Lecture #2: Design Principles #2

WPI CS4516 Spring 2019 D term

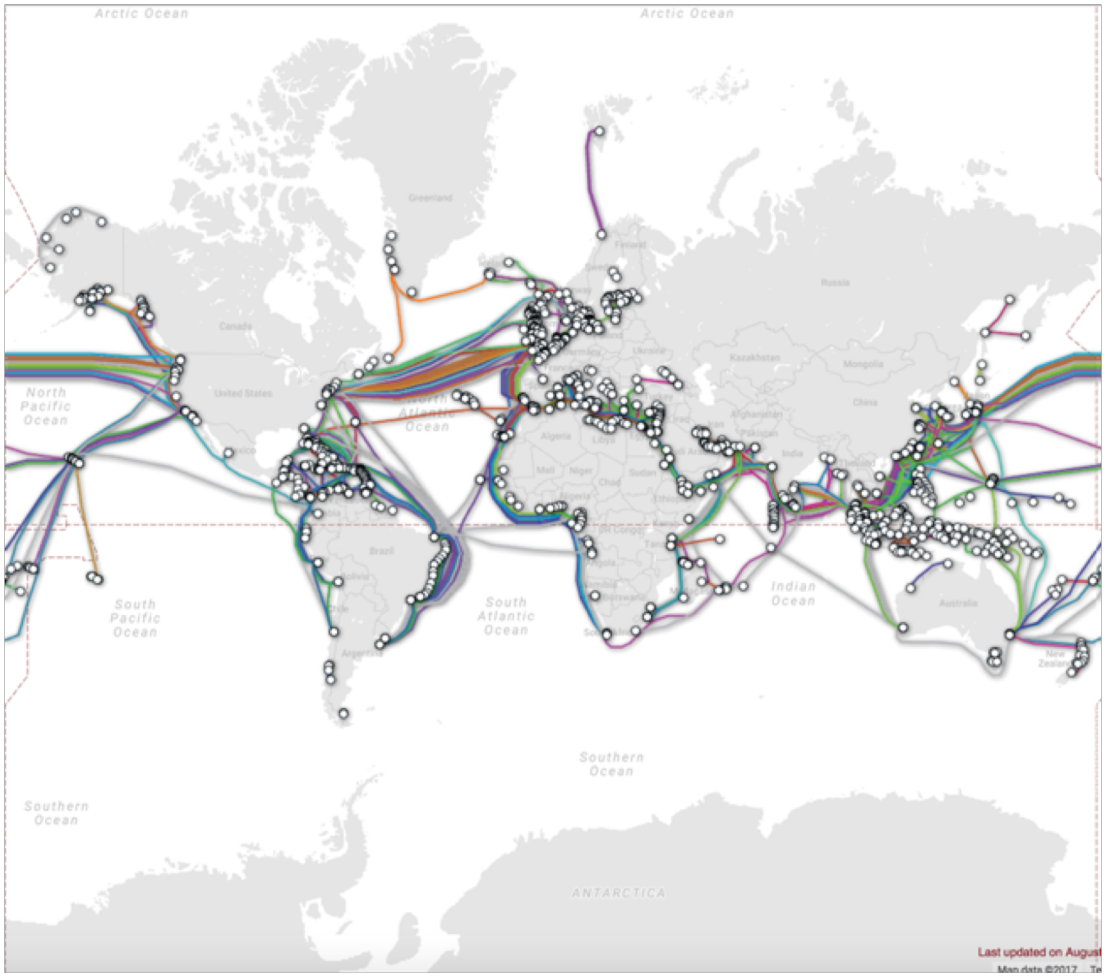
Instructor: Lorenzo De Carli (ldecarli@wpi.edu)

(slides include material from Craig Shue, WPI and Christos Papadopoulos, CSU)

How should we design the Internet?

- The Internet is so pervasive that we may not even notice it exists…
- …let alone think about how it works
- However, it is a very real and very physical infrastructure

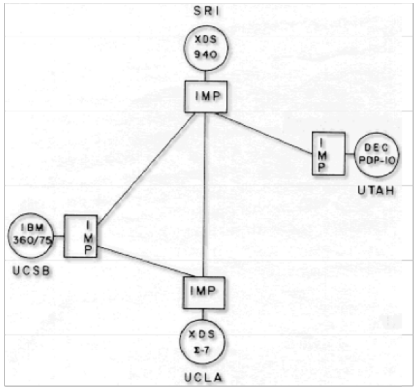
The Internet is a physical object!



https://www.submarinecablemap.com/

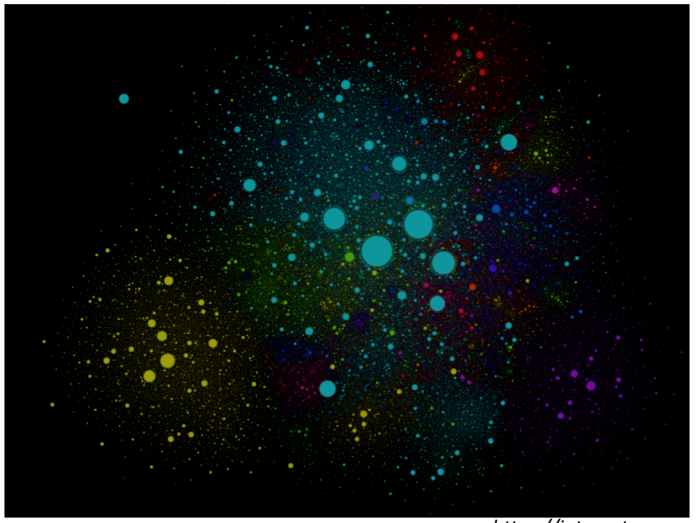
A bit of history

- The Internet grew **incrementally** from a small sets of hosts interconnected **as a result of military and academic projects**



http://histoire.info.online.fr/net.html

The Internet, 1969



https://internet-map.net/

The Internet, 2011

The first idea

- **Early concepts:** early 60s, mostly from BBN technologies
  - J.C.R Licklider, “Man-Computer Symbiosis”, 1960 (among other things, idea of a large interconnected network)
- Initial research funded by DARPA (1962) to interconnect their three computers (ARPANET)
- **Network governance slowly transitioned from military to civilian agencies** (NASA, NSF) in the 70s and 80s
- Other similar networks ”grew” in other parts of the world and were also interconnected

The early Internet

- **Commercial use** began in the late 80s/early 90s
- The **end of NSF stewardship** in 1995 gave a significant impulse in the transformation of the Internet from a **research network** to a **public and commercial tool**
- The **rise of WWW** (Tim Berners-Lee, CERN) also helped turning the Internet in something useful to the general population
  - **World-wide web:** a system of **documents** containing **text and graphics**, viewed in **browsers**, interconnected by **hyperlinks**, hosted on top of a global network

The modern Internet

- Browsers slowly turned from utilities to display hypertexts into **fully-blown operating systems**
  - Many website are full-blown client-server applications running in browsers instead of displaying native UIs
- However the **mobile revolution** (early 2010s to now) means that many network applications **do not run in browsers anymore** (although they may use similar technologies)
- **Cloud computing** contributes to **centralization of content** into large datacenters
  - Compare with the early Internet, where content was distributed among many smaller individual servers

Now, let’s go back to the beginnings…

There are many possible Internets… and you are in one of them!

- Building a network interconnect that spans the entire work was an herculean task…
- … and required a great deal of engineering effort from many smart people!
- Still, **what they built is only one of the many possible realizations of "the Internet"**

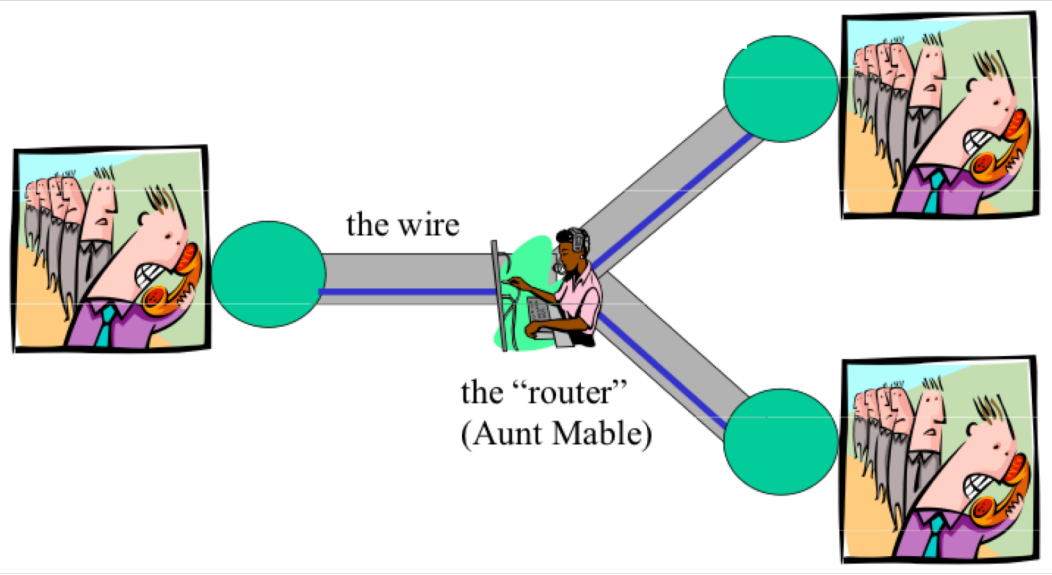
Goal of this lecture

- Understand why the Internet was designed the way it was
- Understand which ones of the original assumptions/priorities still make sense today, and which ones don’t

Basic design aspects

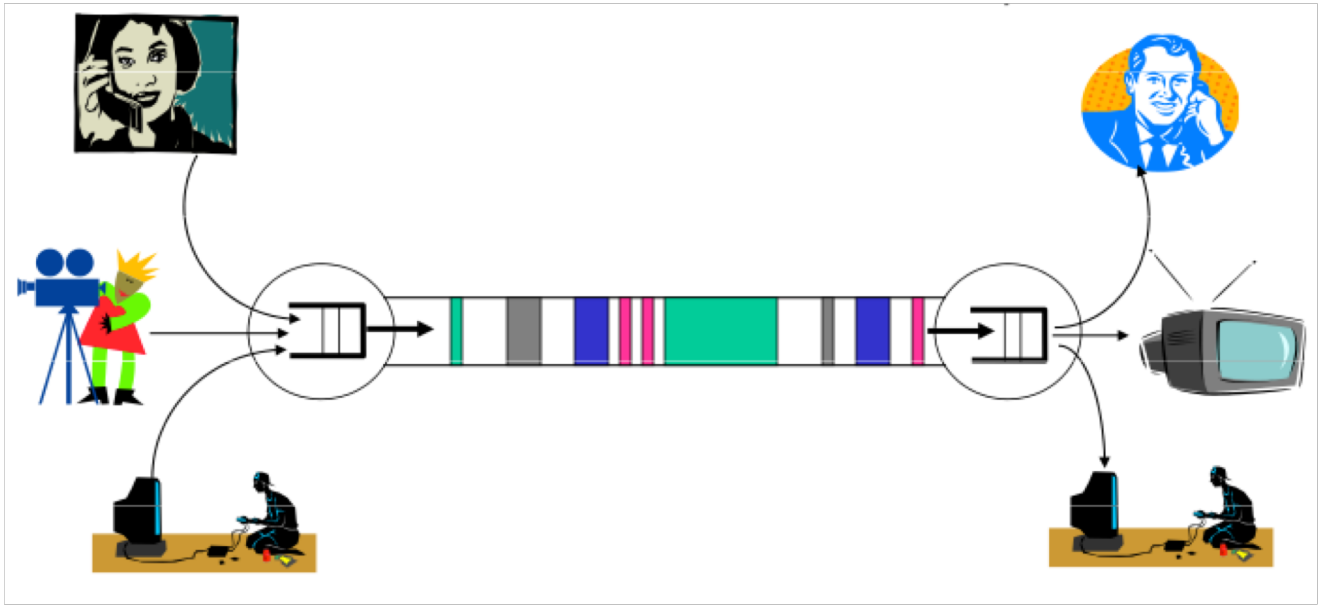
Circuit vs Packet Switching

- **Circuit switching:** data between two nodes follows a static path which must be reserved and set up prior to communicating. Concurrent communication between different host pairs must use different paths.



Circuit vs Packet Switching - II

- **Packet switching:** data is broken into packets which are forwarded across links between nodes. Communication flows between different host pairs can be multiplexed on the same link, simply by alternating packets from different flows on the wire

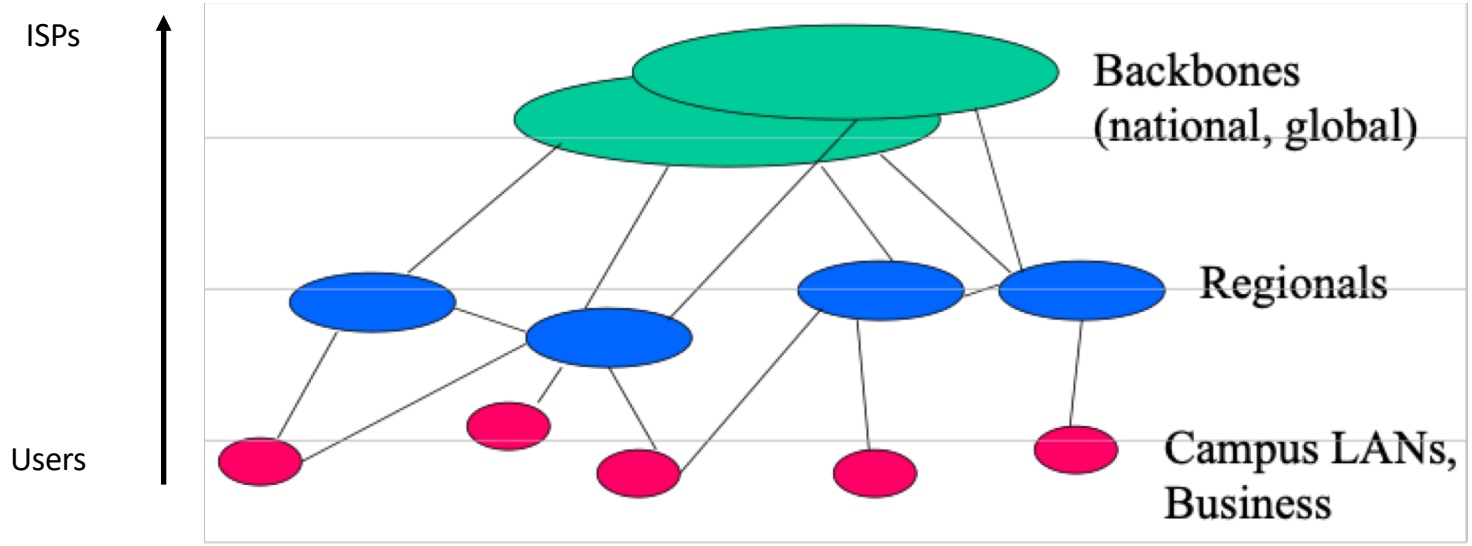


- Most modern networks are based on packet switching technology (although *virtual circuits* are sometimes used)

Store-and-forward technology

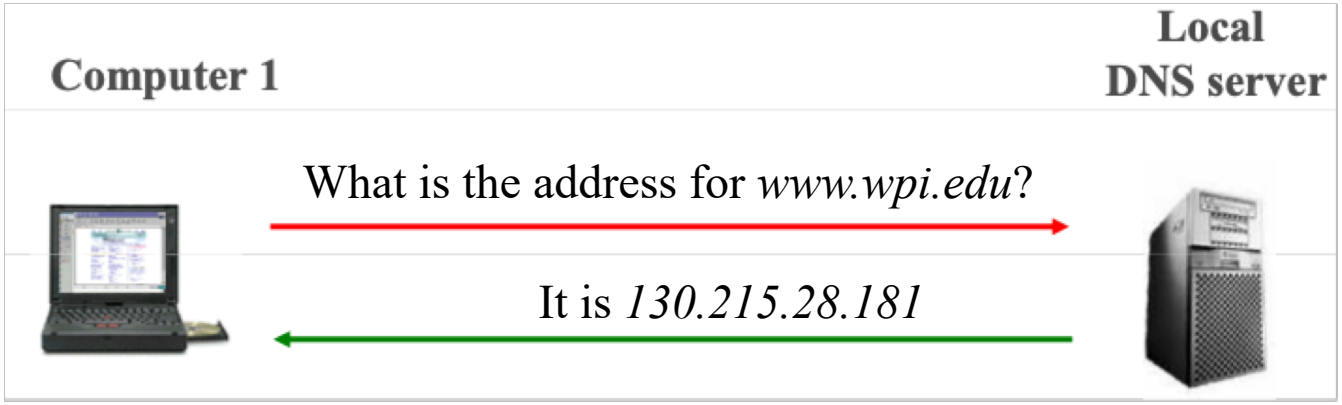
- **Store-and-forward** means that routers on the Internet receive and store (in their memory) an entire packet, before sending it on the next link
- Advantages:
  - **Packet can be processed** before being sent on the next hop
    - Fragmentation, defragmentation, NAT, etc.
  - Some types of **errors** can be **identified**
    - E.g., whole-packet checksum

Internet structure

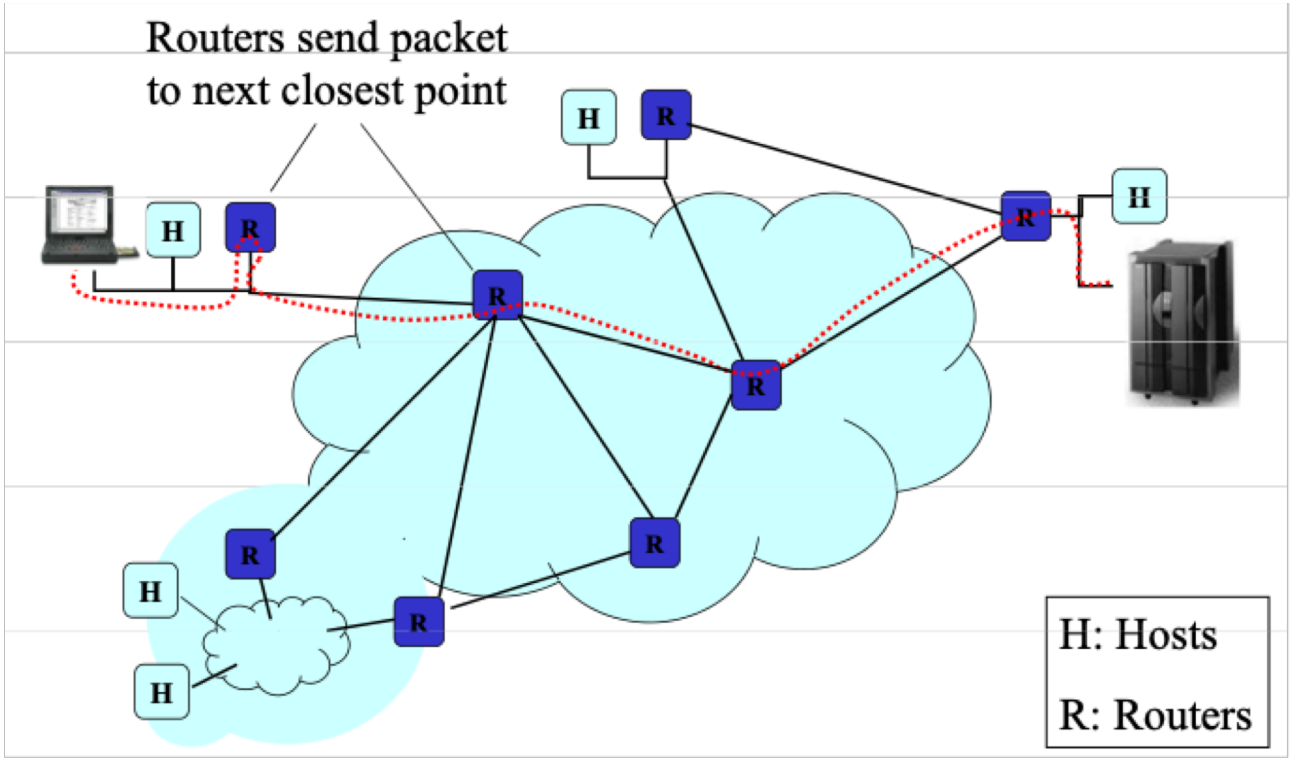


How does anyone find anyone else on the Internet?

- Each node on the Internet is assigned a unique identifier (IP address) (not really, but acceptable simplification)
- IP address are numerical and hard to remember - hence public nodes are typically associated with a mnemonic: a string in the **domain name system (DNS)**

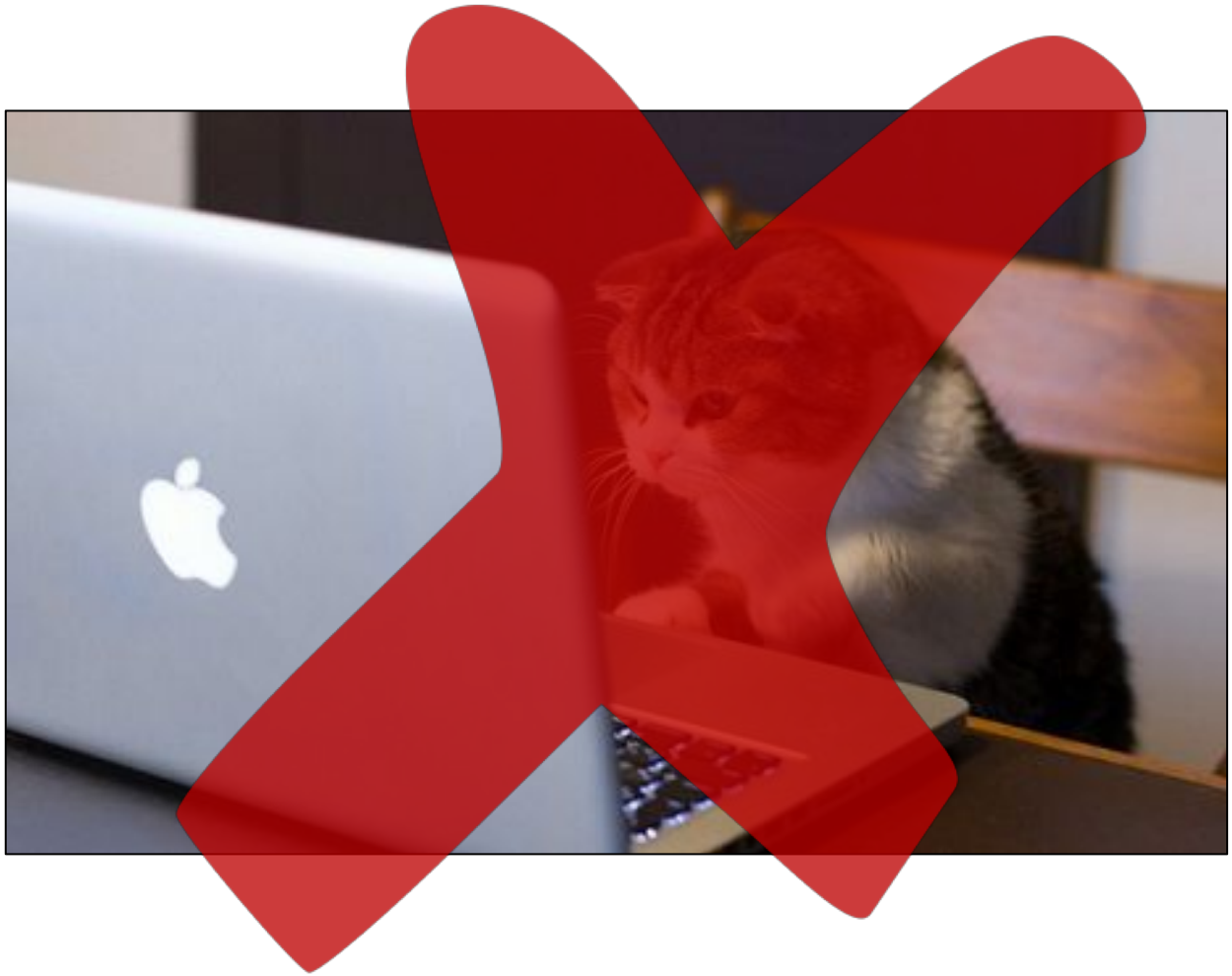


I have the IP address, now?



What Was the Point of the Internet?

- Cat videos!



What Was the Point of the Internet?

- Multiplexed communication over **existing** networks.
  - In particular, the ARPA packet radio network
- Each network had separate administration team.
- The interconnection of these networks was to yield The Internet.
- The creation of the Internet was due to the pragmatic decision to maximize utility of infrastructure.

How is this reflected in the design?

- Gateways
- Packet switching
- Store-and-forward
- Minimal assumptions about the underlying network technology

Other goals

Goal #1: survivability

- Can you explain me **what it means**?
  - Communication continues despite failures.
- Can you tell me **why it was so important**?

They SHOULD Be Shot–the DOGS!

Berlin, Germany.—(AP)—The mayor of a small upper-Bavarian town found something had to be done about poaching dogs. So he had signs put up, warning dog owners "Those, who have dogs and run around without leashes, will be shot."

A critical citizen said it should be made clear that the dogs would be shot, and not the owners. The mayor had the sign changed to read: "Those, who have dogs and run around without leashes, will be shot—the dogs."

THE DEMOCRAT
AND LEADER

94TH YEAR—No. 298. DAVENPORT, IOWA, FRIDAY EVENING, SEPTEMBER 23, 1949. 26 PAGES PRICE 5 CENTS

Weather
Fair and warmer Saturday.
(See page 23 for details).

RUSSIA HAS ATOMIC BOMB!

U. S. Learns of Recent Explosion, President Discloses

Pentagon Shows Excitement, But Officials Mum

Bargaining In Steel Resumed

Strike Clouds Over Industry Are Lightened.

Pittsburgh. — (AP) — Big Steel and the CIO United Steelworkers went back to the bargaining table Friday in an effort to settle their insurance-pension war and wipe out the threat of a nationwide strike.

Union Chieftain Philip Murray and Vice President John A. Stephens of U. S. Steel corporation got together in Hotel William Penn to resume negotiations under a third presidential strike truce.

Both declined comment as they went into conference behind closed doors. Negotiating teams of 30 to

Russ Possession of Bomb Peace Threat, Congressmen Believe

Washington.—(AP)— Senator McMahon (D-Conn) said Friday that if President Truman had announced an atomic explosion in Russia sooner he could have been pressed of rushing pressure o nike senate for the foreign arms bill.

The chairman of the congressional atomic committee offered that argument in the senate in answer to a complaint from Senator Cain (R-Wash) that the senate should have had the information Thursday. The arms measure was passed Thursday night.

Washington.—(AP)— There was one point of fairly general agreement Friday in early congressional reaction to news that Russia apparently has the atom bomb—the threat to peace is real.

On what to do about it there were about as many opinions

ALCOA Fights Bettendorf's Adoption Plan

Prepares for Every Legal Step Needed to Prevent Annexation.

For the first time Friday the Aluminum Company of America plant east of Bettendorf openly entered into the fight against the city of Bettendorf's plan to annex a large area east of the city.

U. S. Still Ahead of Russians

No Reason Seen To Believe War May Be Nearer.

Washington. —(AP)— The United States has lost its monopoly on atomic weapons but is still four years and many accomplishments ahead of the Russians.

Those are the standout facts in estimates by the best informed government officials of the impact from Russia's atomic explosion on the balance of power between Russia and the United States.

State department officials generally took the line that they did not believe the danger of

JOHNSON QUESTIONED ON ATOM BLAST—Secretary of Defense Louis Johnson, (right) holds his jaw as he is questioned by white house newsmen after leaving a cabinet meeting Friday at which President Truman

Details on Blast in USSR Not Revealed; Vishinsky Urges Pact

No Change Made in Disposition of American Forces, Johnson Declares; Political Repercussions Expected.

New York. — (AP) — Soviet Foreign Minister Andrei Y. Vishinsky called on the Big Five powers Friday to conclude a peace pact among themselves.

He made no mention of an atomic explosion in the Soviet Union in his general policy address to the United Nations assembly.

Vishinsky prefaced his peace proposal with his usual slashing attack on the western powers.

He charged the United States and Britain are leading plans for an aggressive war.

Washington.—(AP)—The United States has evidence of a recent atomic explosion in Russia—news indicating the Communists at long last have learned to make an A-bomb.

President Truman disclosed this in a statement Friday. He then held an hour-long session with his cabinet about it.

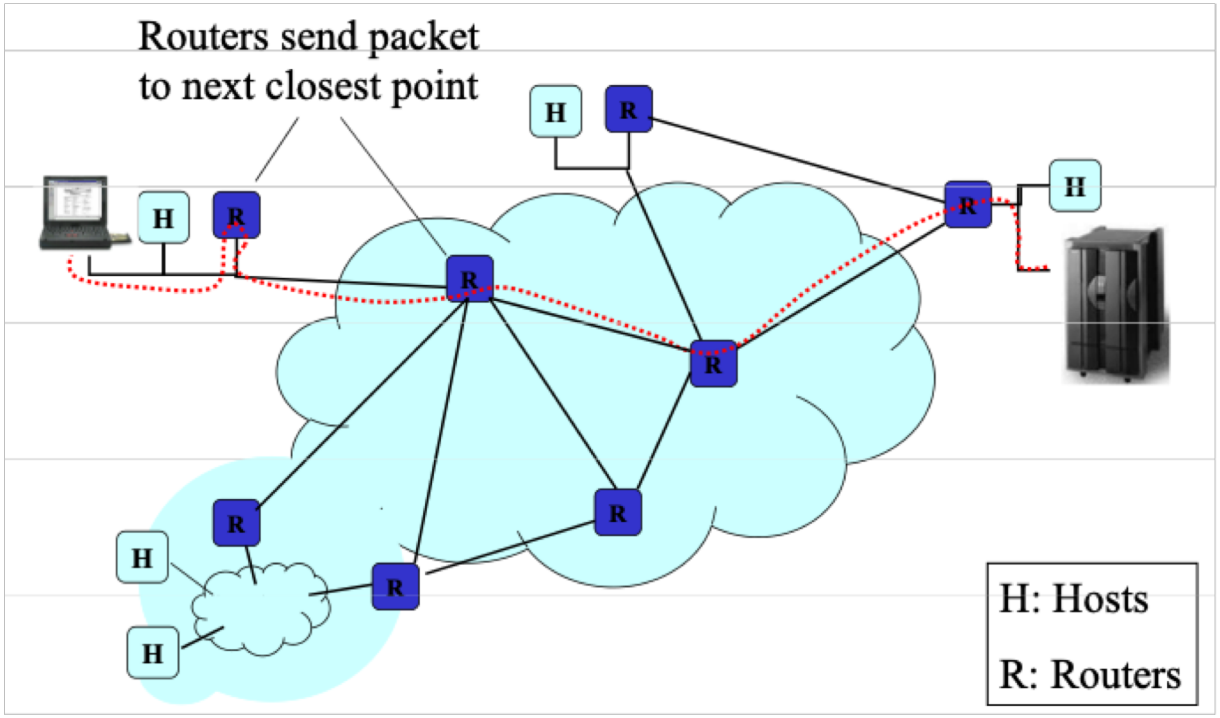
Mr. Truman said the development emphasizes the necessity for "truly effective, enforceable, international control of

Survivability

- Provide communication service **even if gateways fail.**
  - **Temporary failures** should not require **reconnection of flows.**
- …which in practice, means that:
  - (i) The network must be able to **route around failures**
  - (ii) The **state** pertaining a communication flow **must be protected**
- Only a complete partition should deny connectivity.

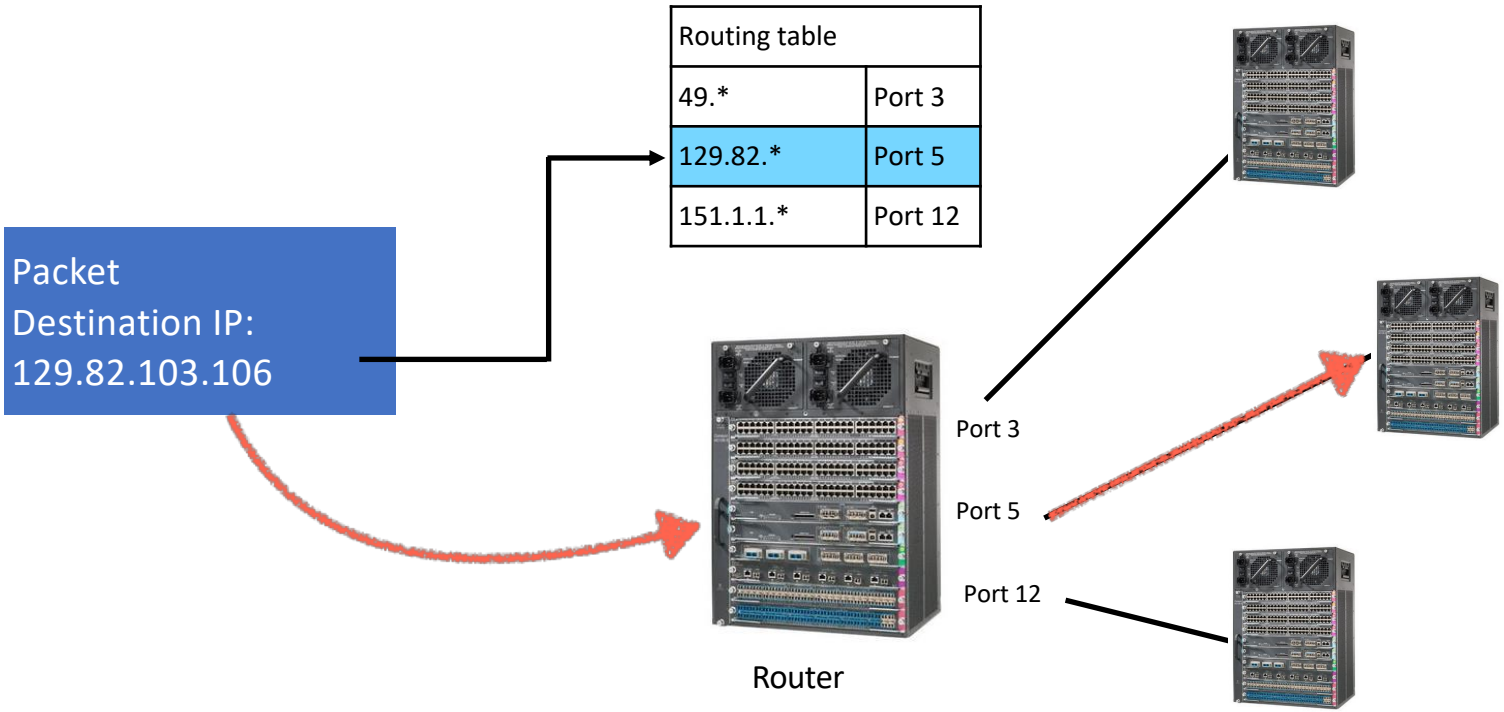
How does the network route around failures?

- Can you give me an overview?



Yes, but how does it work?

- Routers maintain forwarding tables which specify how to reach all portions of the assigned public IP address space



What happens if a link goes down?

- Routers typically store **multiple routes** w/ different costs **for a given destination**
- If route through an endpoint becomes unavailable, simply **select another one**
- **Lots of underlying complexity** to ensure convergence, avoid loops, etc.!

Protecting state: fate sharing

- **Issue:** oftentimes carrying a communication between two nodes requires maintaining state (think of TCP connections) that must be stored somewhere
  - **In the network:** reliability issue: what if network nodes (switches, routers) malfunction?
  - **In the host:** the state is only lost if the host goes down - but this is acceptable, because implies that the state is lost together with the entity using it. This principle is called **fate sharing**.

Fate sharing

- What if I really really want to **maintain some state in the network**?
  - For example?
- It is acceptable to maintain some communication state in the network…
  - **…as long as it can be easily recreated if lost**
- **Idea:** endpoints install state in the network; network can **discard state at any time for any reason**, in which case **endpoints can reinstall it**
- Example: **RSVP**: receiver uses it to reserve resources across network; if network conditions change and state is lost, the receiver reinitialize the state

Implications

- What is the **main implication** of fate-sharing and soft state?
- Intermediate network links **can fail without affecting communication** between endpoints!
- In fact, **there is no way for IP to report network malfunctions to upper layers**

Goal #2: multiple types of communication services

- It does not make sense to build a complex, massive network infrastructure if then we can only run one application on top of it
- Therefore, a top Internet goal was to **support multiple types of communication services**

Types of services

- Virtual Circuit
  - Transmission Control Protocol (TCP)
- Network debugging (XNET)
- Digitized speech
  - Timeliness and regular intervals, but loss okay
- User Datagram Protocol
  - For those who want to build it all themselves
- **Easier to add reliability to a datagram network than to convert reliable to datagram**

Simplifying…

- There are **two broad classes of services**:
  - Applications that need **reliable transport** (“virtual circuit”)
  - Applications that are willing to **sacrifice reliability for better performance** (AKA less jittering, higher bandwidth, lower delay)
  - The latter are easier to accommodate than the former (just don’t do anything special!)
- How would you build a network that **supports both**?

Datagram-based delivery

- **Easier to add reliability to a datagram network than to convert reliable to datagram**
  - Can you explain this in terms of the **end-to-end principle**?
    - Hint: reliability is an application-level concern
- So, build a **datagram-based network**…
  - …and let applications **directly access the datagram primitive** if they want **performance** (**UDP**)
  - …then, build a **middleware on top of datagrams** that implements **in-order, error-free delivery** (**TCP**)

Effects of not making reliable transport a network primitive

- Originally there was **only one combined network+transport protocol:** NCP
- …then, it was decided that NCP was doing too much, for the reasons described above
- So, it was decided to **replace NCP with TCP/IP**
- How? Any idea?
  - 1983: **Internet flag day**!
  - Basically, a reboot of the entire network

Some other aspects of TCP

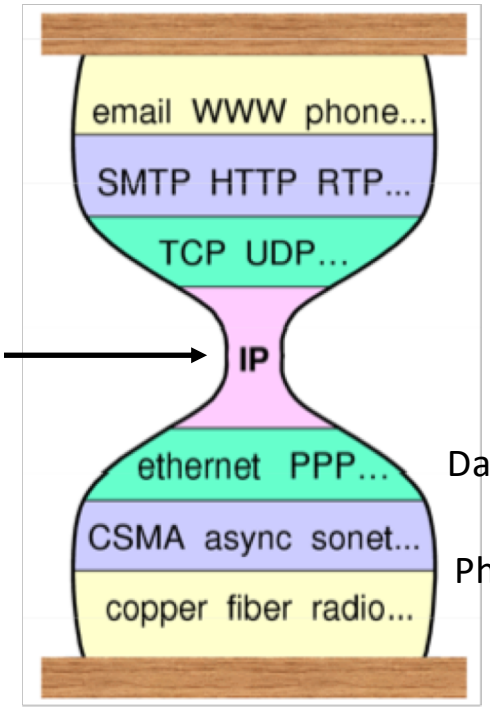
- **Flow control** regulates byte delivery
  - No significance **in packet segmentation**
- Reasons
  - Enables **fragmentation** (but later moved to IP, so now we have both!)
  - Byte-based segmentation useful in **retransmissions** when using small packets (enables merging a bunch of small packets into one when retransmitting)

Goal #3: support for heterogeneous network technologies

- This is **related to**, but **different from** the goal of connecting different existing networks
- Explain?
  - Connecting different existing network was a **"political" goal**: networks already existed, and it did not make sense, in terms of economics and opportunity, to restart from scratch
  - Supporting different network technologies is a **technical goal**: means that the Internet must be designed from the ground up to support as many different data-link layers as possible

Hourglass design

- Support for **heterogeneous network technologies** is implemented in the **design of the IP protocol**
- The hourglass design:



Application (high-level data stream)

Transport (reliable transmission of packets)

IP provides datagram abstraction, on top of which more complex services can be build

Network (routing across multiple networks)

Data-link (transmission of frames on link)

Minimal assumptions about underlying medium

Physical (characteristics of physical signal)

Varied networks

- LANs (Ethernet, rings)
- Satellites
- Radios
- Serial links like T1s
- **How can all those be supported?**
- Versatility by **not assuming much about the medium**
  - It can transport bytes (e.g. 100 byte min)
  - Reasonably reliable, but not perfect
  - Addressing scheme if multiple hosts

Any other assumption?

- Not really!
- What are the things that are **not** expected/required?
- Some examples:
  - Reliable delivery
  - In-order delivery
  - Broadcast/multicast
  - Explicit representation of link properties
  - Error notification

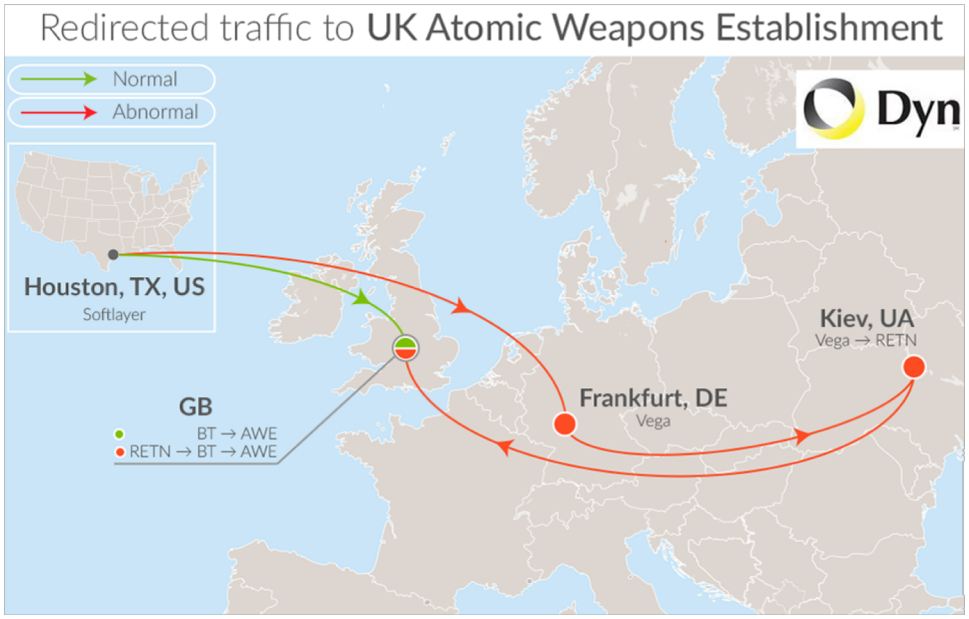
Properties that were not so relevant/important

Distributed management

- In a sense, this goal has been implemented fairly well…
  - The Internet is (and was designed to be) a graph of **autonomous systems** (AS), which are in general **independently managed entities**
- But there are problems!
  - **Routing policies are still hard to enforce** in a robust and secure way

Robust distributed management?

- This is beyond the scope of the paper, but…



https://dyn.com/blog/uk-traffic-diverted-ukraine/

- BGP hijacking is a real problem!

Cost-effectiveness

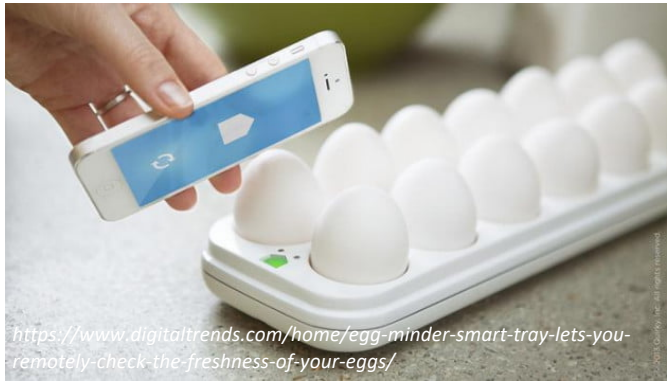
- Some design choices are somewhat **suboptimal** in terms of **efficiency**
- Chiefly, header size of TCP packets (20 bytes for IP header + 20 bytes for TCP header) is quite large
  - Especially considering that this ”header penalty” must be paid even when sending 1 byte of data

Cost-effectiveness II

- This was a more significant issue in the early day of the Internet due to:
  - Limited line rates
  - Prevalence of interactive traffic (Telnet)
- Nowadays, neither factor is very significant

Ease of adding a host

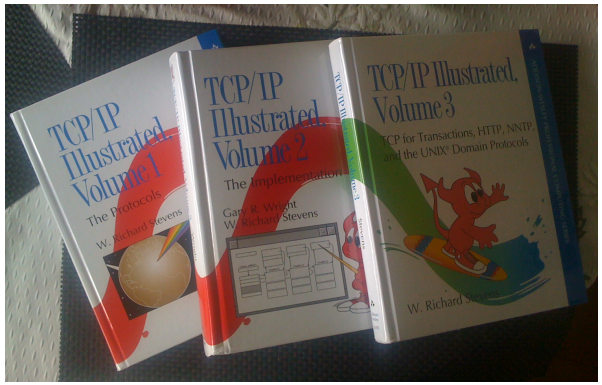
- Nowadays, even egg trays have network connectivity…



(yes, it was a real product)

(because reading the expiration date on the box is **so** 20th century)

- So we tend to forget that a **TCP/IP stack** is really **complex** and **hard to implement right**!



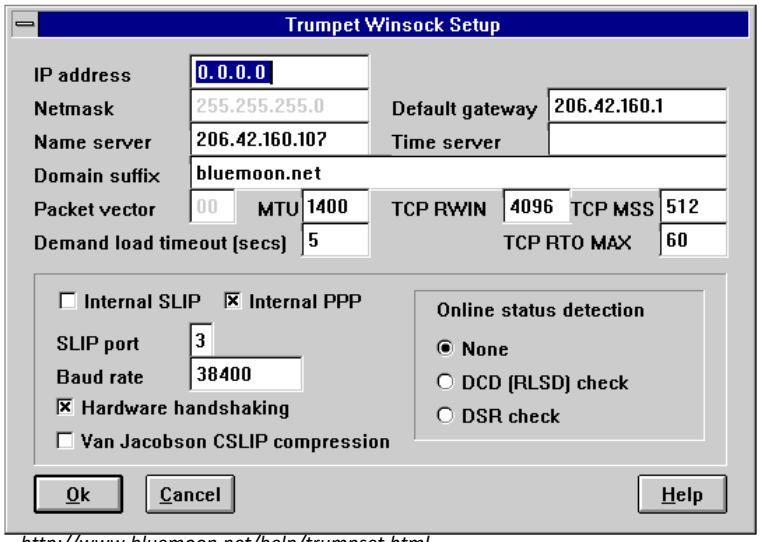
http://lispmeister.github.io/posts/2012/02/tcpip-illustrated-vol-1-3/

Ease of adding a host - II

- Much of the complexity in ”talking to the Internet” is in TCP, which is a **end-host protocol**
- Therefore, much on the **implementation burden** is on the **end-host designer**

Aside

- Microsoft did not release a TCP/IP stack with their OS until 1994…
- …so if you wanted to do some browsing before that, you had to buy and install a separate commercial stack



http://www.bluemoon.net/help/trumpset.html

Aside - II

- The idea that a software company may consider TCP/IP not worth the effort may sound strange, but in the early days of the Internet it was not clear that this was **the** network technology to rule them all
- *"In this time when Microsoft and Novell were duking it out over proprietary LAN technologies and the third-party Winsock stack market was prospering, TCP/IP was this weird thing that was mostly used by minicomputers and expensive workstations running Unix. It was the technological glue that held together this equally strange thing called the Internet that few businesses and almost no individual users had used yet, if they'd heard of it at all. The rise of TCP/IP and the Internet was so peripheral to the main PC industry that Microsoft almost completely ignored it until the second half of the 1990s." (Warren Young, https://tangentsoft.net/wskfaq/articles/history.html)*

Accountability

- In other words, **keeping track of the amount of network resources** used by each node
- Yeah, the Internet is not really designed for that…
- …and in general for a lot of things that relate to **monitoring** and **understanding what Internet users are doing**
- …will talk about this again when we discuss **net neutrality**!

Take-away points

- The Internet we have **depends on the order in which required properties were considered**
- A **different order** (or different properties altogether!) would have resulted in a **different Internet**
- Different Internets are **possible**, but **hard to implement!**

Example of alternative paradigm

- **Named Data Networking (NDN):** identify data, not host
  - Because basically nobody cares about which specific server you are talking to anymore…
  - …we just care about **reaching the content we want**
- In a nutshell:
  - **Names identify resources** (content)
  - To reach a piece of content, I send an **interest** for that content
  - Routers **route the interest** until it reaches the **provider of that content**, which send the content back
  - Routers can **cache content in the network** to make communication more efficient
- Current versions run on top of TCP/IP

Additional readings

- Some Internet (and TCP/IP) history:
  - https://www.internetsociety.org/internet/history-internet/
  - https://en.wikipedia.org/wiki/History_of_the_Internet
  - https://tangentsoft.net/wskfaq/articles/history.html
- Named data networking:
  - Zhang et al., “Named Data Networking”, SIGCOMM CCR, July 2014