Cloud Computing and Global Communications

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

Cloud Computing

What's a Cloud?

- A cloud is a traditional way to represent a network
- This "three-cloud network" picture is from 1982
- But—today "cloud" refers to computing services provided via the Internet by an outside party.
- (The modern usage seems to date to 1996: http://www.technologyreview.com/news/425970/who-coined-cloud-computing/)

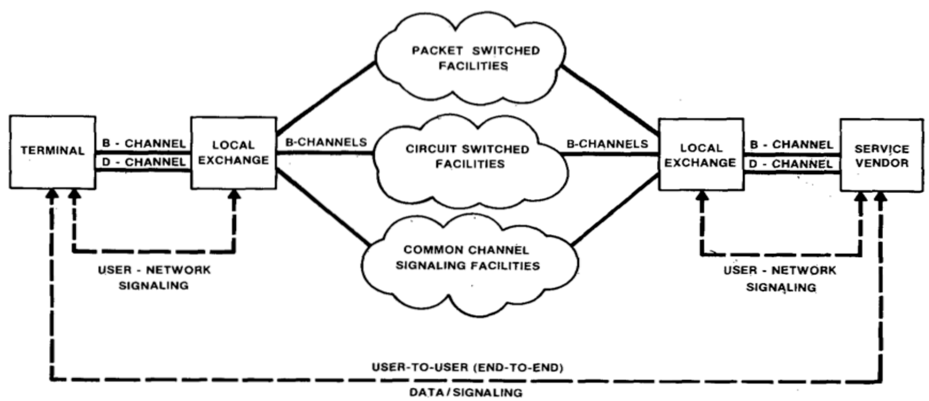


Fig. 7. The three-cloud network.

"Via the Internet"

- The service is not provided on-premises
- An Internet link is necessary
- This link provides an opportunity for interception, lawful or otherwise

“Outside Party”

- By definition, cloud services are provided by an outside party
  - Similar in spirit to the computing and time-sharing service bureaus, which date back to the 1960s
- *Not* the same as a company’s own remote computing facility
  - Organizations can have a “private cloud”, but the legal issues may be very different

Computing Services

- Many different types of services
  - Storage
  - Computing
  - Applications
  - Virtual machines
  - More

Storage

- Disk space in a remote location
- Easily shared (and outside the corporate firewall)
- Often replicated for reliability
  - Replicas can be on different power grids, earthquake zones, countries, continents, etc.
  - Data can be moved—or move "by itself"—to be closer to its users
- Expandable
- Someone else can worry about disk space, backups, security, and more
- Examples: Dropbox, Google Drive, Carbonite (for backups), Amazon S3
- Mental model: secure, self-storage warehouse

Computing

- Rent computing cycles as you need them
- Pay only for what you use
- Often used in conjunction with the provider's cloud storage service
- Examples: Amazon EC2, Microsoft Azure, Google Cloud
  - Dropbox is a cloud service that uses a different provider's cloud storage
- Mental model: calling up a temp agency for seasonal employees

Applications

- Provider runs particular applications for clients
- Common types: web sites, email services
- Less common types: shared word processing, payrolls
- Well-known providers: Google's Gmail and Docs, Microsoft's Outlook and 360, Dreamhost (web hosting)
- Mental model: engaging a contractor for specific tasks

Playing an Active Part: Google Docs

- Someone, using a Web browser, creates a document
  - Standard formatting buttons: font, italics or bold, copy and paste, etc.
- Others who have the proper authorization (sometimes just a special URL) can edit the document via their own Web browsers
- The changes made by one user show up *in real time* in all other users' browser windows
- In other words, Google is not just a passive repository; it is noticing changes and sending them out immediately

Virtual Machines

- Normal desktops: an *operating system* (e.g., Microsoft Windows) runs the computer; applications run on top of the operating system
- Virtual machines: a *hypervisor* running on a single computer emulates multiple real computers. A different operating system can run on each of these emulated computers—and each one is independent of the others and is protected from it
- Net effect: many computers that consume the space and power requirements of a single computer
- Mental model: rented office space

Location of Cloud Servers

- Responsiveness of and effective bandwidth to a server is limited by how far away it is
  - The problem is the speed of light—and not even Silicon Valley can overcome that limit!
  - It takes a *minimum* of a quarter-second to set up a secure connection from Washington to Paris, and twice that to New Delhi
- For performance reasons—and independent of political and legal considerations—large cloud providers therefore place server complexes in many places around the world
  - Also: take advantage of cheap power and cooling

Where is Data Stored?

- Modern email: on the server *and* on one or more devices
  - Users can't easily tell what's on their device (e.g., phone or laptop) versus what is retrieved from the server on demand
  - It differs for different devices at different times, and may depend on the user's recent activity
  - What if the device and server are in different jurisdictions?
- (A bad fit for the assumed behavior model of Stored Communications Act)

Security and Privacy Issues

- Gmail: Google applications scan email and serve up appropriate ads
- Dropbox: uses Amazon S3 for actual storage; encrypts data so that Amazon can't read it—but Dropbox can
- Spider Oak: data is encrypted with the user's password; Spider Oak can't read it
- Outlook.com: blocks file attachments that frequently contain viruses
- Many: check pictures for known child pornography
- Many: spam filtering

Interconnections

Interconnections

- The Internet is a collection of interconnected ISPs
- There are several types of ISPs
  - Individuals and organizations connect via an *access provider*
  - *Transit networks* talk to each other and to access networks
  - *Content distribution networks* ship out large, seldom-changing files—pictures, music, movies—on behalf of large content providers
- Architecturally, they're the same—but some are bigger than others and have faster links
- Most connections (and in particular most Web traffic to major sites) use all three types

ISP Architectures

- Internal architectures of all ISPs are highly engineered
- Twin goals: performance and reliability (and of course cost matters)
- Reliability is achieved through redundancy: there are alternate routes for *everything* (except, in general, the “last mile” link to customers
- Links generally run at <50% capacity—leave headroom for load spikes and to provide backup capability in event of a failure elsewhere

Building a Network

- Networks are composed of *links*—wires or fiber optic cables—and *routers*
- Routers are highly specialized computers that receive *packets* from one link and send them out over another, either to an end system (i.e., a computer) or to another router
  - There are often many outbound links from a router; the router has to choose the right one
- If a router or a link fails in the middle of a conversation, subsequent packets can take a different path
- Links are *always* shared; packets from many different conversations are intermixed on any link

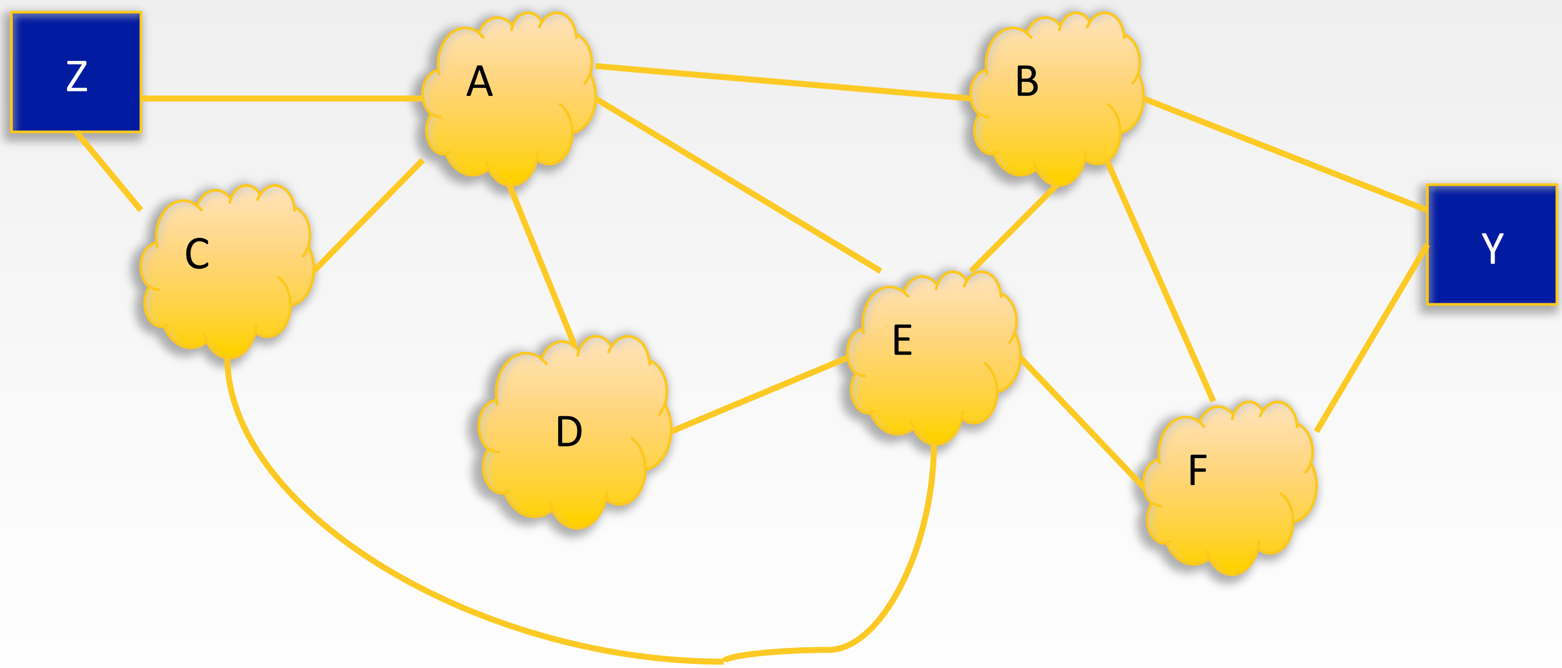
Links

- Inside a home: primarily WiFi
  - Reasonably secure *if* you use WPA2 encryption and a good password
- Businesses: primarily Ethernet (100M bits/sec or 1g bits/sec); some WiFi
  - Intelligence agencies can probably monitor unencrypted Ethernet
- ISPs: point-to-point fiber at 10G bps and higher; often leased from telcos
  - Tremendous capacity available because of *Dense Wave Division Multiplexing* (i.e., using subtly different colors for different channels)
  - Popular myth: fiber isn't tappable
  - Intelligence agencies can do it—and they can also ask a telco for access

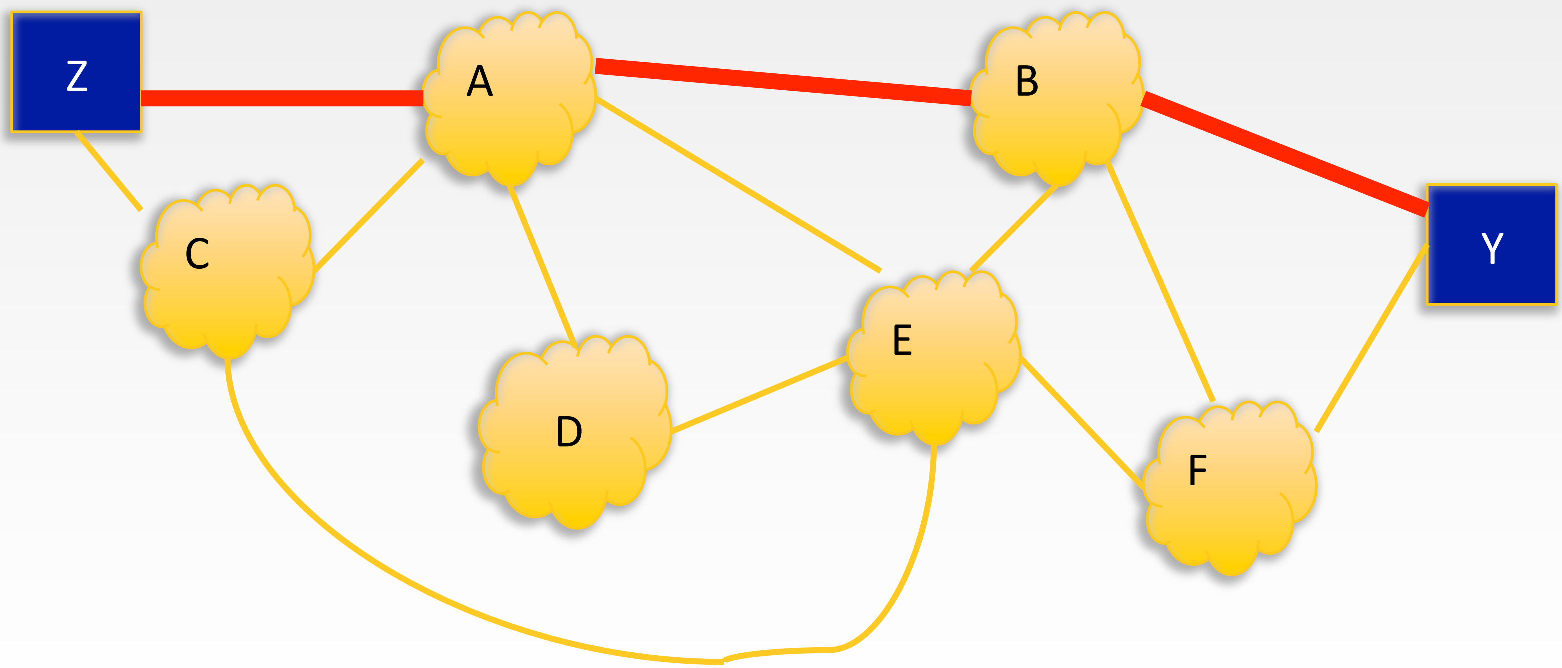
Inter-ISP Routing

- Connections between ISPs are governed by complex, generally confidential contracts
- Wide variety of payment terms and conditions: no fee, payment if traffic in one direction exceeds traffic in the other direction by some amount, payment for excess peak-hour bandwidth, etc.
- Wide variety of policies on what sorts of traffic can be sent over the link, and in particular what the permissible sources and destinations are
- (Much of the net neutrality debate is about these two points.)

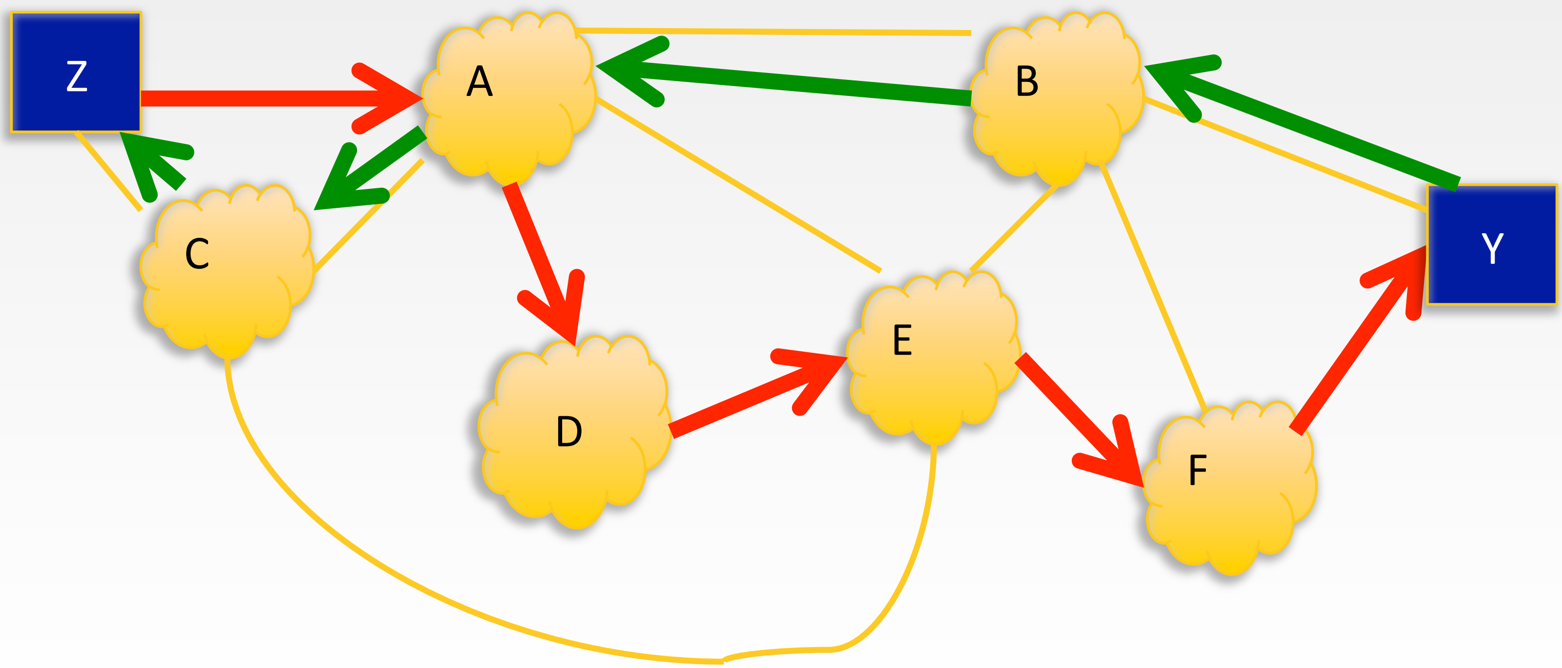
Inter-ISP Routing—Which Path?



Z-A-B-Y is shortest—but do contracts permit it?



It could be Z-A-D-E-F-B-Y — and the path from Y to Z could be completely different



What Does Z Know?

- In general, each entity—"node"—knows only the next hop
- Z does not know the full path, nor even its length
- Z cannot control the path except for the first hop, i.e., via A or C
- ISPs learn the next hop via a very complex technical process using "routing protocols". Routing protocols take into account efficiency, business contracts, cost, load-balancing among different links, current outages, and more.
- International routes often take a non-obvious (and counterintuitive) path
- For complex reasons, the reverse path may be completely (and very frequently is) completely different

The Philosophy of Routing

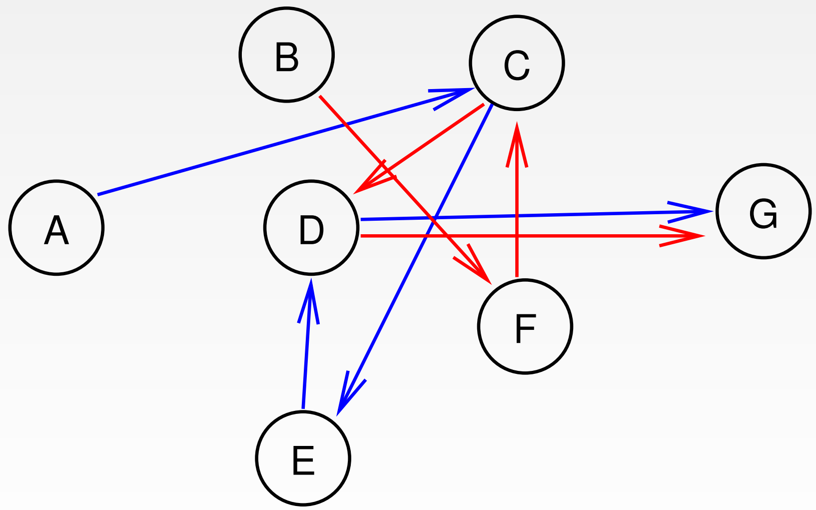
- Generally speaking, ISPs want to get rid of packets as soon as possible: let someone else bear the expense of carrying the traffic
  - But this isn't always true…
- Packets are routed in a way that makes economic and technical sense—and generally without regard to national boundaries
  - Some countries, e.g., China, do impose policy restrictions
- The Internet grew up in a deregulatory era, and without the legal legacy of older, highly regulated telecommunications technologies

Where is Y?

- In the abstract, Z cannot tell
- In practice, a number of companies offer *IP geolocation* services that tell you where some other computer is
  - There are several different technologies for doing this
  - Accuracy varies, but 90-95% is probably a reasonable guess
  - Geolocation is frequently used for geographic marketing rights (e.g., can a site show a movie in a given country?) and by gambling sites to avoid coming under the scope of US law
  - It's also used to target ads and to show content in the local language
  - Location is often—but not always—spoofable
  - Locations reported by smartphones are generally more reliable

Tor: The Onion Router

- Computer A picks a sequence of Tor relays (C→E→D)
  - D is the exit node, and passes the traffic to destination host G
  - All of these hops are encrypted
- B picks relays F→C→D
  - G can't tell which is from A and which from B
- Neither can anyone else monitoring G's traffic
- Many use Tor for anonymity: police, human rights workers, spies—and criminals (e.g., Ross Ulbricht of Silk Road fame)
- Mental model: nested, sealed envelopes



Location Accuracy

- The NSA actually has a patent (US 6,947,978) on one technology—roughly, triangulation based on the time (which is distance at the speed of light in fiber) from known locations to the target
- A clever target may be able to introduce great uncertainty, but possibly only at a considerable cost in performance
- Virtual Private Networks (VPNs), which are frequently used by business travelers, can mask location
- If you tap a link going to an overseas router, you know where the next hop is —but you don't know the location of the ultimate destination

Identifying Computers

- IP addresses identify computers, but…
  - For computers other than servers, IP addresses are assigned temporarily
  - Some residential ISPs *deliberately* change customers' IP addresses, to make it harder to run servers at home
- Home computers and computers in public hotspots—hotels, coffee shops, this room, etc.—generally share a few *global* IP addresses
  - On the inside, they each have a different *private* IP address that the border router modifies using *NAT* (Network Address Translation)
  - In other words: you often need a precise timestamp and cooperation from the network operator to track down a computer given its IP address
- There are sophisticated ways to spoof even global IP addresses—definitely used by spammers

Identifying People

- Hacking attacks almost never originate from the apparent origin
  - For decades, hackers have used *stepping stones*: use one computer to hack a second, use that to hack a third, launch the real attack from that one
- It's harder to spoof use of services where a password is needed—but of course passwords can be guessed or stolen
- Family members often share a computer and perhaps an email login
- Nation-state attacks are very hard to attribute
  - Use modus operandi
  - Use programming style
  - Correlate technical details with other forms of intelligence

Encryption

Encryption

- Can provide secrecy
- Can provide authentication
- *Very* hard to design good encryption mechanisms
  - These days, it's a branch of applied mathematics
- Often hard to *use* encryption securely
  - One of the major reasons the British could crack the German Enigma machine during World War II was operational mistakes by the Germans

What is Encryption?

- What you want to protect is called *plaintext*
- You feed the plaintext and a *key*—a long, random number—into an encryption algorithm to produce *ciphertext*
- You need the key and the ciphertext to produce plaintext
  - Protecting keys is *crucial*
- You do this in a stylized form called a *cryptographic protocol*
- No one in the unclassified community knows what the NSA (or other intelligence agencies) can break—but it's pretty certain that breaks aren't free; it probably takes a lot of computation and time for each message
- The NSA has stated that certain common algorithms are good enough for TOP SECRET traffic—*if* used correctly. But they take advantage of mistakes

Conventional Encryption

- The same key is used for encryption and decryption
- Keys must be shared in advance
- If you receive a message encrypted in a key, you have reasonable assurance about who sent it *if* you've shared the key with only one other person
- But you can't prove that to a judge; you have the key, too, so you could have forged the message
- Key lengths: 40-80 digits

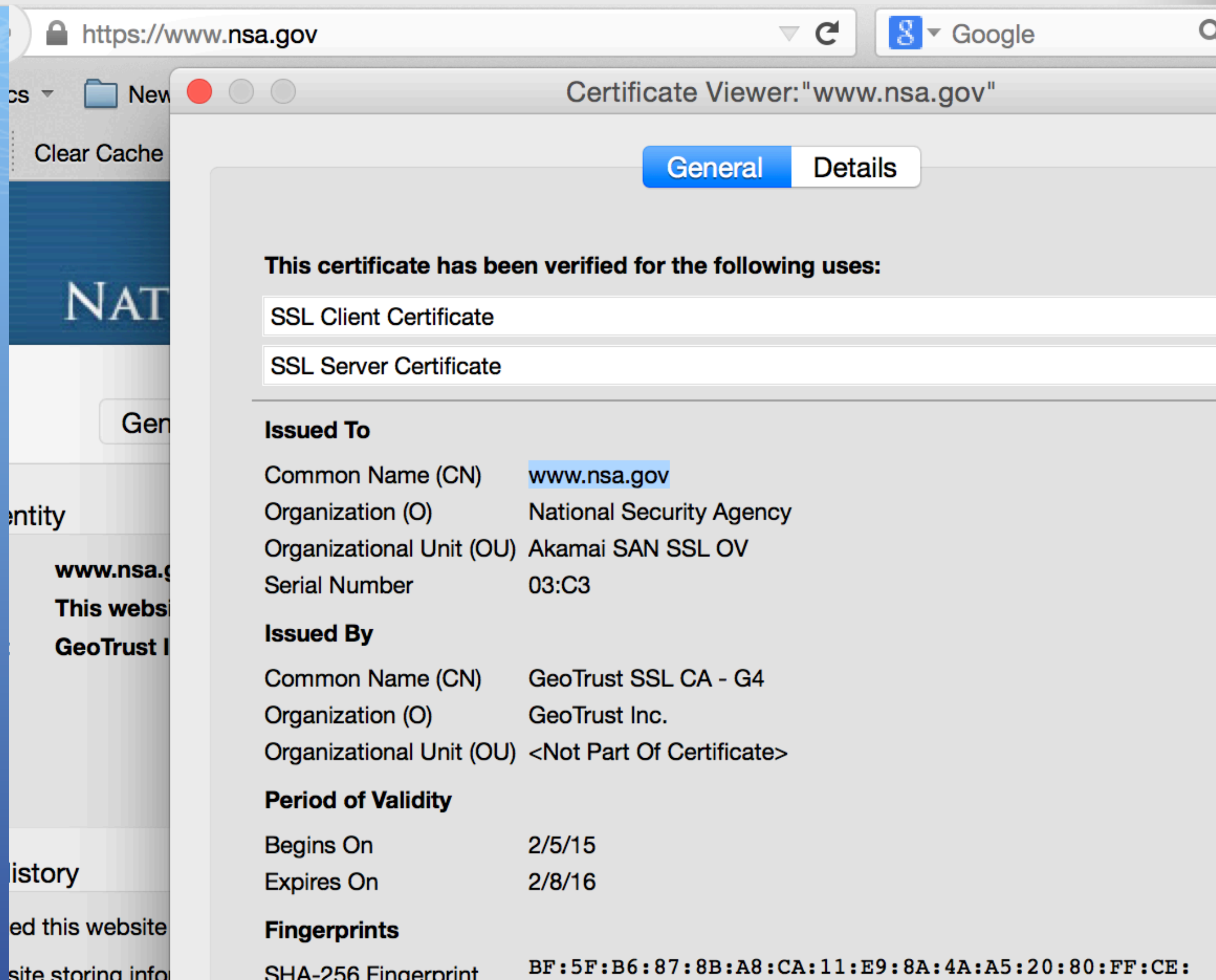
Public Key Encryption

- Separate keys are used for encryption and decryption
- You can publish your public (encryption) key; anyone can use it to send you an encrypted message
  - *Only you* have the private (decryption) key
- If you encrypt a message with your private key, it's called a "digital signature"
- Anyone who has your public key can verify the signature, and demonstrate this publicly
  - Note: no longer deniable, unless you can show that your key was stolen
- Key lengths: 600 digits

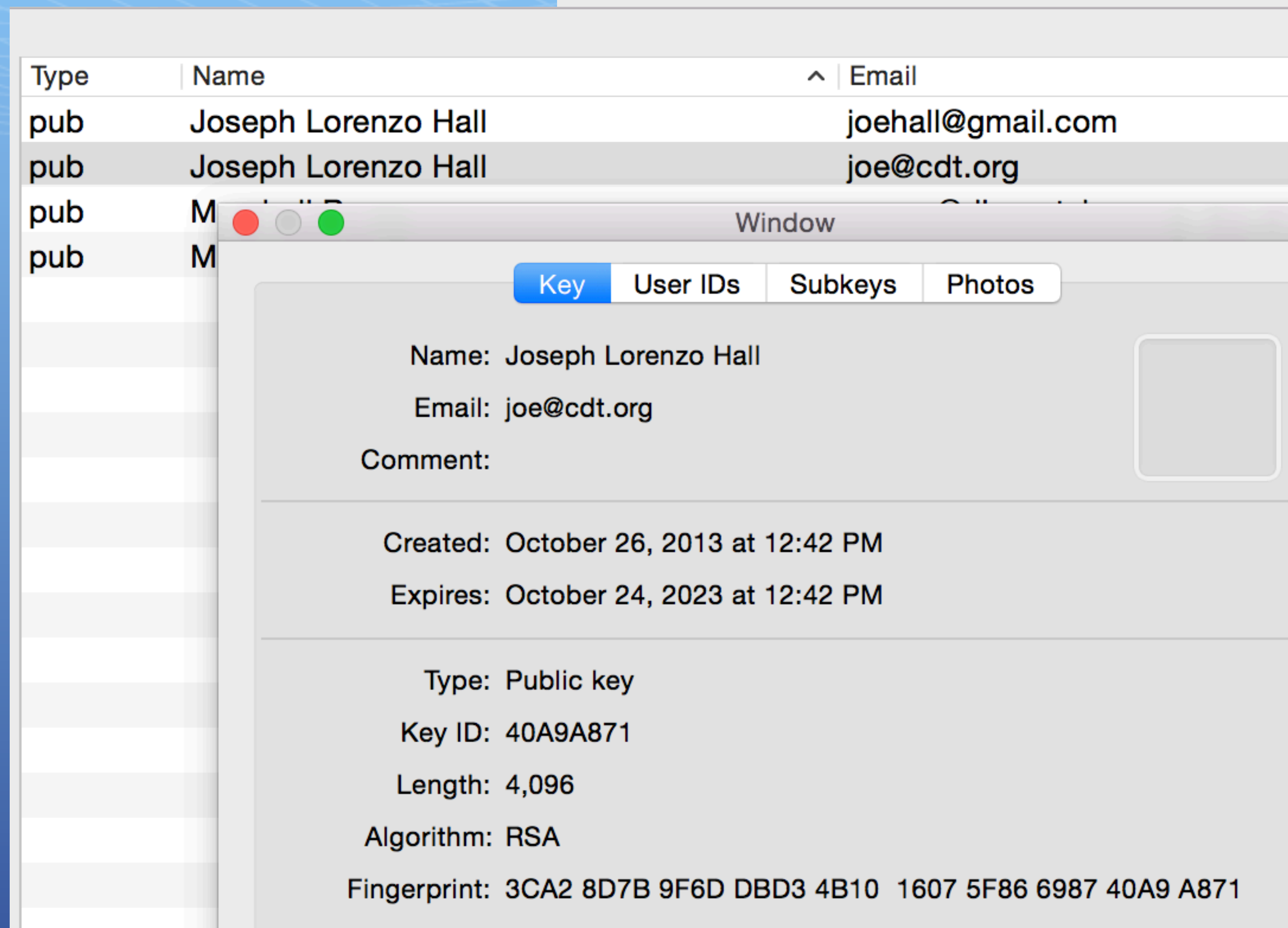
Usage Issues

- Who owns a key?
- How is the key protected?
- How do you know it is legitimate?
- On the Web, we use *certificates*
  - Someone else has vouched for the identity of the key owner (using cryptography)
- Who can vouch for it?
  - On the Web, many hundreds of *certificate authorities*

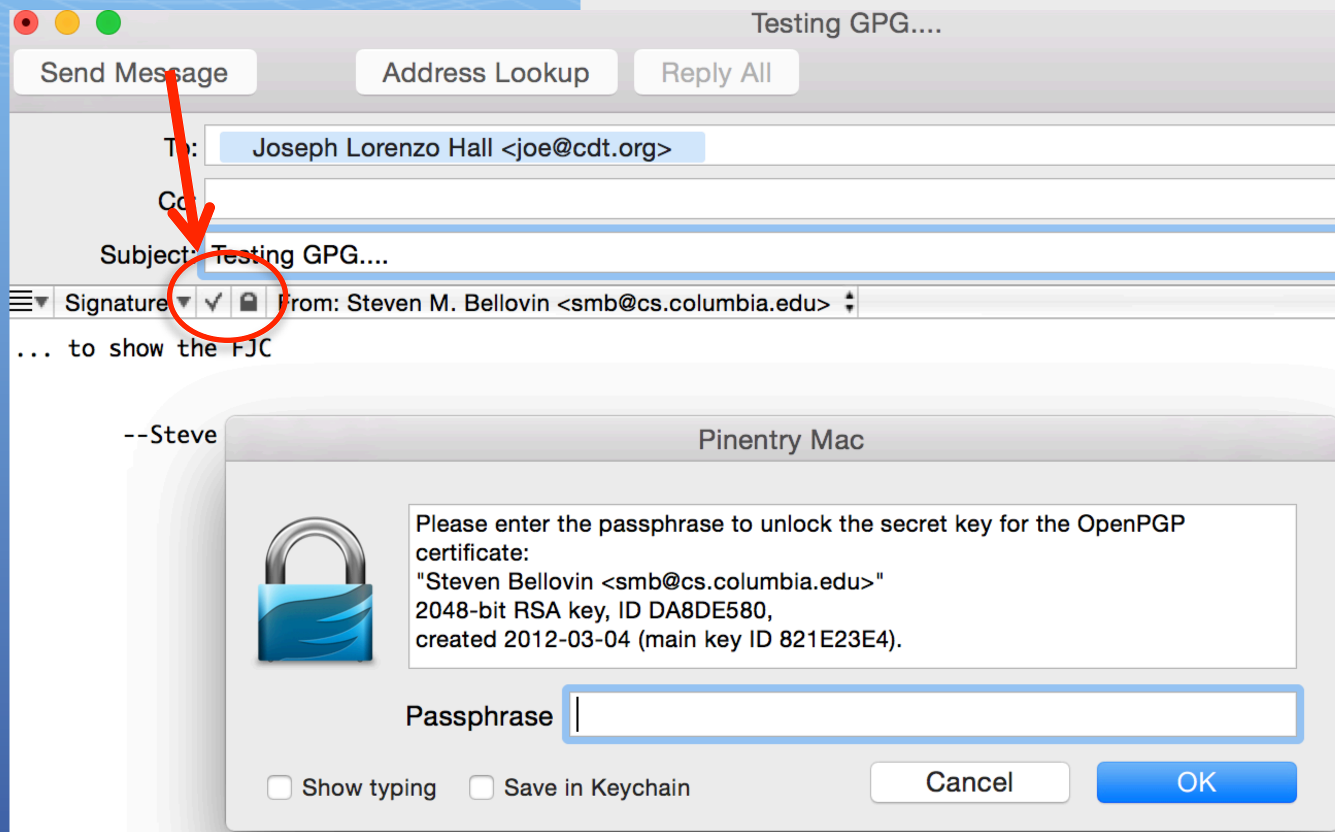
The NSA's Web Certificate



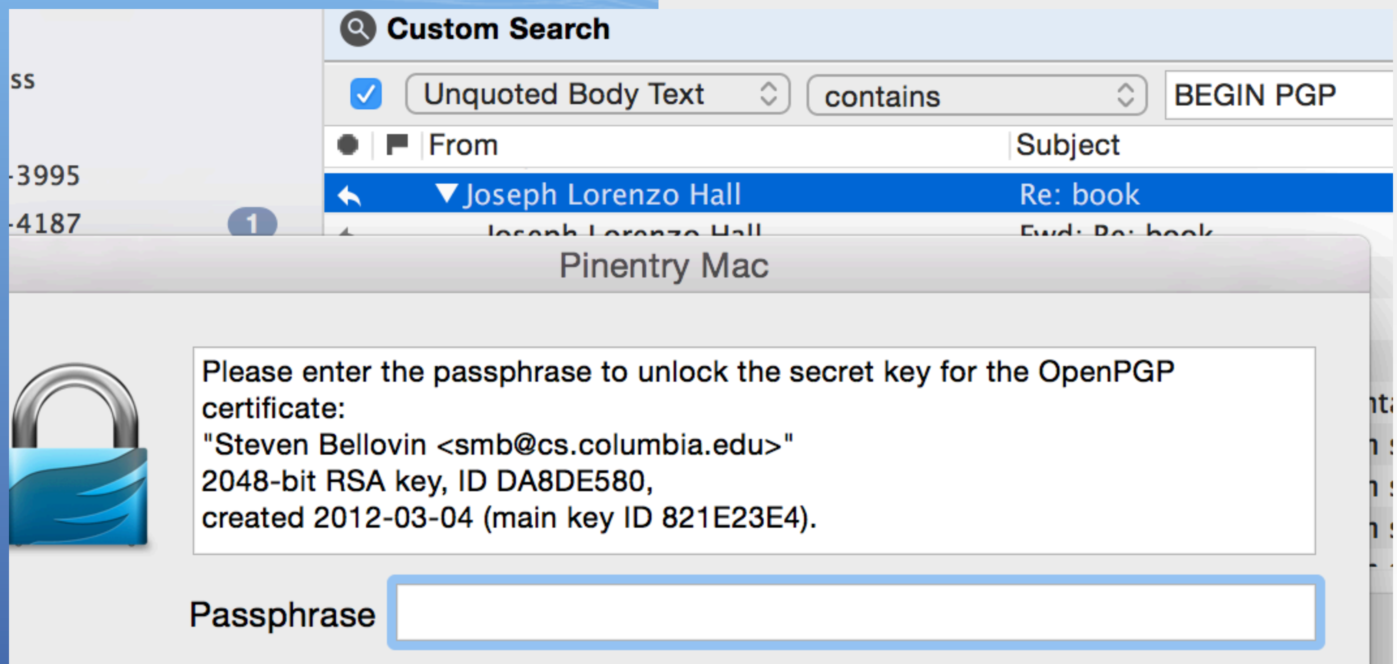
Selecting an Email Key



Sending Encrypted, Signed Email



Receiving Encrypted Email



iPhone Encryption

- (Important) memory is encrypted with a randomly key generated by the phone itself
- This key is itself encrypted
- That key is stored in a secure area of the chip and encrypted with the user's PIN
- Because of the secure storage, the only way to decrypt it is to try all PINs—and PINs can now be very long