# History of telecommunications and the Internet

Week 12a - April 10

# Creating a research poster

- Any word processor, drawing, or page design software will work
  - PowerPoint is well-suited for making posters
- Design poster as single panel or modular units
  - Single panel posters
    - Have a professional look (if well designed)
    - Should be printed on large format printers
  - Modular units
    - Easier to design and transport
    - Print on letter paper (optionally, mounted on construction paper)

# Research poster content

- Don't try to present your whole paper
  - Convey the big picture
  - Don't expect people to spend more than 3-5 minutes reading your poster
  - 500-1500 words, *maximum*
- Introduce problem, your approach, and results
- Provide necessary background or glossary
- A picture is worth 1000 words
  - Graphs, diagrams, etc.
- Use bullets and sentence fragments, similar to making slides
- Don't forget to include title and author

# Research poster design

- Use a modular design
- Each section of your poster can go in a box
- Use a large, easy-to-read font
  - Most text should be at least 20 point font
  - No text less than 14 point font
  - Headings should be larger and in bold
- Use color consistently
- Arrange elements for a sensible visual flow

# Presenting your research poster

- Be prepared to give a 1-minute overview of your poster and answer questions
- Let people read your poster without interrupting them
- Consider bringing a laptop if you have software to demo or a video to show
- Consider making handouts available with abstract, web URL for obtaining your paper, and your contact information

## Motivation

- First P3P Search Engine by Byers, et al.*
- Dual purpose engine:
  - Provide P3P-enabled search
  - Facilitate P3P deployment research
- Goals:
  - Improve scalability and usability
  - Add research functionality

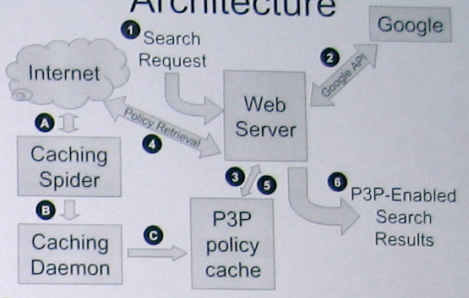
*Byers, S., Cranor, L. F., Kormann, D., and McDaniel, P. Searching for privacy: Design and implementation of a P3P-enabled search engine. In *Proceedings of the 2004 Workshop on Privacy Enhancing Technologies*, pp. 26-28.

# Enhanced P3P-Enabled Search Engine

**Damon Smith**

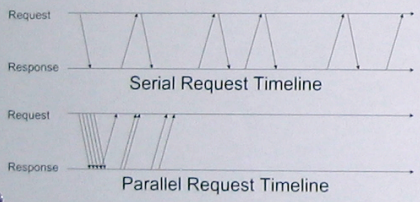
15508 Privacy Policy, Law, and Technology

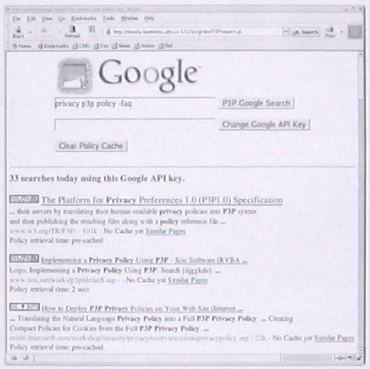
## Architecture



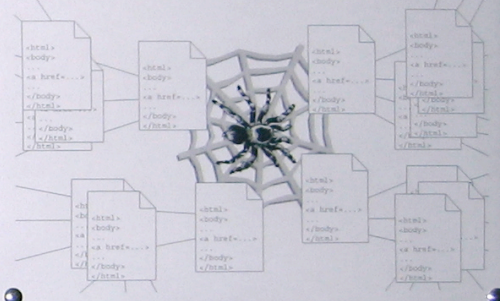

## Policy Retrieval

- Problem: performance bottleneck
- Solution: parallel policy requests




## Proactive Caching



## Presentation

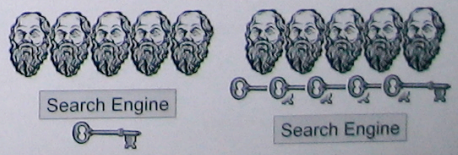
- Custom APPEL rule sets in addition to predefined low, medium, high rule sets
- Save rule set choice in cookie
- Privacy Bird Icons
  - Green: policy passes rule set
  - Yellow: no policy
  - Red: policy fails rule set

## Cache Query

- P3P policy cache as a research tool
- Use special queries:
  - p3p:<website> *display cached policy*
  - p3pstat: *stats about policies in cache*
    - Total policy count
    - Percent deployment
    - Multiple policies/site count
  - appel:<ruleset> *test policies against rule set*
    - Percent matching policy

## Scalability

- Google API allows 1000 queries per day
- Let users input their own Google API key




**1000/N** queries per user

**1000** queries per user

# April 26 Poster Fair

- During class, in meet in NSH 3305
- Arrive on time!
- Other faculty and students are invited
- 32x40 inch foam core boards, 9x12 inch construction paper, glue sticks, and thumb tacks will be made available
  - Pick them up from Jennifer Lucas in Smith 231A jmlucas@cs.cmu.edu
- Use this as an opportunity to get feedback you can use to improve your final paper!

# Design and History of the Internet

# Design and History of the Internet

Layperson misconceptions

WWW = Internet = Email = online = broadband

- Some questions to think about
  - Who owns the Internet? we all do (collective)
  - Who controls the Internet?
  - Is the current system OK?
    - Security
    - Scalability
    - Usability

# Structures of the Telecom Industry

- Government Dept.
- Government company (PTT)
- Regulated Monopoly
- Competition
  - Splits within sectors
    - IXC – InterExchange Carrier (Long Distance)
    - ILECs – Incumbent Local Exchange Carrier ("Baby Bells")
    - CLECs – Competitive Local Exchange Carrier

# Government Departments

- Losing ground
- Privatization big push
  - Type 1
    - Public Assets privatized and then regulated
  - Type 2
    - Government carrier becomes one of many players

# PTT

- PTT: Abbreviation for postal, telegraph, and telephone (organization). In countries having nationalized telephone and telegraph services, the organization, usually a governmental department, which acts as its nation's common carrier.

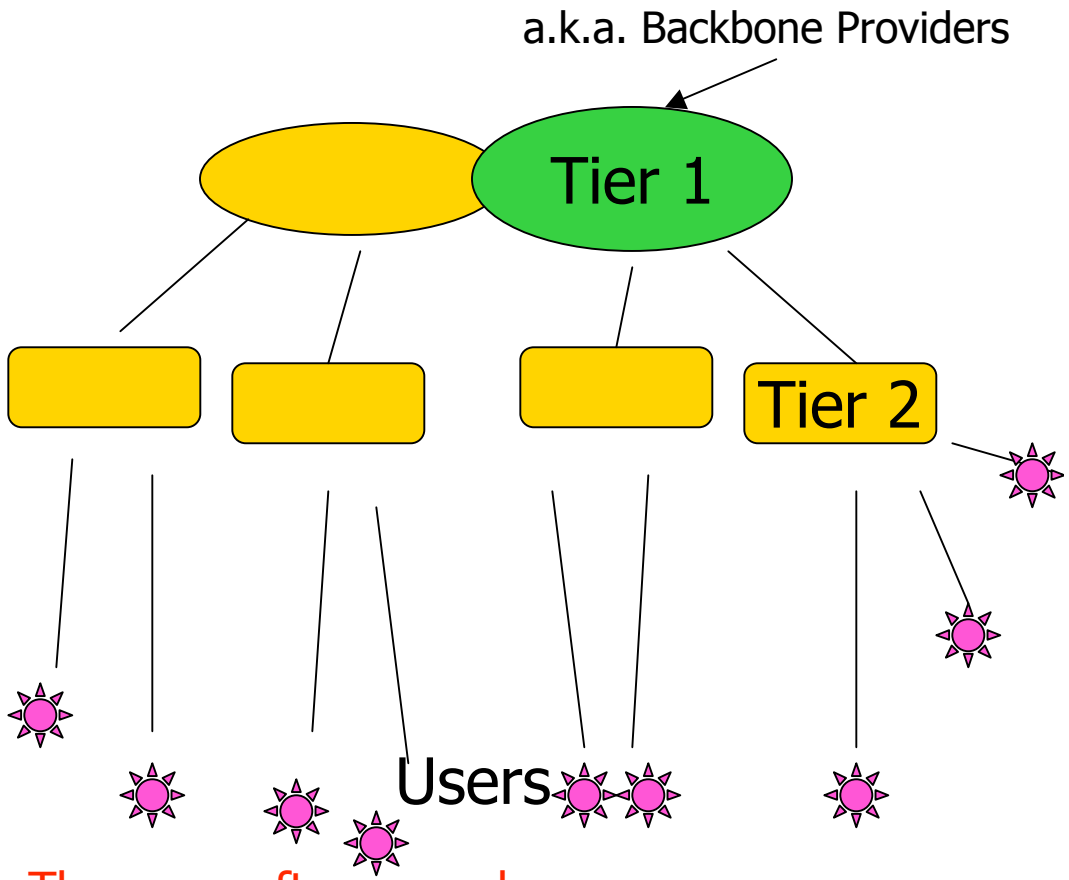
# "Call/Transaction" Completion Charges

settlement

- Mail
  - Flat Rate
- Telephony
  - Usage based *or* flat rate
- Internet?
  - Depends on what user (residential, commercial, bulk, etc.)

# What is the Internet?

- The global (public) network built from hundreds and thousands of internetworking independent networks.
- No single entity "runs" the Internet
- Operates on standards
- Built on a modified hierarchical structure
- Packet Switching




- There are often more layers
- There can be interconnections other than at a backbone

# What makes the Internet the Internet?

- Open architecture
  - Standards and protocols allow applications and communications without caring of the underlying infrastructure or system
    - "The Cloud"
  - Anyone can access anything (is public)
- Resiliency (mesh design)
- End to end system

# How big is the Internet?

- Many metrics
  - Number of Service Providers
  - Number of Hosts
  - Number of Subscribers
  - Size of Interconnections

  - (see outside sources such as CAIDA, Hobbes Internet Timeline, etc.)

# Brief History of Internet Evolution

| | | |
|---|---|---|
| 1969 ARPANET | 50 kbps | UCLA, UCSB, SRI, and Utah |
| 1970 | 56 kbps transcontinental | adding BBN, MIT, RAND |
| 1972 | 50 kbps | 23 hosts |
| 1973 | 75% of traffic on ARPANET is email | |
| 1981 CSNET (in parallel) | 56 kbps | 213 hosts |
| 1983 | TCP/IP mandatory, DNS created | 562 hosts |
| 1985 NSFNET initiated | 1.544 Mbps | 1961 hosts |
| 1987 UUNET created for commercial access | | |
| 1990 ARPANET disbanded in favor of NSFNET | | 313,000 hosts |
| 1992 NSFNET (+ a few pvt. Backbones) | 45 Mbps upgrade complete | 1,136,000 hosts |

# Brief History of Internet Evolution (cont.)

- 1994 NSFNET 145 Mbps ATM 3,864,000 hosts
  (+ a few pvt. Backbones of 56 kbps, 1.5 Mbps, and 45 Mbps)
- 1995 NSFNET privatized to 4 players 6,642,000 hosts
- 1996 MCI 622 Mbps
- 1996 - Now upgrading to 2.5 and 10 Gbps *IP links*

*This history has helped shape US Internet architecture in terms of competition and layout (peering)*

# Peering

- Where backbones come together
  - Major design issue (relates to cross-connection)
- Public Peering
  - Network Access Points (NAPs)
    - Started with 4, but now there are more
    - Usually done by equals
      - Give as much traffic as receive
- Private Peering
  - Commercial (private)
- International peering is more limited (links are much more expensive)

# Open Systems Interconnection (OSI) Model

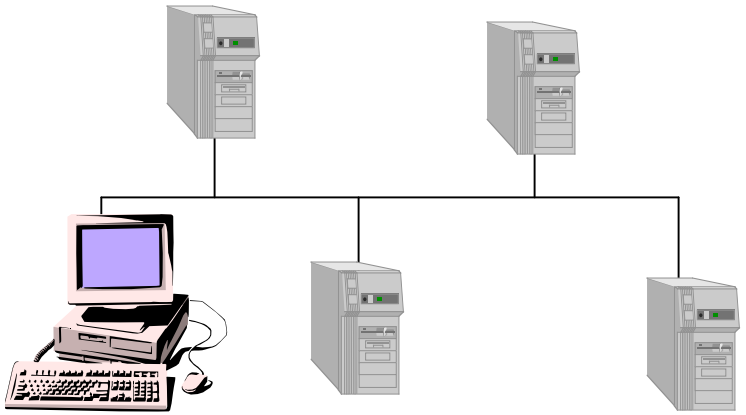
| OSI MODEL | | | *examples* |
|---|---|---|---|
| 7 | **Application Layer**<br>Type of communication: E-mail, file transfer, client/server. | Interface : MESSAGES<br>User Interacts with these | FTP, Ping, HTTP, etc. |
| 6 | **Presentation Layer**<br>Encryption, data conversion: ASCII to EBCDIC, BCD to binary, etc. | Translation and encryption : MESSAGES | |
| 5 | **Session Layer**<br>Starts, stops session. Maintains order. | Remote Procedural Calls (RPCs), Error Checking : MESSAGES | |
| 4 | **Transport Layer**<br>Ensures delivery of entire file or message. | Reliability, Error-checking : SEGMENTS<br>end-to-end validity | TCP |
| 3 | **Network Layer**<br>Routes data to different LANs and WANs based on network address. | Software Address, Routers : DATAGRAMS<br>establishes routes (extends nodes…) | IP |
| 2 | **Data Link (MAC) Layer**<br>Transmits packets from node to node based on station address. | Hardware Address, Bridges, Intelligent hubs, NICs, Error Checking : FRAMES<br>node-to-node validity | Ethernet, ATM |
| 1 | **Physical Layer**<br>Electrical signals and cabling. | Pins, Wires, Repeaters, RS-232, Volts, etc : BITS<br>Deals with the medium | SONET/SDH |

# Ethernet

- A standard for networking at Layer 2
  - Based on physical hardware address (12 Hex numbers)
- First started within the LAN
- Started of as a shared bus (from the Aloha Packet Radio network – Bob Metcalf)
- New versions are full-duplex, switched
  - Amenable for optical, longer reach
- Graceful evolution (backwards compatible) between 10/100/1000 Mbps
- Ethernet Frames are between 64 and 1518 bytes in size
- IEEE is the standards body (802.xx working groups)

# Ethernet Operation (traditional)

- Carrier Sense Multiple Access/Collision Detect (CSMA/CD)
  - All machines wait to see if medium is free
  - If so, they transmit
  - Sometime, packets can collide
  - In that case, the transmitters wait a random period of time, and re-transmit
  - If yet another collision, will wait longer period of time ("exponential back-off")



- Limitations
  - Effective bandwidth was modest
  - Distances were limited
  - Non-duplex

# TCP/IP

- Suite of protocols for networking
- Based on logical address for devices
- Most popular standard worldwide – built into most OS
- Like most other packet switching, is
  - Connectionless
  - Statistical (non-deterministic)
    - No inherent Quality of Service (QoS)
  - Most of IP routing is unicast
- Packets carry lots of information
  - Source Address, Destination Address, etc.
  - Special instructions such as priority
  - Port number (meaning application ID)
    - E.g., Port 80 - http

# IP Addresses

- Each device connected needs a unique IP address
  - Exception is “private” IP addresses used within non-global networks
    - Home gateways can use this
    - Gateway “router” translates between public and private IP addresses
- 32 bit addresses in current version (IPv4)
- 4 8-bit portions
  - Dotted decimal is popular for convenience
  - 128.2.72.44 is same as 10000000.00000010.01001000.00101100

CMU

# IP Addresses (cont.)

- IP addresses have 2 portions, network and host
  - Networks are uniquely controlled. e.g, 128.2.x.y. is CMU's network
- Earlier, IP addresses were class-based to differentiate

| Class | First Octet | Network/Host [octets] | # of Networks | # of Hosts per Network |
|---|---|---|---|---|
| A | 1 – 127 | 1/3 | 126 | 16,777,214 |
| B | 128 – 191 | 2/2 | 16,384 | 65,534 |
| C | 192 – 223 | 3/1 | 2,097,152 | 254 |

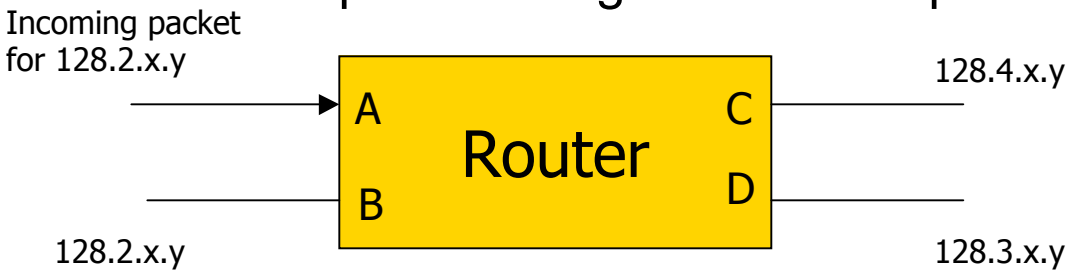
- Newer system is classless; can arbitrarily demarcate network and host
  - A.B.C.D/24 implies first 24 bits are for network portion
  - More efficient
  - "Subnet Mask" is used to identify network portion
- Most people don't own their own network; they take a portion from their service provider

# Network boundaries

- LANs used to predominate
  - Old rule of thumb: 80% traffic inside 20% outside
  - Often were Layer 2 networks
  - "Intranet"
  - Can make an outside, non-global network
    - "Extranet"
    - Often using private (leased lines)
- Outside world
  - Layer 3 connections (IP)
- Many types of interconnections, e.g., varying by
  - Speed
    - Dial-up
    - Dedicated connection – Just a pipe to the "cloud"
  - Protocol
    - IP, IPX, Appletalk, etc.

# Routers

- Forward packets based on destination address
- They know the route to every network
  - Once the packet gets to the network gateway, it internally finishes the routing
- Today's Internet is roughly ~200,000+ routes in size (advertised prefixes [2006 estimate])
- Routing is done on a hop-by-hop basis
  - A routing table is built up in each router
  - Incoming packet's destination address is looked up
  - A match is made, and the packet is forwarded to the appropriate port which gets it one step closer to the destination




Routing table knows which port (interface) is most closely connected to a particular network(s)

# IP Routing

- Core Routing
  - Internet-sized routing tables
  - Optical interfaces
- Edge Routing
  - Traditional edge players (aggregators)
  - Metropolitan Area Network/GigE edge players
  - Wide Area Networking is different from LAN, even though many protocols are the same
- Access (Customer Edge)
  - Often the bottleneck
  - Earlier, relied on the ILEC (e.g., Verizon)
  - Now, new carriers want to bypass the ILECs
    - Often use new technologies and standards

# Communications Components

- Transport
  - Now, typically optical, except the "last mile"
- Termination
  - Different devices (typically) for different layers
    - Phones, Video-conf. phones, routers, modems, etc.

cell phones

- Switching
  - Cross Connects / Add-drop Multiplexers (ADMs)
  - Class 4/5 switches
  - IP switches (Routers)

# Network Intelligence

- Quality-of-Service (QoS)
  - Today's Internet is "best-effort"
    - Need to differentiate different packets
  - Issues of identification, authentication, and billing
  - Critics content some schemes amount to violation of Net Neutrality
- Moving Intelligence to the Edge
  - Filtering, monitoring, and "differentiating"
  - Lets the core be super-fast
- Security
  - Today's internet is inherently insecure
  - Higher layers are used for security
    - E.g., SSL in browswers
  - New designs are being worked on for more security

# Internet is built on: Principles, not Laws

- Registration (databases) are believed because people think they are correct
  - Domain Name System
    - Handles names for humans vs. binary for machines
    - Root names are the last .xyz, e.g., .com, .edu, .org, .mil, .ca, .tv
    - Just 13 root servers in the world
      - Many copies made for practical purposes
- Borders define responsibilities
- Best effort (democratic)
- Robustness

"Be liberal in what you accept, and conservative in what you send."

- Jon Postel

# Standards and Regulation

- Many bodies, sometimes with overlap
  - IETF handles the engineering of the network
  - W3C handles web standards such as html, xml, etc.
  - IEEE handles some standards
- Requests for Comments (RFCs) are how things get standardized
  - Draft is circulated
  - Modified, debated, etc. (**many** versions often)
  - Becomes a standard by vote.
    - Companies often try and tilt emerging standards

# Registries and Domain Names

- Numeric address space is coordinated
- Domain Names initially managed by ISI (Jon Postel)
- National Science Foundation (NSF) hired contractor to administer
  - Network Solutions Inc (NSI)
- NSF stopped paying NSI, allowed NSI to charge for .com, .net, .org
  - $70 for two years
- NSI becomes enormously profitable
- NSF responsibilities passed to Commerce Dept.
  - The US government controlled key element of the Internet (!) *so*
- NSF establishes ICANN (Internet Corporation for Assigned Names and Numbers)

** Based on information from Jon Peha*

# Domain Names (cont.)

- ICANN decisions
  - Protect trademark owners
  - Oppose cybersquatting
  - Do not create more top level domains
  - Divide NSI responsibilities
    - Registry: manage database, NSI monopoly
    - Registrar: consumer interface, competition
- NSI claims to own the .com, .net, .org database
  - Do they have to give it up or share it?
- ICANN says that NSI must be accredited
  - NSI refuses to sign agreement with ICANN
  - NSI does not recognize ICANN's authority
  - NSI protects its revenue stream
- What happened in the end?
  - NSI was acquired by VeriSign, then spun off

# Domain Names (cont.)

- ICANN critics
  - NSI and friends, many academics
  - ICANN is the evil face of governance in the Internet, which needs no governance
  - ICANN is an unrepresentative, unelected group with unlimited power
    - Rest of World (especially developing countries) particularly dislike the entire process (not just ICANN)
  - Meet behind closed doors, create taxes ...
- ICANN supporters
  - ICANN, many high-tech companies, trademark owners.
  - NSI is an unregulated monopoly that must be stopped.
  - Engineers seeking consensus, do not address policy.
  - A neutral group of experts making necessary decisions.
  - ICANN people are just "plumbers“
- Remains a major issue: Internet Governance
  - What is the debate about?

# Issues in the Internet

- Scalability
  - Internet is growing* at 75-300%
  - Running out of IP addresses
    - Long term solution: IPv6
      - 128 bit addresses (millions per square meter)
  - Protocols and equipment are straining
- Security
  - Distributed Denial of Service are an example
  - Viruses
- Quality of Service
  - Voice
- Usability

# Issues in the Internet (cont.)

- Privacy
- Anonymity
- Identity
- Regulation
  - Universal Service Obligation
  - Taxation
  - Encryption (and it's a technology issue)
  - Digital signatures
- Digital Divide

# Policy Issues (Discussion)

- Are "Terms of Service" sufficient to disallow Domain resolution?
  - E.g., GoDaddy vs. Seclists.org [dispute over MySpace complain]
- How do we do CALEA on the Internet?
  - Can we?
  - Should we?
  - What about Skype?
    - Is not a phone service, but a "voice IM" (?)