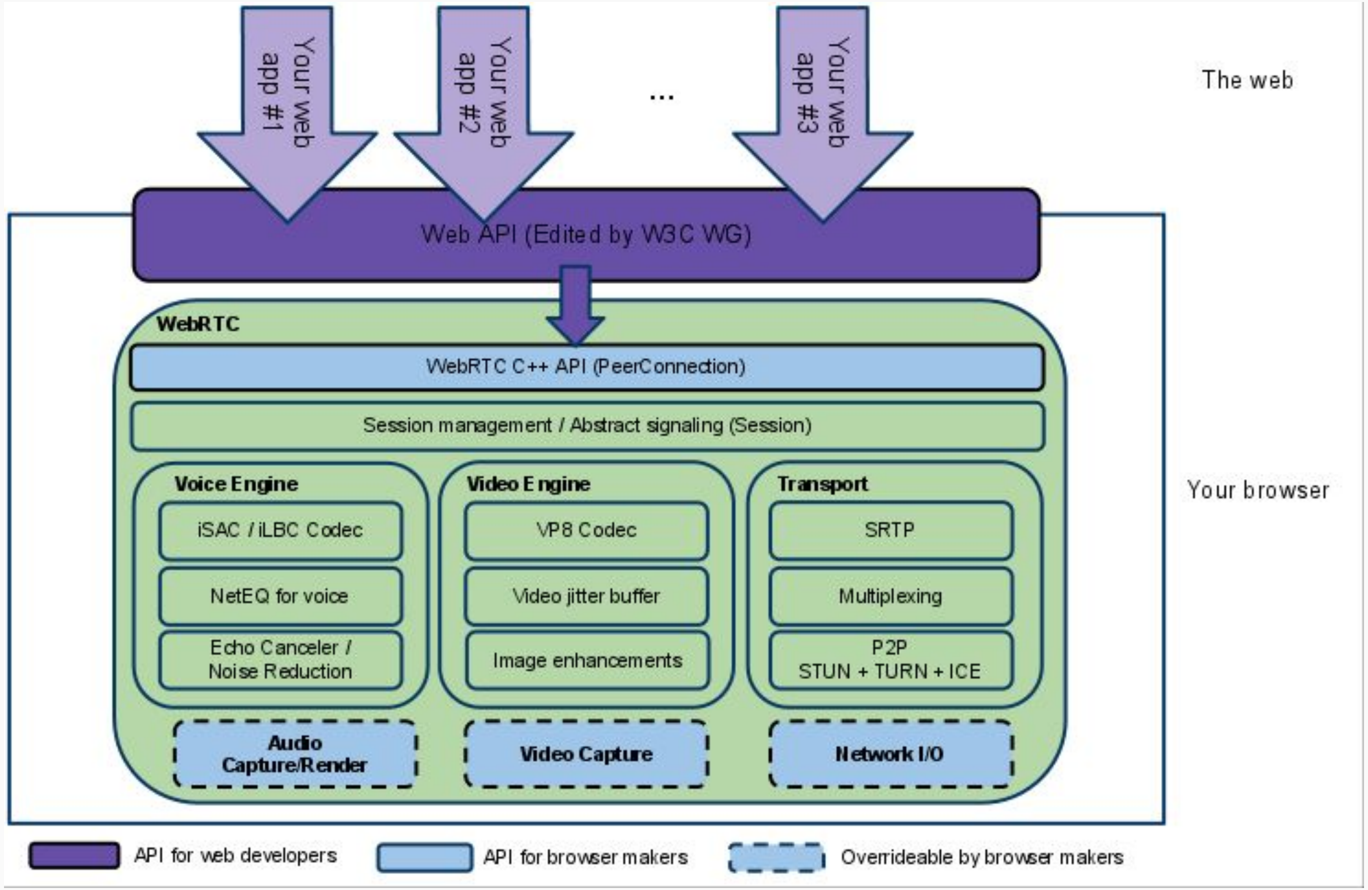


multiple services, one client

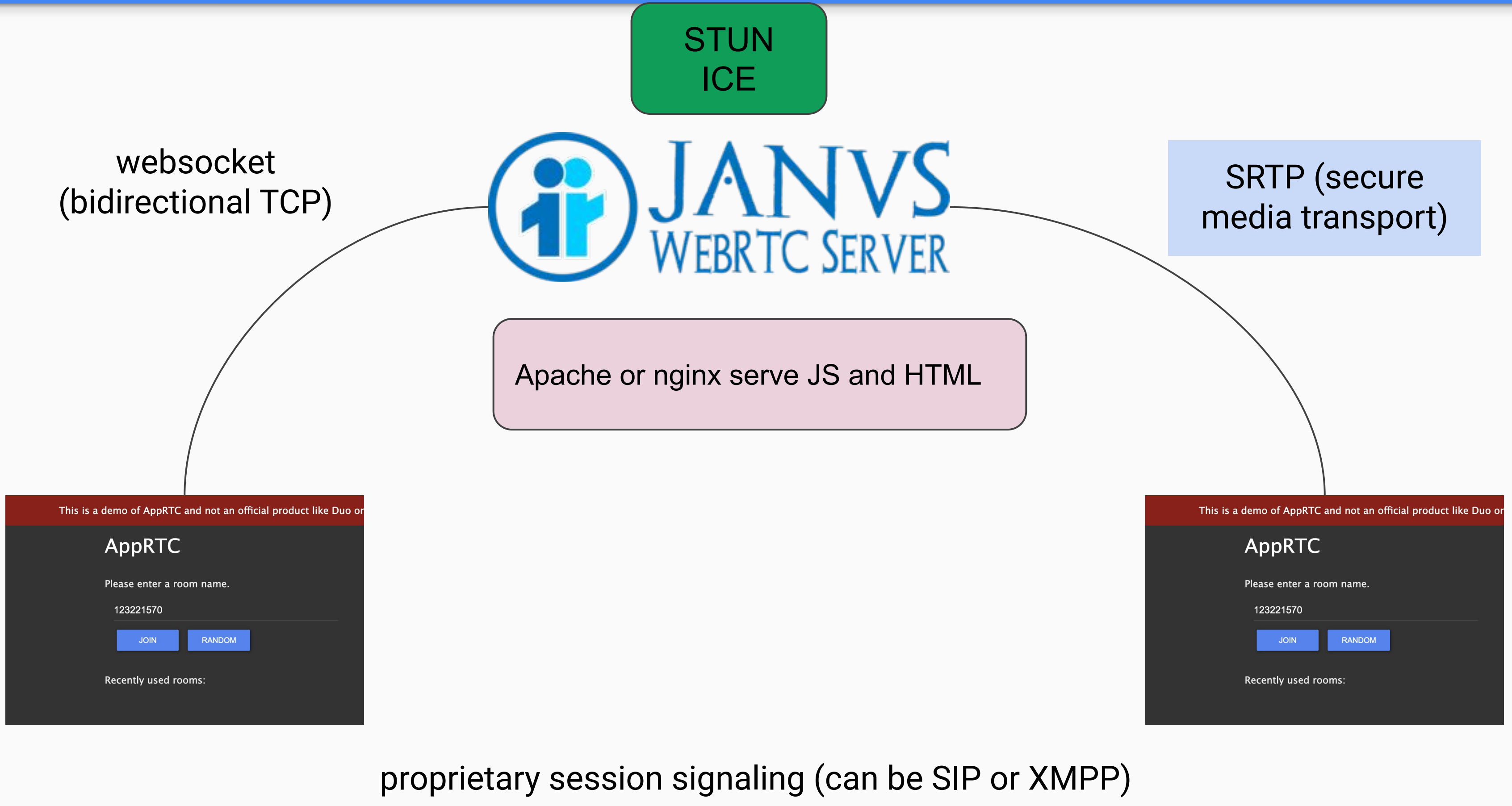
no installation - one “page” per service
switch browsers & maybe platforms
no interoperability between services

No interoperability between services

WebRTC architecture

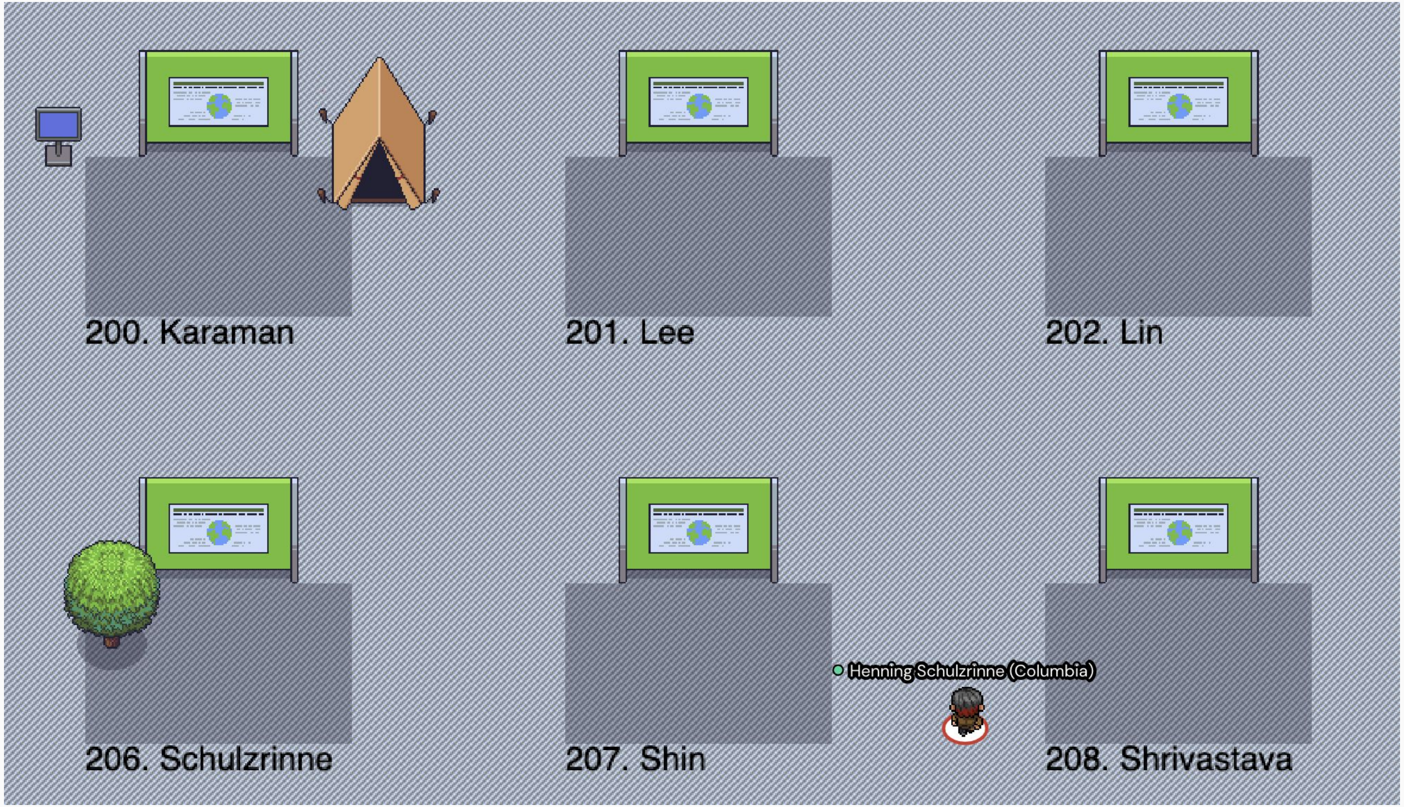


Typical WebRTC architecture



Good for non-square UIs

gather.town



advantages to break-out rooms?

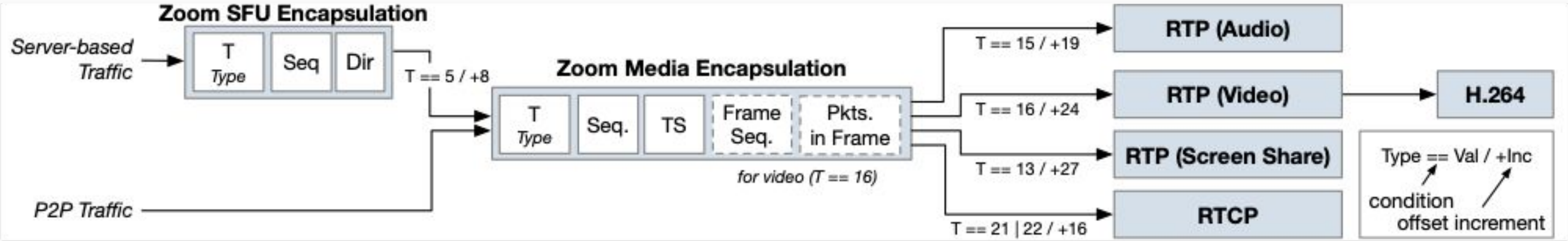
Or lower level still - browser as VM

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

WebTransport: multiple cancellable streams: datagrams + bidirectional reliable streams

WebCodecs API: direct access to codecs

Zoom: vestigial standards compliance



Value	Packet Type	Offset
16	RTP: Video	24
15	RTP: Audio	19
13	RTP: Screen Share	27
34	RTCP: SR + SDES	16
33	RTCP: SR	16

makes it easier to interoperate with SIP and H.323 room systems!

Oliver Michel, Satadal Sengupta, Hyojoon Kim, Ravi Netravali, and Jennifer Rexford. 2022. Enabling Passive Measurement of Zoom Performance in Production Networks. In Proceedings of the 22nd ACM Internet Measurement Conference (IMC '22)

Partial standards re-use is common

TABLE II: Comparison of the RTC applications under test. Under *Redundant data*, "F" stands for FEC and "S" for Simulcast. Under *DNS domains*, "B" stands for easy to block, "C" for company-specific and "S" for social networks. Under *Other*, "N" means it uses less than four server-side ports and "T" means that PTs are used in a static fashion.

Application	Protocols				Operation			Identification		
	RTP	STUN/TURN	DTLS	Other	P2P	Redundant Data	Other	Own AS	DNS Domains	Other
Skype	✓	✓		✓	✓	F,S		✓	B	N,T
Google Meet	✓	✓	✓			S	✓	✓	C	N,T
Jitsi Meet	✓	✓	✓		✓				B	
WhatsApp	✓	✓			✓	F		✓	B	N,T
Telegram		✓		✓	✓			✓	B	
Facebook Messenger	✓	✓	✓		✓			✓	S	T
Instagram Messenger	✓	✓						✓	S	N,T
Facetime	✓	✓			✓		✓	✓	C	N,T
HouseParty	✓	✓	✓						B	T
Microsoft Teams	✓	✓		✓	✓	F,S		✓	B	N,T
Webex Teams	✓	✓				F,S	✓	✓	B	N
Zoom	✓			✓	✓	F			B	N,T
GoTo Meeting				✓					B	N

A. Nistico et al, *A comparative study of RTC applications*, IEEE ISM 2020

GoToMeeting = AVTP (IEEE 1722-2011)

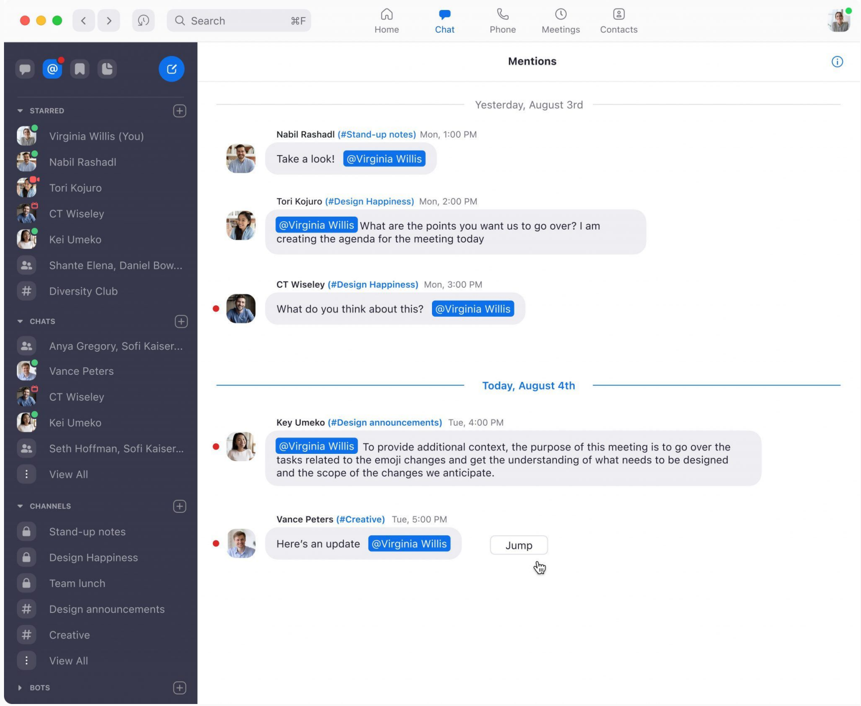
Bifurcation

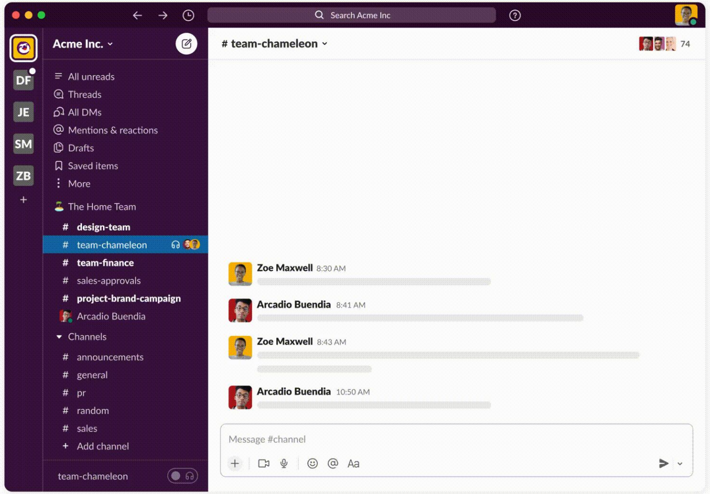
Communication out front applications: collaboration, social interaction, telemedicine

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

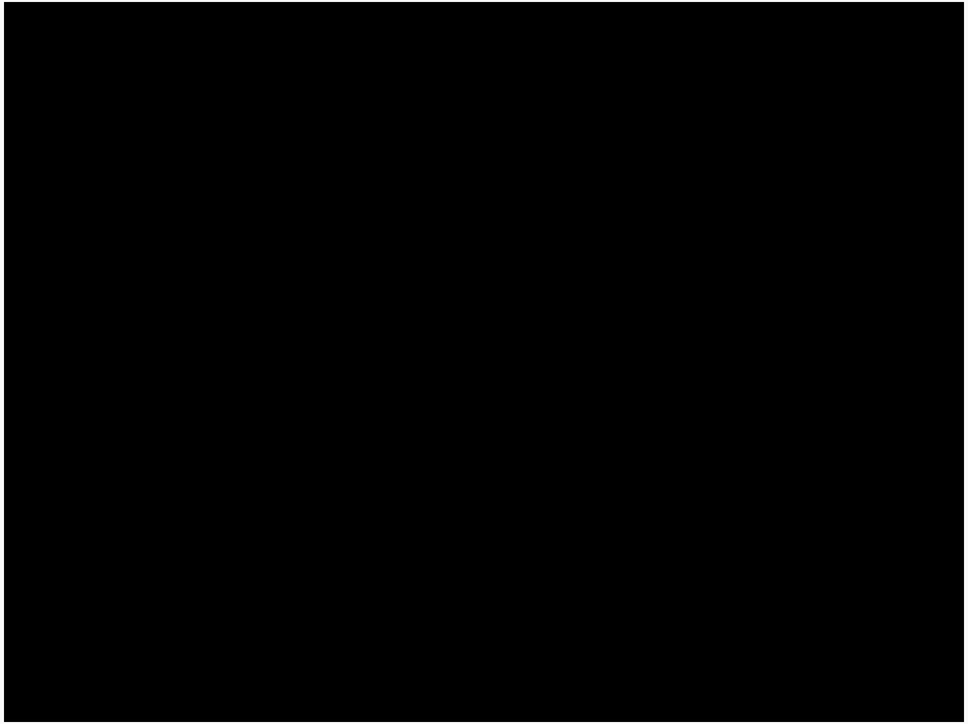
Video in back applications: monitoring (traffic, agriculture, security, …) → consumers are ML applications

The uneasy coexistence of synchronous and asynchronous collaboration





Or maybe we’ll just be avatars



https://www.meta.com/work/workrooms/

And the typical group project has…

slack

zoom

Lucidspark

Groups.io

box

SLIMWIKI

each with their own login, groups, privileges, …

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)