The Wiretap Act Meets the Internet

Joint work with Matt Blaze, Susan Landau, and Stephanie Pell

Steven M. Bellovin
https://www.cs.columbia.edu/~smb

cc BY NC

The Internet is for *Katz*?

Katz v. United States, 389 U.S. 347 (1967)

- Held that wiretapping was a search under the Fourth Amendment, and hence required a warrant
  - Overruled *Olmstead v. United States* (1928)
- “The Fourth Amendment protects people, not places”
- (*Katz* was actually about bugging, not wiretaps)

Smith v. Maryland, 442 U.S. 735 (1979)

- “Pen registers” are not searches and do not require warrants
- (What’s a pen register?)
  - Records dialed numbers; cannot pick up voice
- One of the seminal cases for the “third party doctrine”
- “All telephone users realize that they must ‘convey’ phone numbers to the telephone company, since it is through telephone company switching equipment that their calls are completed.”

Today’s Status

- *Katz* led to the 1968 passage of the Wiretap Act (18 U.S.C. §2510 *et seq.*)
- The Wiretap Act was amended in 1986 in the *Electronic Communications Privacy Act (ECPA)*, which extended it to cover data
    - (Also covered stored data, such as email not in transit)
- ECPA also provided for pen registers (18 U.S.C. §3121 *et seq.*)

What is the Difference Between *Katz* and *Smith*?

- *Katz*: conversations between two people are private and hence protected
- *Smith*: you have no privacy interest in information you voluntarily give to a third party. i.e., the phone company
    - When you dial a call, you are talking to the phone company
- Rough intuition: if you speak a different language, the phone call will work. If you send different tones when dialing—and remember, these were landlines, where touching digits sent audible tones—the call will not work
- Voice is *end-to-end*; dialing is not
- How does this work on the Internet?

Katz, *Smith*, and the Internet

- We need to understand what is end-to-end communication and what information (“metadata”) is given to third parties

- It seemed obvious for phone calls

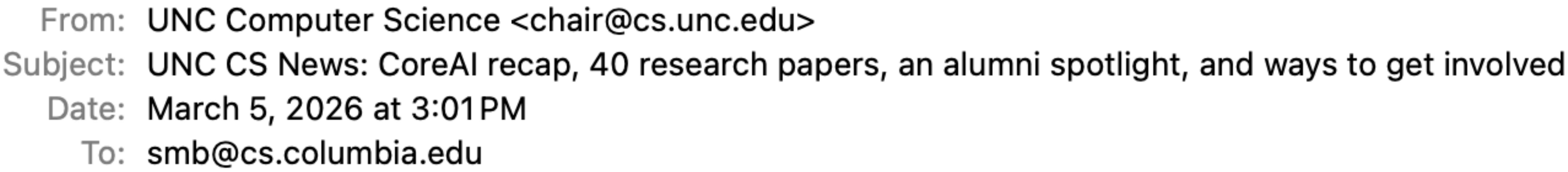
- You can’t understand this for the Internet without knowing how the Internet works

- DoJ and courts have gotten this wrong

    - (Actually, though the *Smith* court didn’t realize it, almost contemporaneous with the decision the phone network got significantly more complex)

A Simple(?) Case: Email

- We all use email, and we’ve all seen standard email and its headers: `From:`, `To:`, `Cc:` `Date:`, `Subject:`, `Bcc:`, plus of course the body of the message
- What is end-to-end and what is *voluntarily* given to a third party?

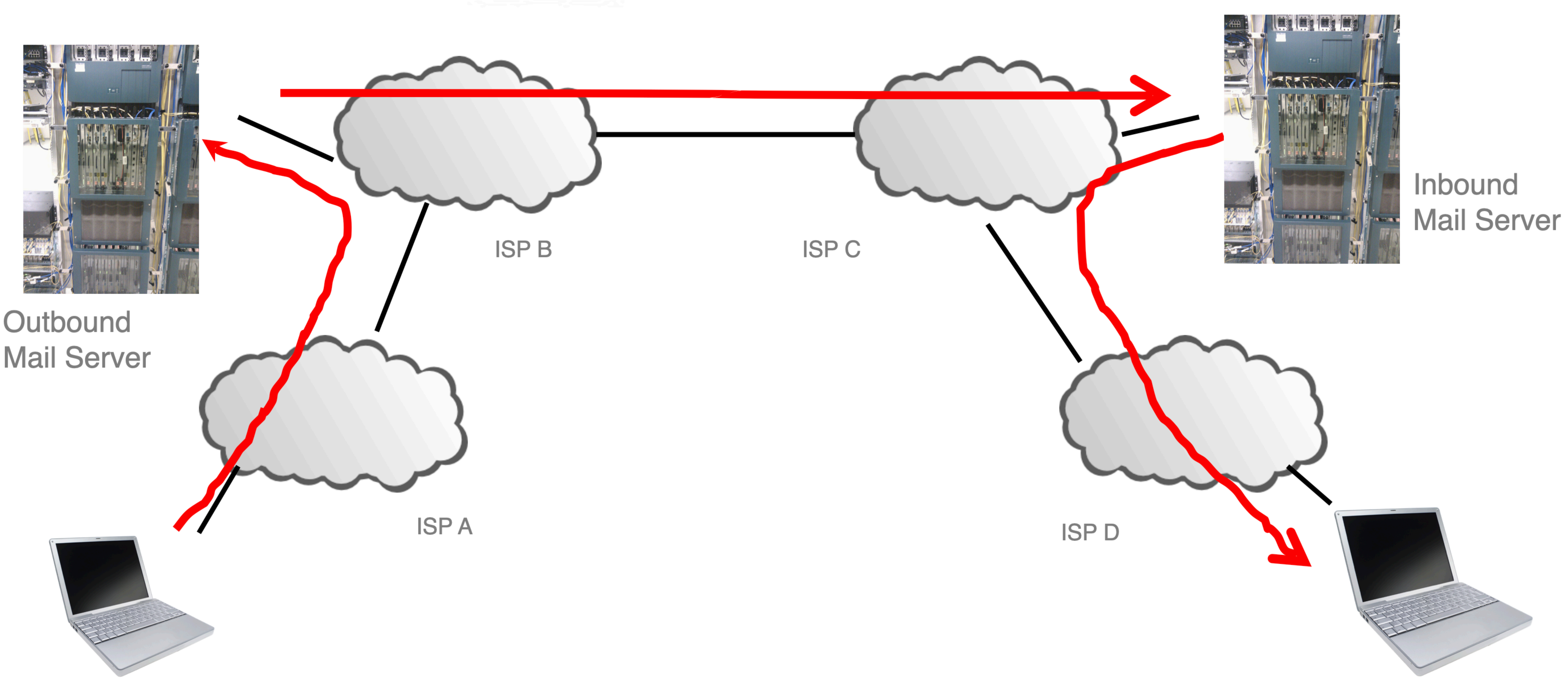


From: UNC Computer Science <chair@cs.unc.edu>
Subject: UNC CS News: CoreAI recap, 40 research papers, an alumni spotlight, and ways to get involved
Date: March 5, 2026 at 3:01 PM
To: smb@cs.columbia.edu

Answer: *All* of that is End-to-End

- Yes, the mail system needs to know to whom to deliver the message
- Yes, *usually* the address information in the header lines is used to create the delivery information
- But it doesn’t have to be!

Sending and Reading Email



Email (Simplified)

- Mail goes from a sender’s device to an “outbound mail server”
- From there, it is sent to the recipient’s “inbound mail server”
- The recipient downloads it from that machine
- The mail servers are generally ISP- or enterprise-operated

Sending Myself Email: SMTP

```
220 machshav.com ESMTP Exim 4.82 Tue, 11 Mar 2014 19:43:03 +0000
HELO eloi.cs.columbia.edu
250 machshav.com Hello eloi.cs.columbia.edu [2001:18d8:ffff:16:12dd:b1ff:feef:8868]
MAIL FROM:<smb@eloi.cs.columbia.edu>
250 OK
RCPT TO:<smb@machshav.com>
250 Accepted
DATA
354 Enter message, ending with "." on a line by itself
From: Barack Obama <president@whitehouse.gov>
To: <smb2132@columbia.edu>
Subject: Test

This is a test
.
250 OK id=1WNSaS-0001z5-1d
QUIT
221 machshav.com closing connection
```

Conversation with a Third Party

```
220 machshav.com ESMTP Exim 4.82 Tue, 11 Mar 2014 19:43:03 +0000
HELO eloi.cs.columbia.edu
250 machshav.com Hello eloi.cs.columbia.edu [2001:18d8:ffff:16:12dd:b1ff:feef:8868]
MAIL FROM:<smb@eloi.cs.columbia.edu>
250 OK
RCPT TO:<smb@machshav.com>
250 Accepted
DATA
354 Enter message, ending with "." on a line by itself
[redacted]
.
250 OK id=1WNSaS-0001z5-1d
QUIT
221 machshav.com closing connection
```

What the Recipient Sees

[redacted]

From: Barack Obama <president@whitehouse.gov>
To: <smb2132@columbia.edu>
Subject: Test

This is a test

[redacted]

Things to Note

- The SMTP envelope—that's the technical term!—can have different information than the message headers
- Unlike the phone network, anyone can run their own mail servers
  - I personally run two, one personal and one professional
  - This complicates third party doctrine analysis
- The reality of email is far more complex than I've outlined here
  - Example: many people read their email via a Web browser—and the NSA has stated that even for them, picking out just the From/To information from a Webmail session is very difficult

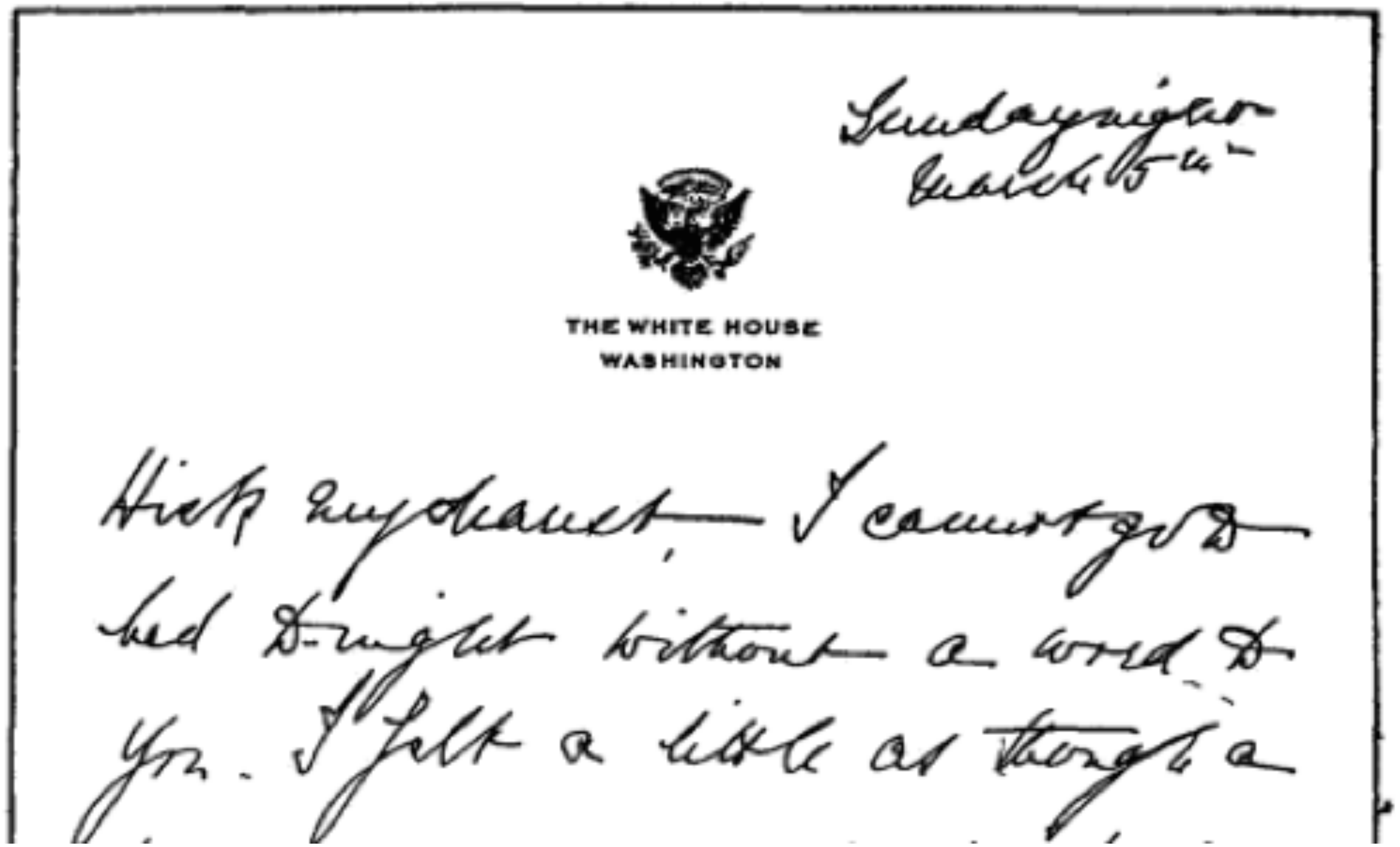
Two Sets of Addresses

- The header From/To lines are not the same as the envelope From/To lines
- But courts and the Justice Department have gotten it wrong:
  - DoJ: “Pen register and trap and trace devices may obtain any non-content information . . . Such information includes … the ‘To’ and ‘From’ information contained in an e-mail header” (*Electronic Surveillance Manual)*
  - Court: “That portion of the “header” which … reveals the e-mail addresses … would certainly be obtainable using a pen register and/or a trap and trace device.” (*In re Application of United States*, 396 F. Supp. 2d 45, 48 (D. Mass. 2005))

It’s Worse than That

- Two of the four authors of this article run their own mail servers
- Mail from me to one of the others has *no* third party involved, even for envelope addresses
- Mail from me to one of the other two does have a third party
- You can’t tell until after you’ve intercepted the SMTP dialog if it was legal to do so!

A Letter from Eleanor Roosevelt to Lorena Hickock (March 1933)



Sunday night
March 5th

THE WHITE HOUSE
WASHINGTON

Hick my dearest, — I cannot go to bed tonight without a word to you. I felt a little as though a

The Web and URLs

- URLs seem simple:

  https://en.wikipedia.org/wiki/Metadata

- There's an "authority" (en.wikipedia.org), which seems to be retrievable by pen registers, and there's a "path" (/wiki/Metadata), which seems to be content

- It's not that simple…

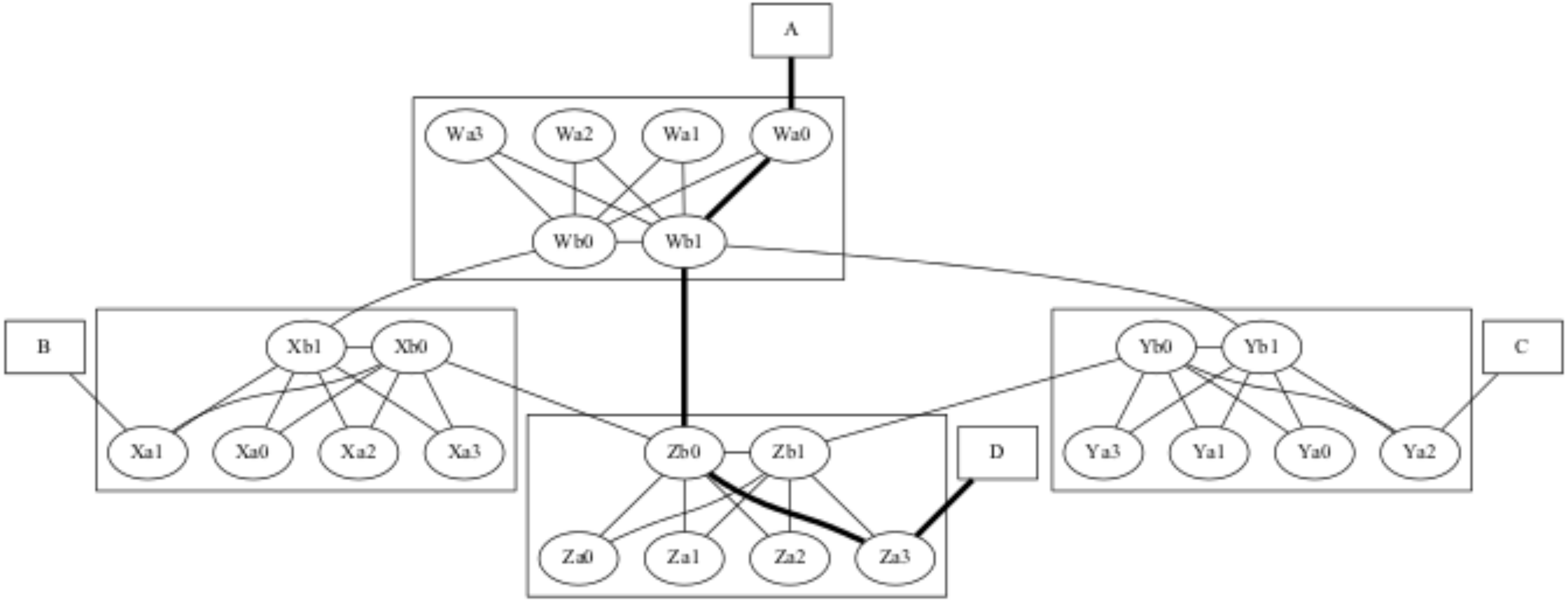
URL Complexities

- If a server hosts multiple web sites, the IP address points to the server, but the actual web name site goes to the web site operator, who may own all of the sites (i.e., there’s no third party)
    - But there are hosting services where the operator does not own the web sites
- Is there a legal difference between patents.google.com and www.google.com/patents ?
- When clicking on a Google search result, the real path may go to Google first—a third party!—but the user doesn’t know this
- The user cannot tell whether or not there are third parties involved, so the data is not voluntarily given

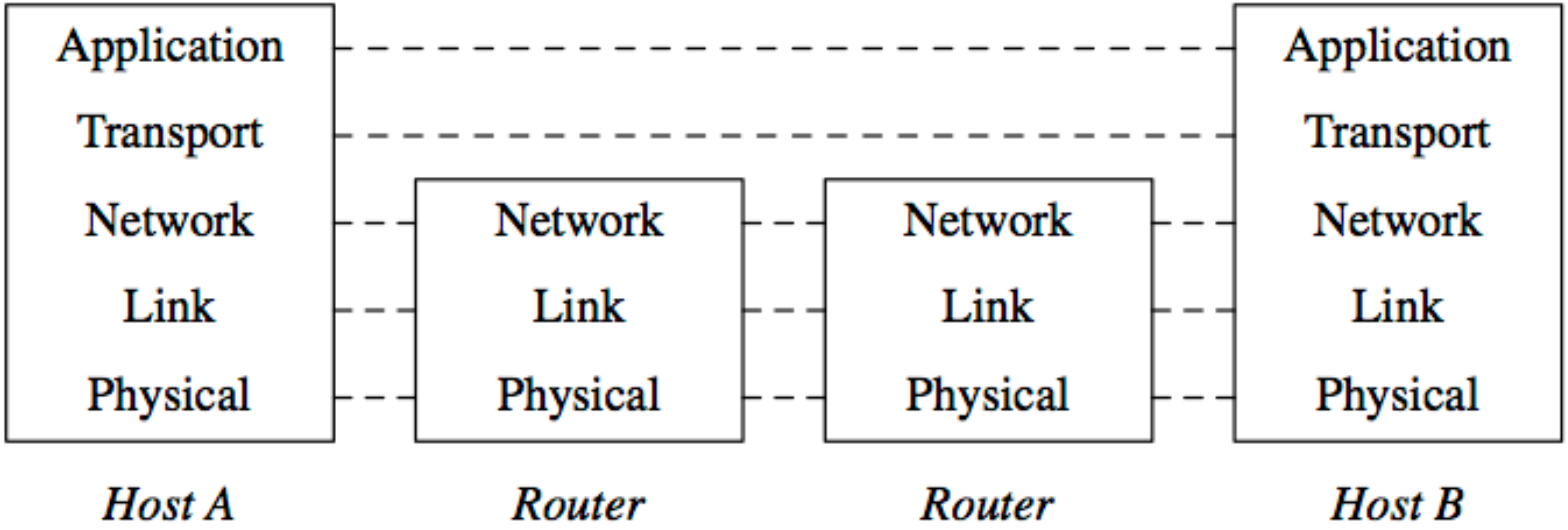
Voice over IP

- Most phone calls today are actually carried via Internet technology
- *Ex Parte Jackson* (96 U.S. 727 (1878)) said that the “outward form and weight” of a letter or package was not protected
- White et al. showed that they could use packet lengths of encrypted VoIP conversations to recover some phrases
- Metadata now reveals content!

Internet Routing



The Internet: A Layered Architecture



Third Parties?

- The transport and application layers are “end-to-end”; information in them is not given to a third party
- The network layer has lots of third parties
- The link layer is (often) local, so the other end-point is typically another device of yours
- It’s not so simple…

Link Layer Addresses

- Link layer addresses (e.g., WiFi MAC addresses) are the kind of thing that pen registers collect
- If the device is used at home these addresses are generally not given to third parties, and hence cannot constitutionally be collected with a pen register order
- But: if used on a public WiFi network, the hotspot operator is a third party
- What if your cable modem is owned and operated by your ISP? Third party?

IP Addresses

- Clearly given to third parties: intermediate routers along the path
- Also: clearly phone number-like information
- But—do most individuals know their computers have IP addresses?
    - In Smith, the court noted how much consumers would know from phone bills, phone books, popular culture, etc.
    - Have you ever gotten a bill itemizing IP addresses you connected to?
- Is the conveyance *voluntary?*

Transport Layer “Port Numbers"

- If an IP address is like a building’s street address, the port number is the room within the building
- Different ports are used for different services: port 25 is for receiving email, ports 80 and 443 are web servers, etc.
- In other words, a port number is a service address; seeing the port number (often) says why someone is contacting another computer

Port Numbers

- They’re part of the transport layer, so there are no third parties involved
- Or are there?
  - Mobile phone users’ connections go through “carrier-grade NAT” (Network Address Translation), which uses port numbers
  - (On the Columbia University campus, there is no NAT—but in the law school, there is)
  - ISPs monitor and sometimes block based on port numbers
- But—do ordinary users know any of that? Is the conveyance “voluntary”?
- Port numbers are taken, not given!

Signaling

- Signaling is the exchange of messages that set up a connection, and is covered by the pen register statute
- On the phone network, this is done by phone switches operated by the telephone company
- Internet signaling is done by the transport layer—and transport is end-to-end, with no third parties involved
- But: NATs used on mobile phones do look at the TCP signaling fields. Voluntary and knowing?

The Internet is Not the Phone Network

- The Internet works differently than the phone network
- The legal standards were created for a much simpler network
- Some information is clearly third party data per Smith—but other information is much harder to classify as content or metadata.
- It is very hard to use “voluntariness” as a touchstone—and Smith relied on voluntariness
- The content/non-content distinction and the third party doctrine are no longer workable rules for an IP-based communications environment.
- We need new constitutional and statutory frameworks to govern law enforcement access to wire and electronic communications data.

The Gory Details

It’s Too Complicated: How the Internet Upends *Katz*, *Smith*, and Electronic Surveillance Law

Steven M. Bellovin, Matt Blaze, Susan Landau, and Stephanie K. Pell

30 Harvard J. of Law and Technology 1

https://jolt.law.harvard.edu/assets/articlePDFs/v30/30HarvJLTech1.pdf

Questions?

Common raven, New York City, March 2024