Understanding the Internet

Allen B. Downey

An engineering approach

The Internet is an engineered artifact.

- Designed under constraints for a purpose.

Therefore to understand the Internet:

- Understand constraints.
- Understand purposes.
- Understand solutions (and create new ones)!

A scientific approach

The Internet is a complex system.

- Understanding components is NBNS.
- Understanding interactions among components is NBNS.

Understanding requires:

- Appropriate abstraction.
- Observation of emergent properties.

Traffic jams

Example: automobile traffic is a complex system.

What is a traffic jam?

- Made of cars, but not always the same cars.
- It's an abstract entity, like a storm front.
- It has properties that its components don't have.

Traffic jams move backwards.

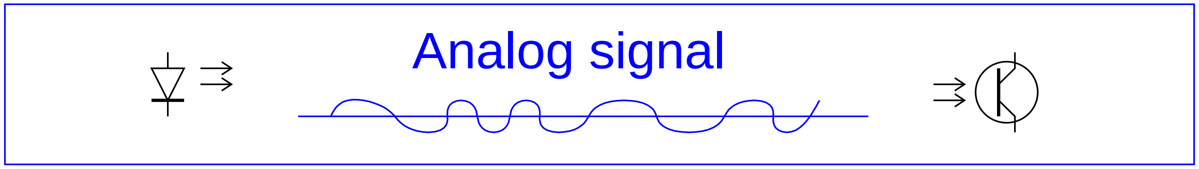
Outline

Exercise these viewpoints:

- The bottom-up Internet.
- The emergent Internet.

Physical layer

Starting with a medium that transmits an analog signal...



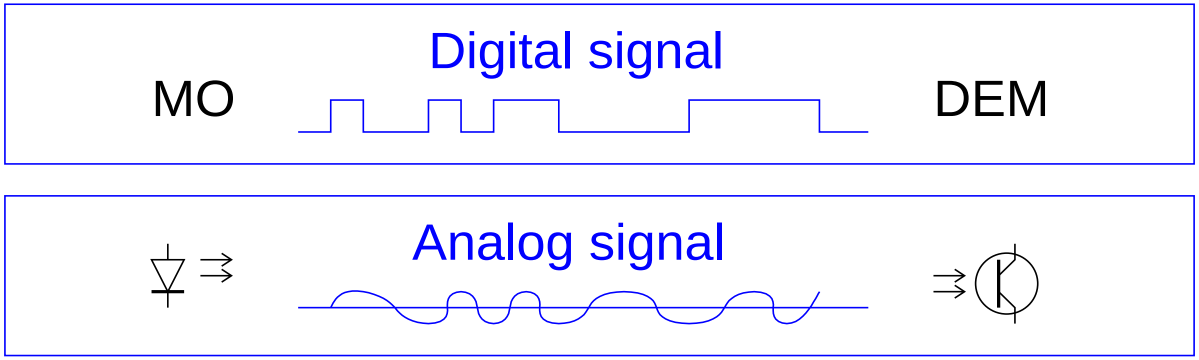
...how can we transmit a digital signal?

Solution: modulation.

- Vary amplitude, frequency, etc.
- Distinguish (at least) two levels.

Signal layer

Physical layer + mo/dem = Signal layer



Abstractly, a modem transmits a digital signal.

Physically, there's no such thing!

HI and LO are abstract symbols with many possible physical representations.

Transmitting bits

- A bit can have two states, 0 or 1.
- A digital signal has two states, LO and HI.
- Naive encoding: LO represents 0, HI represents 1.

Problem?

Transmitting bits

- A bit can have two states, 0 or 1.
- A digital signal has two states, LO and HI.
- Naive encoding: LO represents 0, HI represents 1.

Problem?

1. Clock recovery.
2. Threshold discovery.

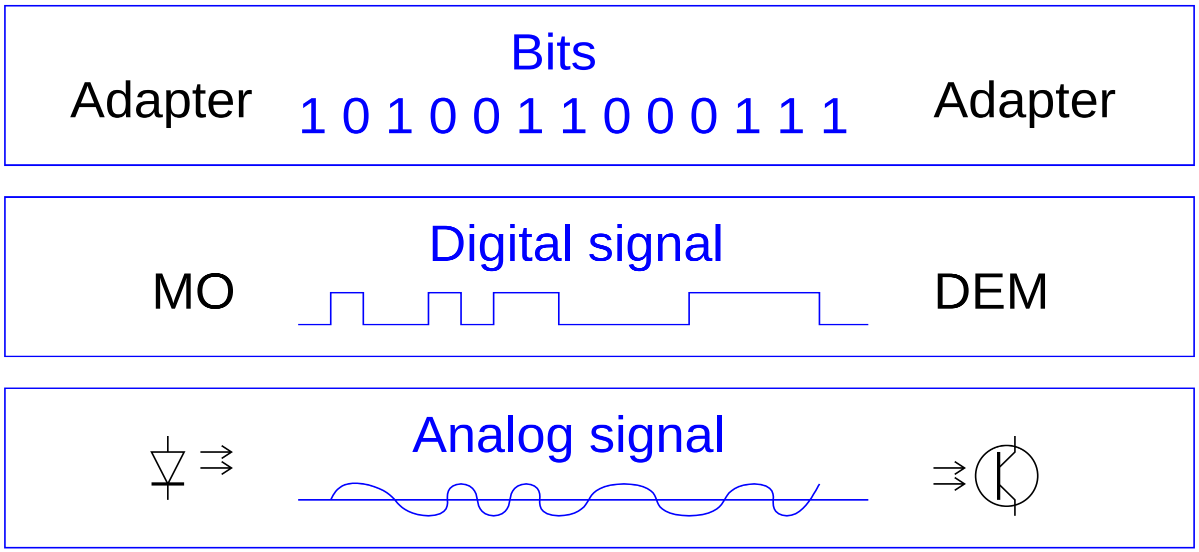
Manchester encoding

LO-HI represents 0.
HI-LO represents 1.

- Lots of transitions; good for clock recovery.
- Equal number of LO and HI; good for threshold discovery.

Bit layer

Signal layer + encoding = Bit layer.



Abstractly, an adapter transmits bits.

Physically, there’s no such thing!

0 and 1 are abstract symbols with many possible physical representations.

Error detection

How do you know the bits you got are right?

- Group bits into frames.
- Attach redundant information to each frame.

How do we identify the beginning and end of a frame?

Framing

A frame consists of:

01111110	data	checksum	01111110

- Special sequence marks beginning and end.
- After the data, send a checksum.
- Checksum is a function of the data bits.

Problem?

Bit stuffing

What if 01111110 appears in the data?

Sender:

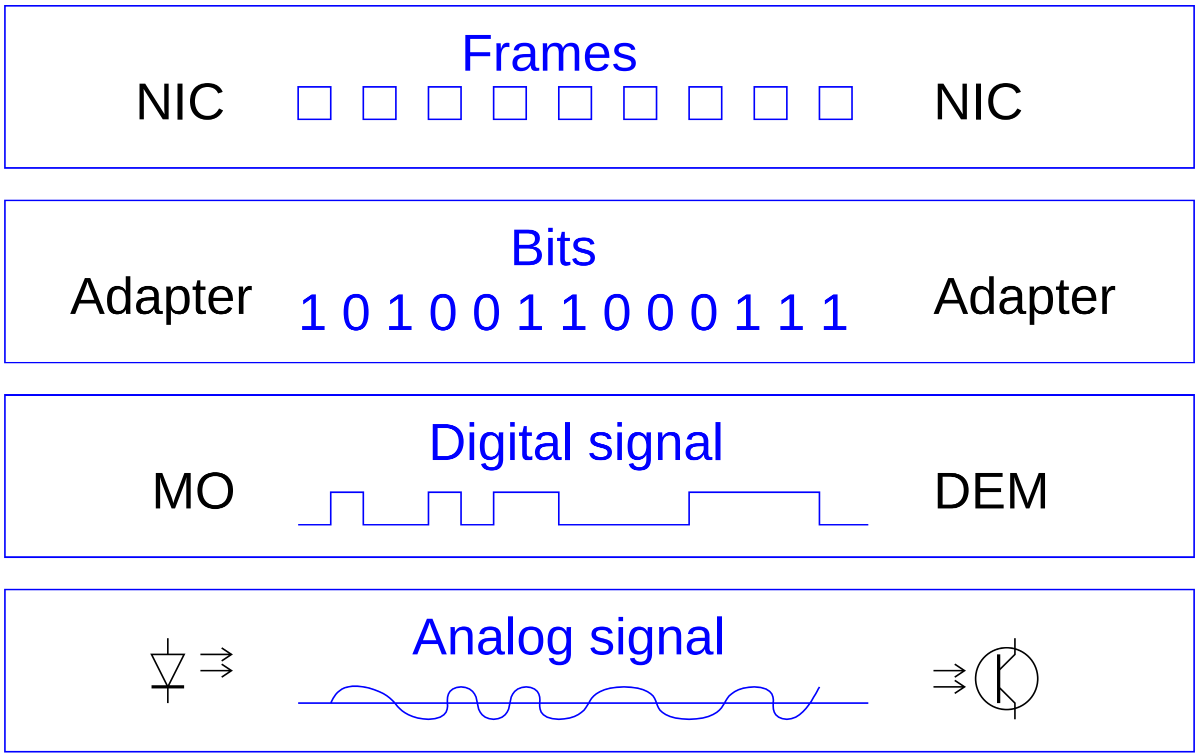
- Whenever you send 011111, stuff in a 0.

Receiver:

- Whenever you see 011111, check the next bit:
  - If 0, remove it.
  - If 1, check the next bit:
    - If 0, that's the end of frame.
    - If 1, there must have been an error!

Link layer

Bit layer + framing + error detection = Link layer.



Abstractly, a NIC transmits frames.

Physically, there's no such thing!

The Internet Protocol

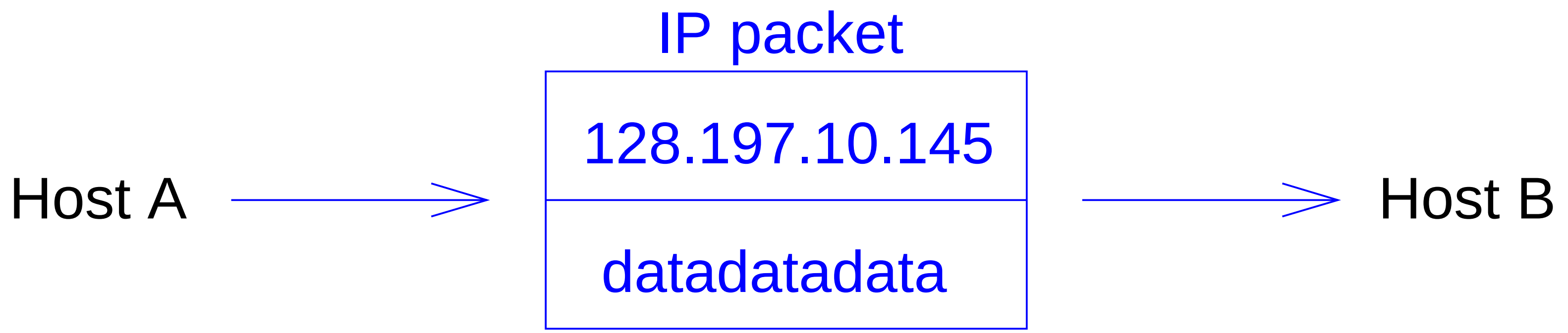
Links are point-to-point.

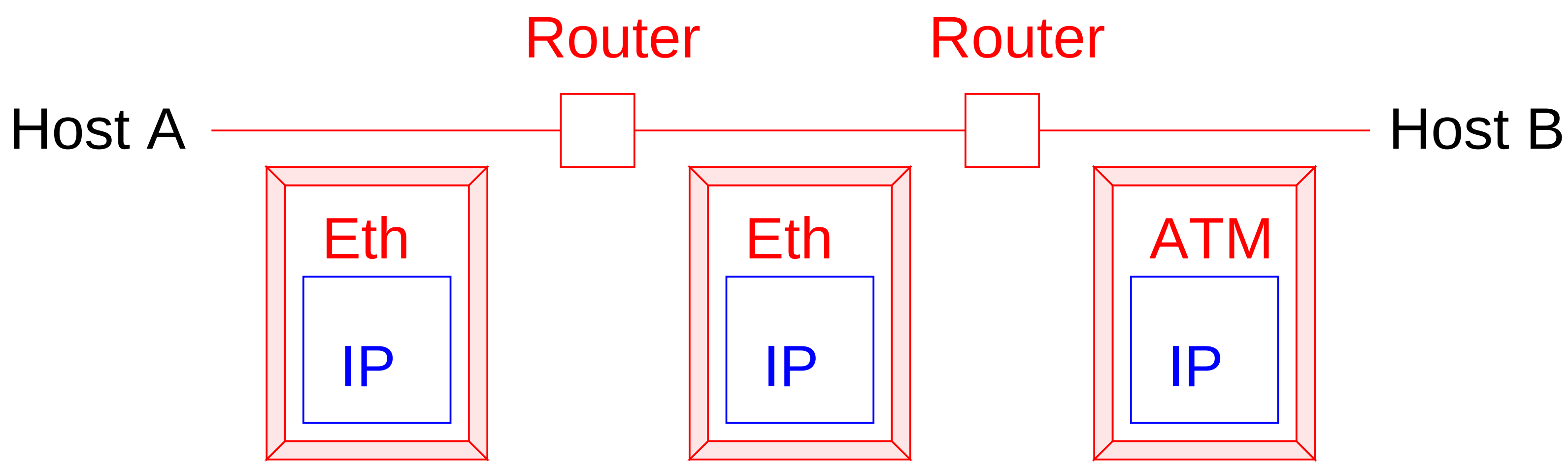
How do you talk to someone you’re not connected to?

- Assign IP addresses (128.197.10.145).
- Relay packets through routers.

What's a packet?

Abstractly, Host A transmits a packet to Host B.

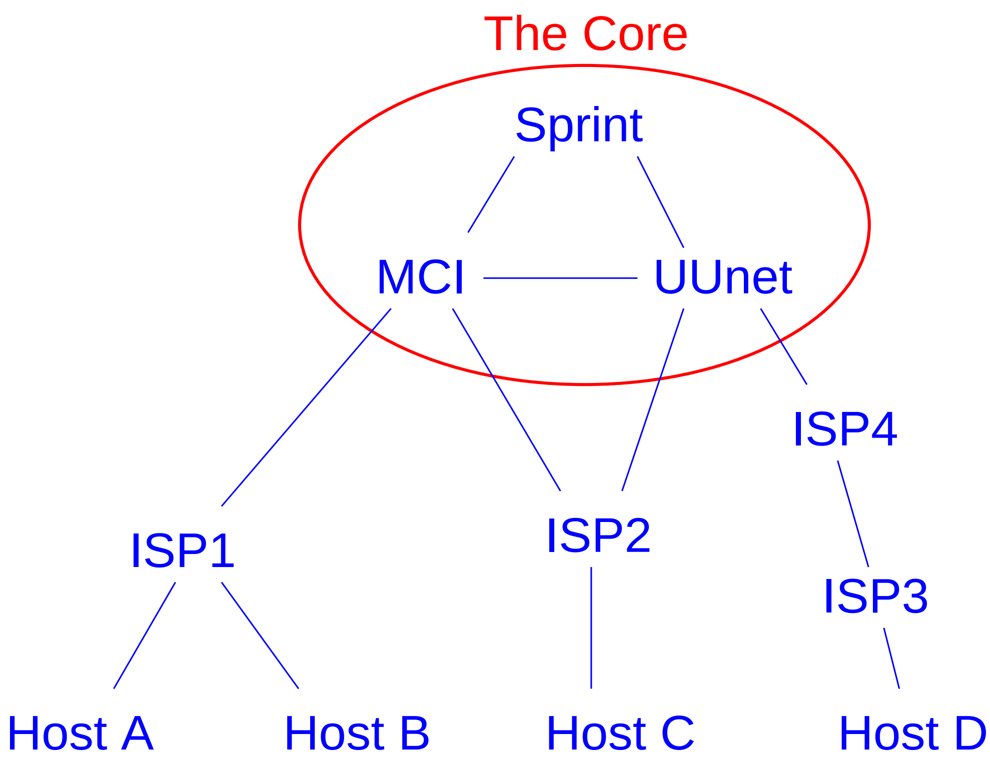




Physically, the packet's data are copied from link to link.

Internet structure

Physical structure is mostly hierarchical.



- Core providers completely connected.

A typical path

0 rosco.bu.edu (128.197.10.145)

1 csnet-gw (128.197.10.1)

2 extgw-dgw-cumm111 (128.197.254.122)

3 bos-edge-02.inet.qwest.net (63.145.1.77)

4 bos-core-02.inet.qwest.net (205.171.28.29)

5 jfk-core-01.inet.qwest.net (205.171.8.19)

6 jfk-core-03.inet.qwest.net (205.171.230.6)

7 jfk-brdr-02.inet.qwest.net (205.171.230.25)

8 so-4-0-3.nycmny1-hcr3.bbnplanet.net (4.25.133.29)

9 p10-0.nycmny1-nbr2.bbnplanet.net (4.0.7.14)

10 so-4-0-0.bstnma1-nbr2.bbnplanet.net (4.24.6.49)

11 p2-0.bstnma1-cr10.bbnplanet.net (4.24.10.62)

12 multilink1.folincollege2.bbnplanet.net (4.24.90.218)

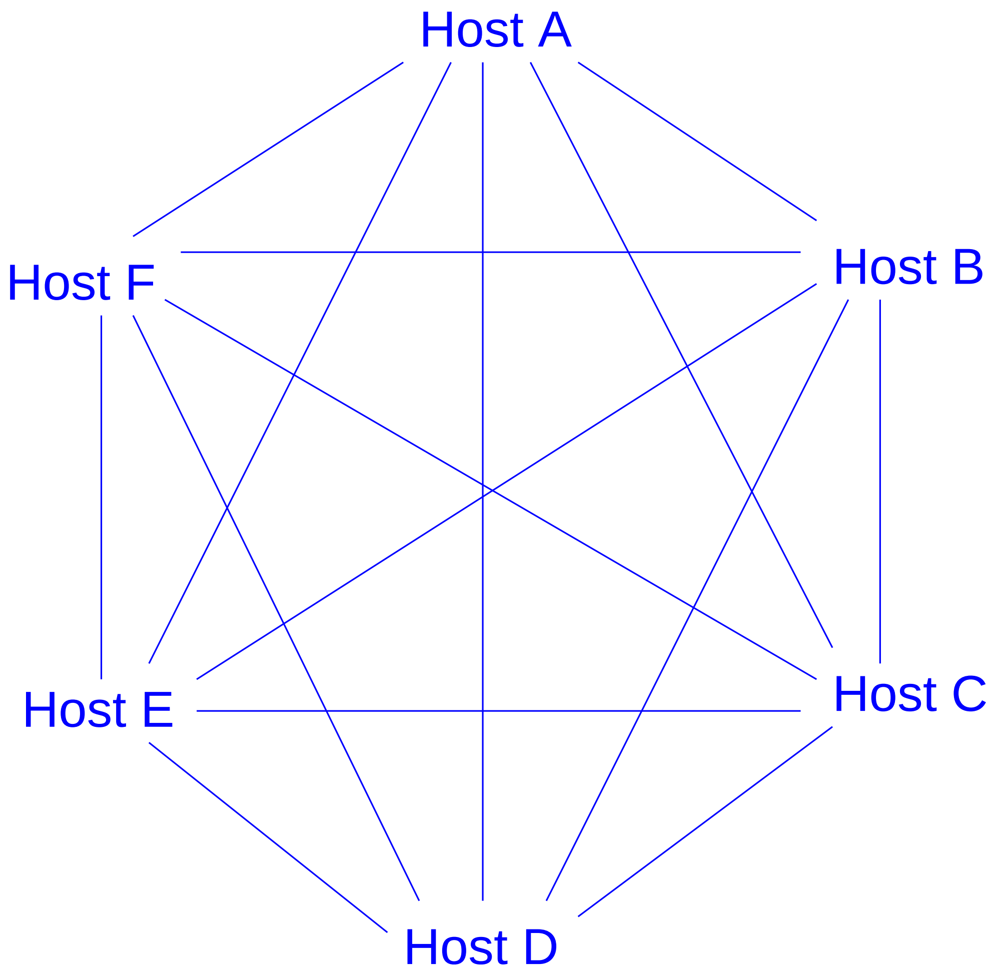
13 4.21.173.194 (4.21.173.194)

14 * * *

15 www.olin.edu (4.21.175.12)

Abstract structure

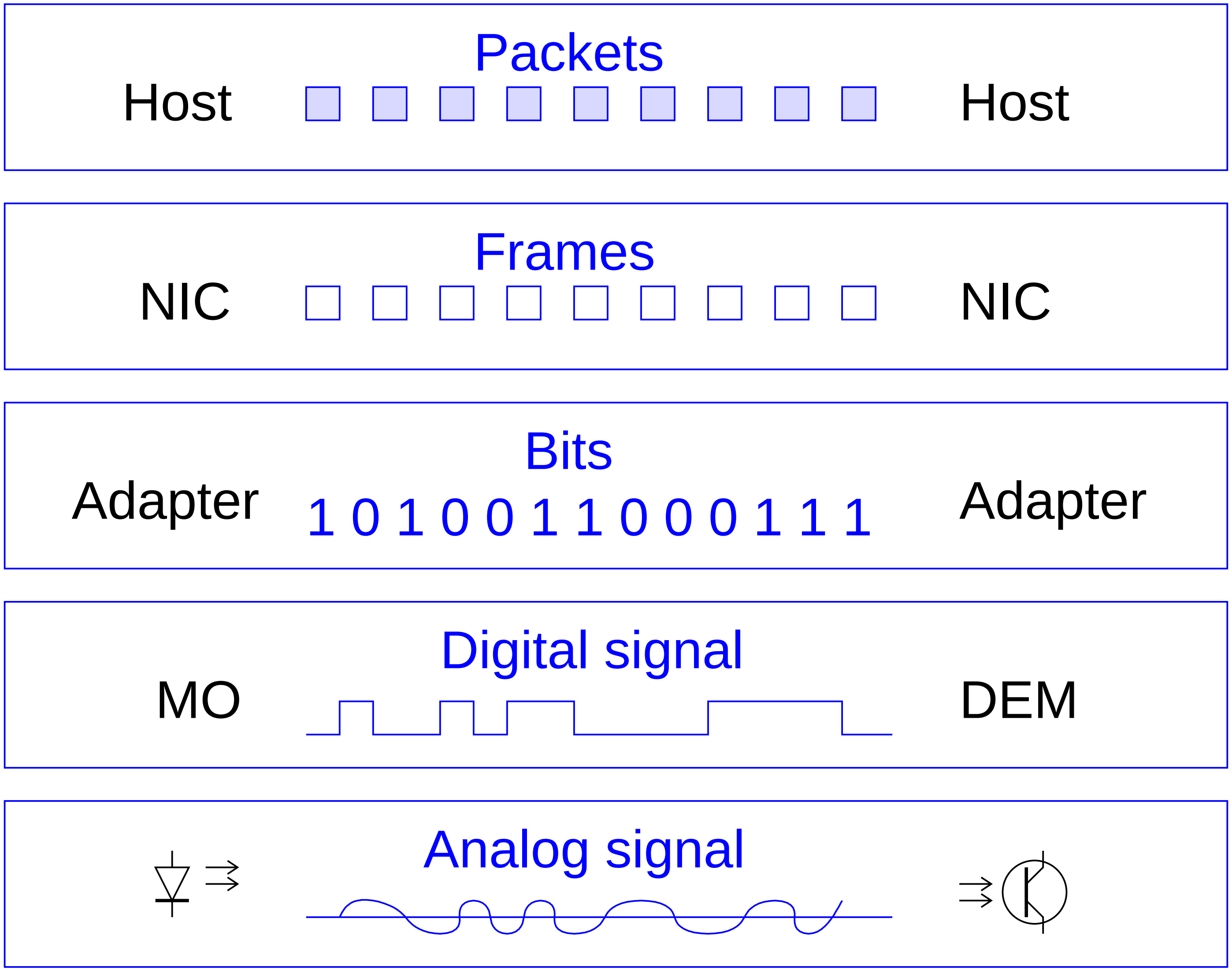
Fortunately, IP hides this complexity!



IP creates an abstract, completely connected network.

The Network Layer

Link layer + addressing + routing = Network Layer



Best effort delivery

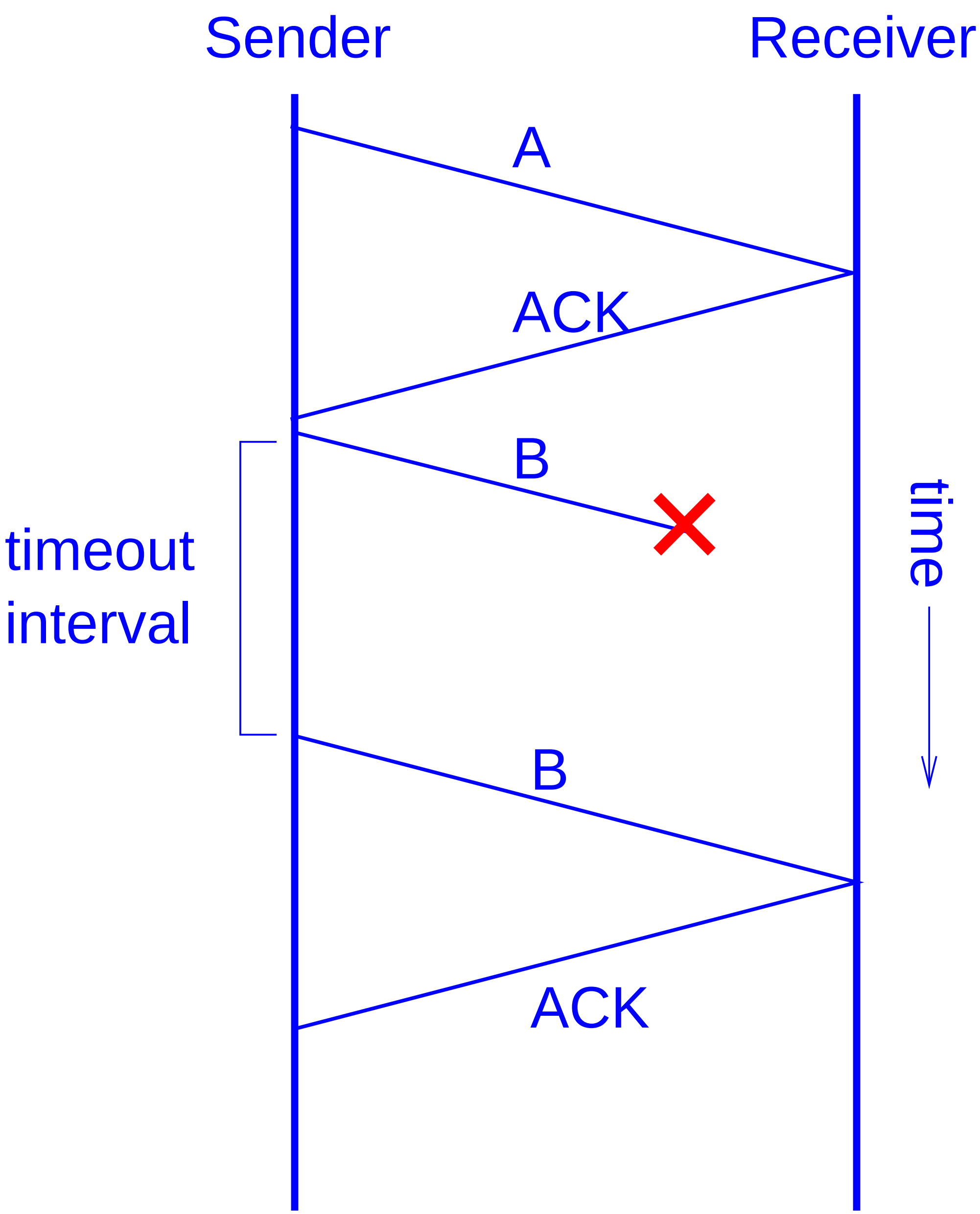
Absolutely, positively might get there eventually.

Who could ask for anything more?

- Exactly-once delivery.
- In-order delivery.

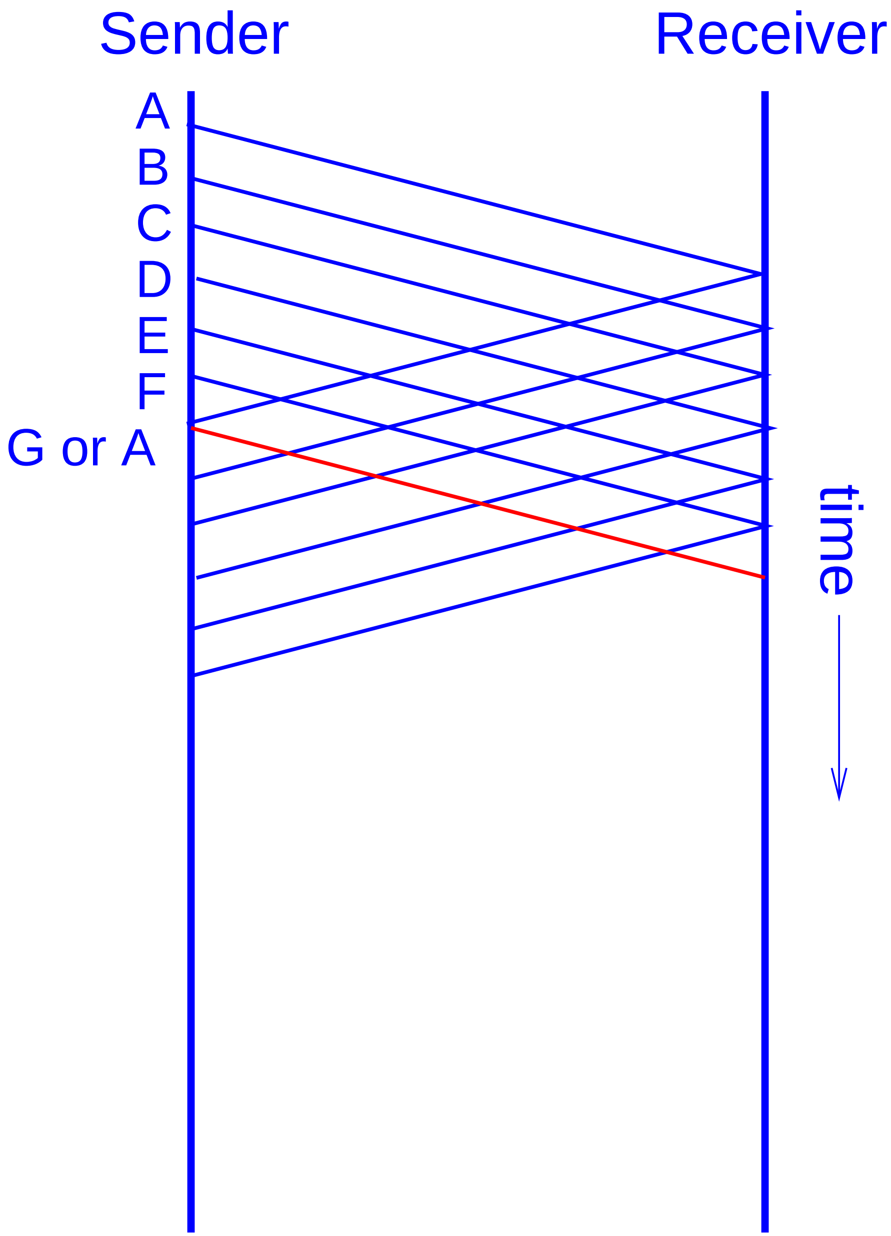
Can you build reliable communication on an unreliable medium?

Stop and wait



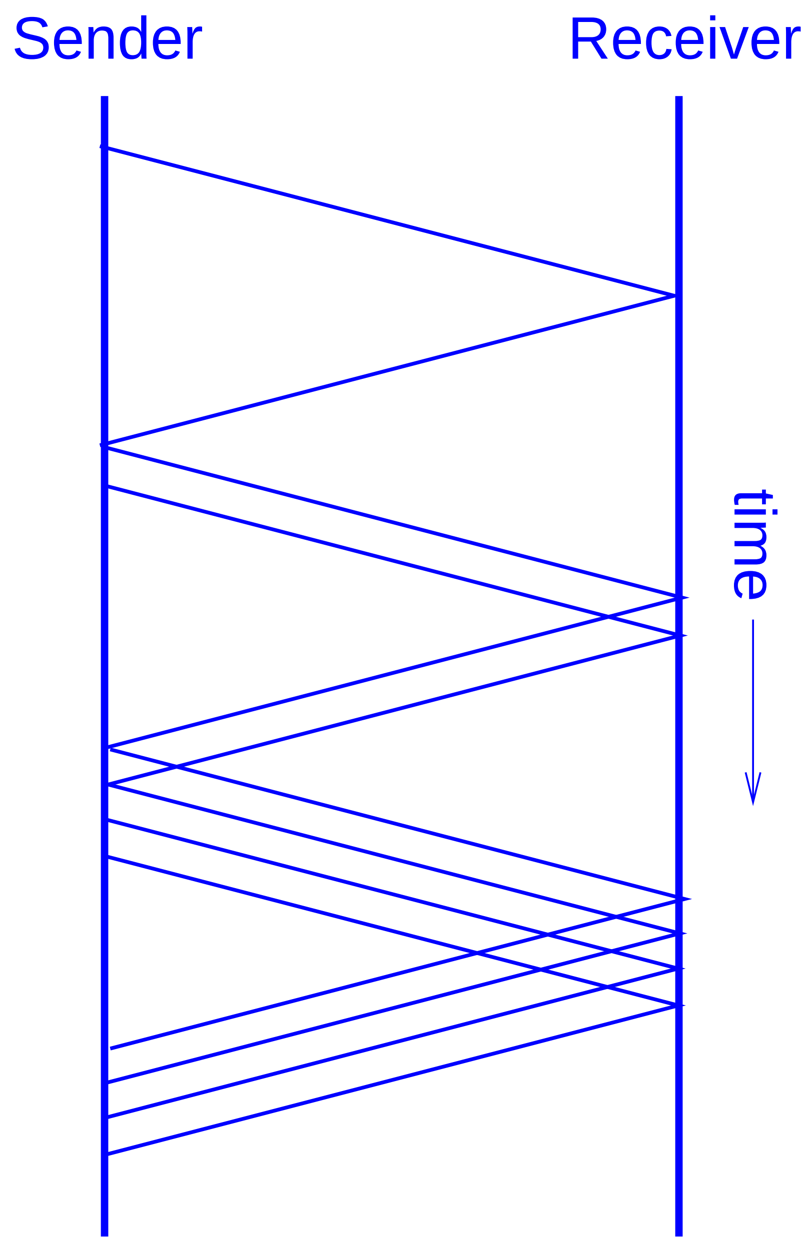
- Transmit.
- Acknowledge.
- Timeout... retransmit.
- Problem?

Sliding window



- Keep a copy of unACKed packets.
- Get an ACK, send a new packet.
- Miss an ACK, retransmit.
- Problem?

Slow start



- Start with one packet.
- Get an ACK, send two packets.
- Watch for a congestion signal.

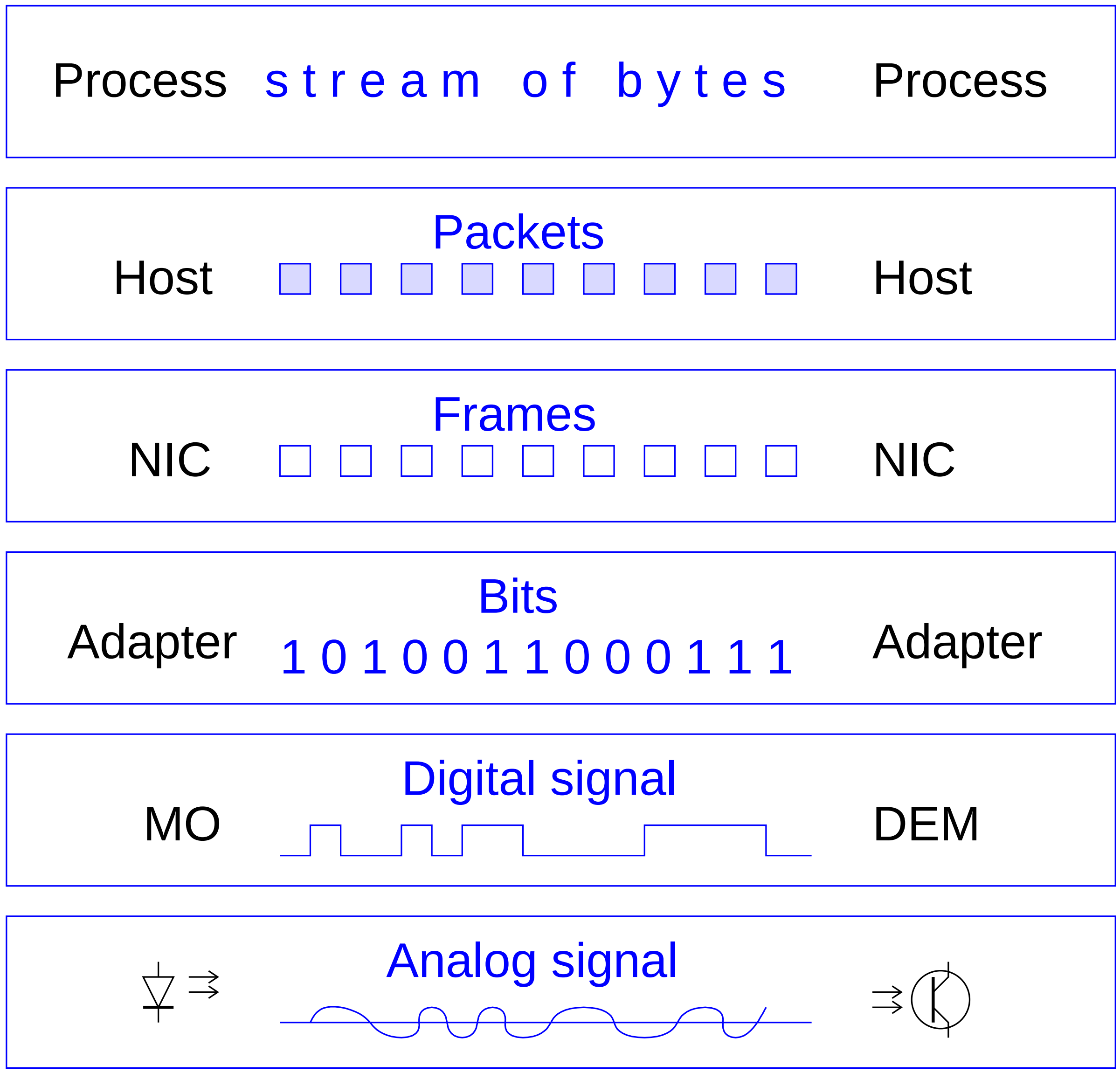
TCP

Transmission Control Protocol

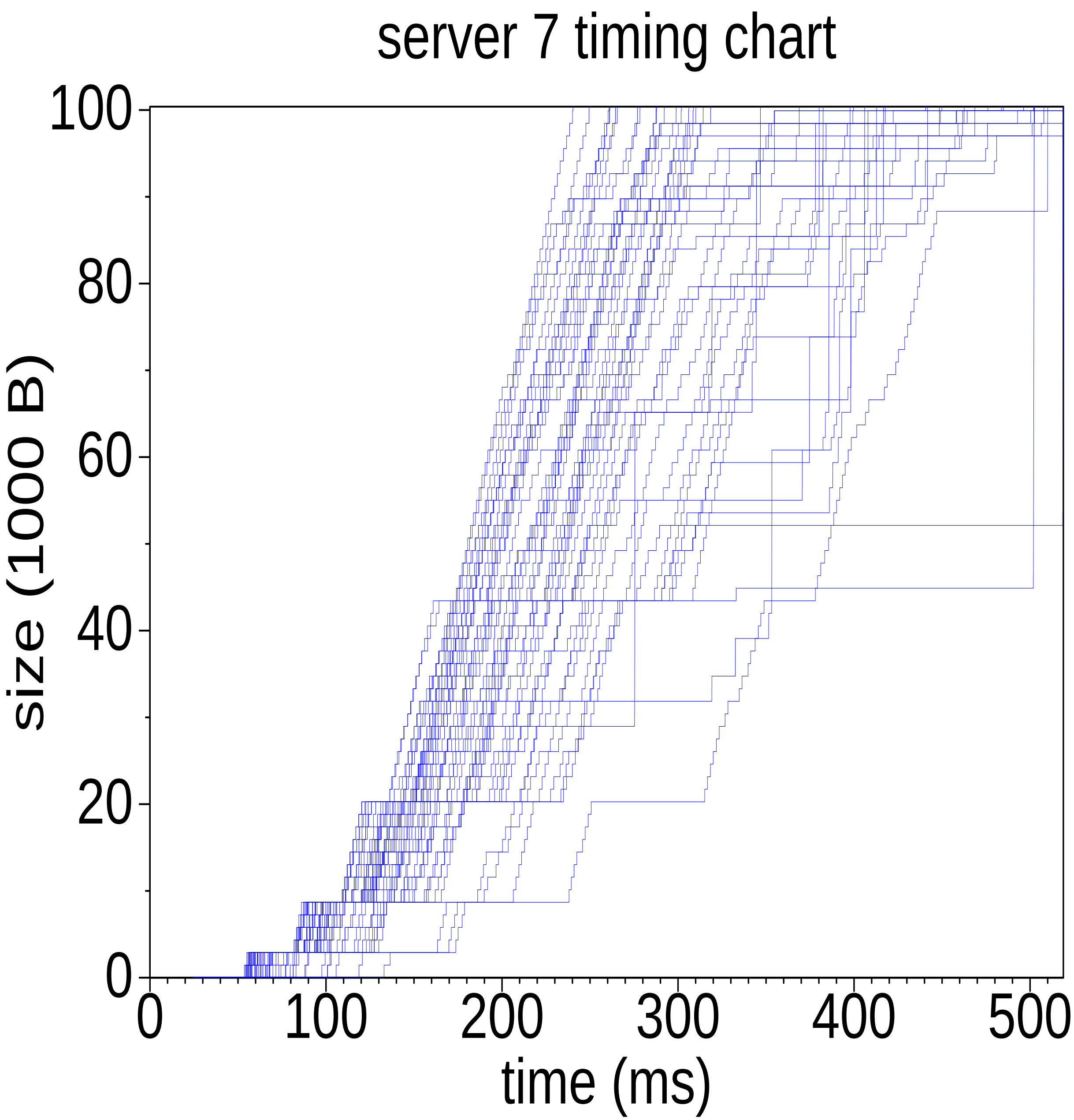
- Reliable delivery (ACK-retransmit).
- Flow control (sliding window).
- Congestion control (slow start).
- Byte stream abstraction.
- Process-to-process communication.

The Transport Layer

Network layer + sliding window + slow start = Transport layer



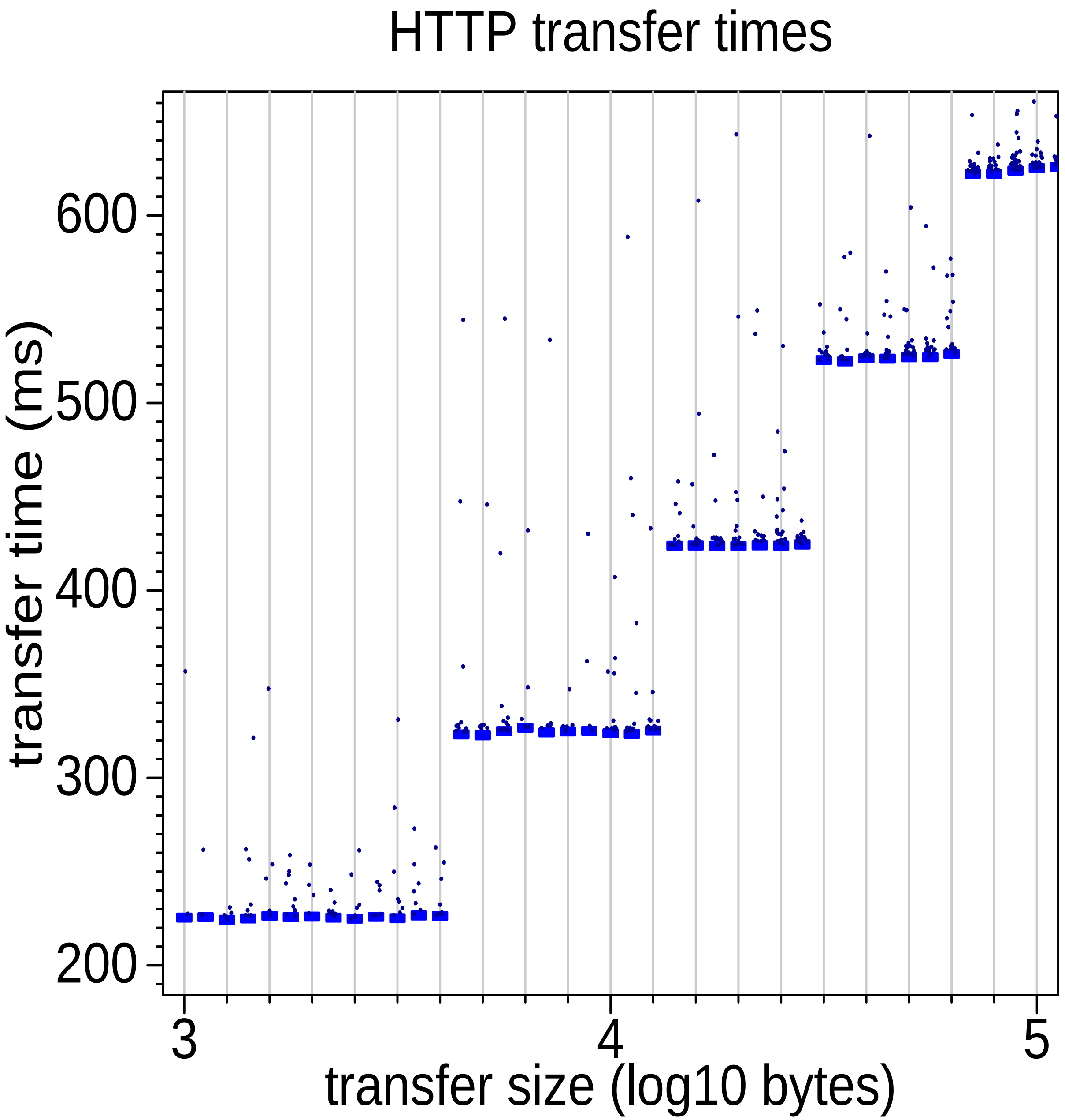
Modeling TCP Performance



- Observe network transfers.
- Estimate path parameters.
- Predict transfer times.

“Modeling TCP Performance,” Downey, 2003.

Modeling TCP Performance



- Transfer times for short files are multiples of the round trip time.

Who uses TCP?

90% of Internet traffic is based on TCP:

- TCP + HTTP = The Web
- TCP + SMTP = email
- TCP + gnutella protocol = peer-to-peer network

Application protocols

A little HTTP:

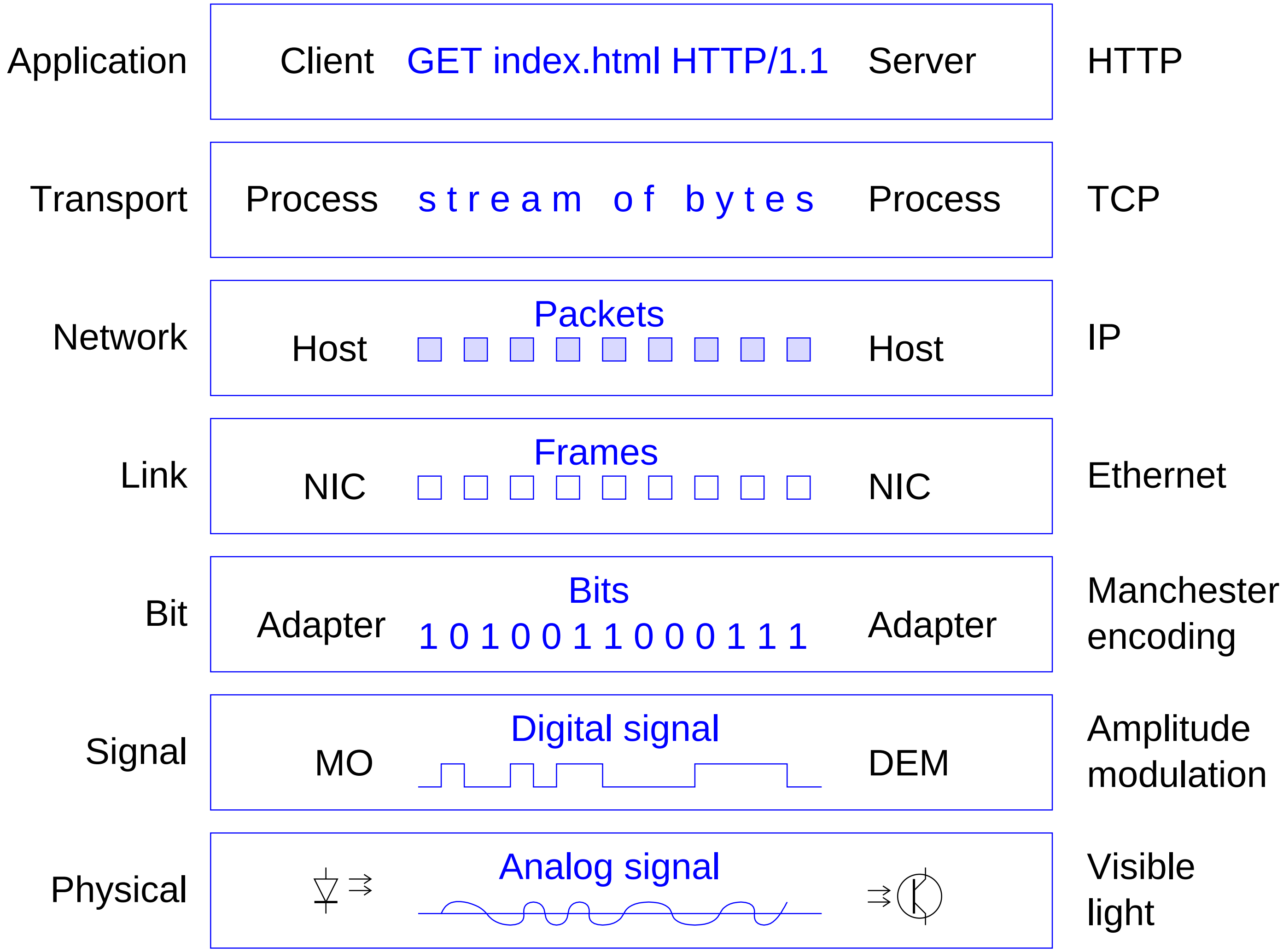
```
GET index.html HTTP/1.1
```

```
HTTP/1.1 200 OK
Server: Apache-2.0.44
Content-Length: 12481
Content-Type: text/html

<html>
 ...
</html>
```

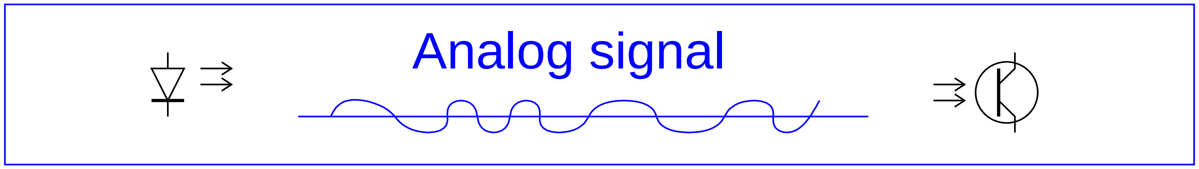
The Application Layer

Transport layer + application protocol = Application layer



So now you know...

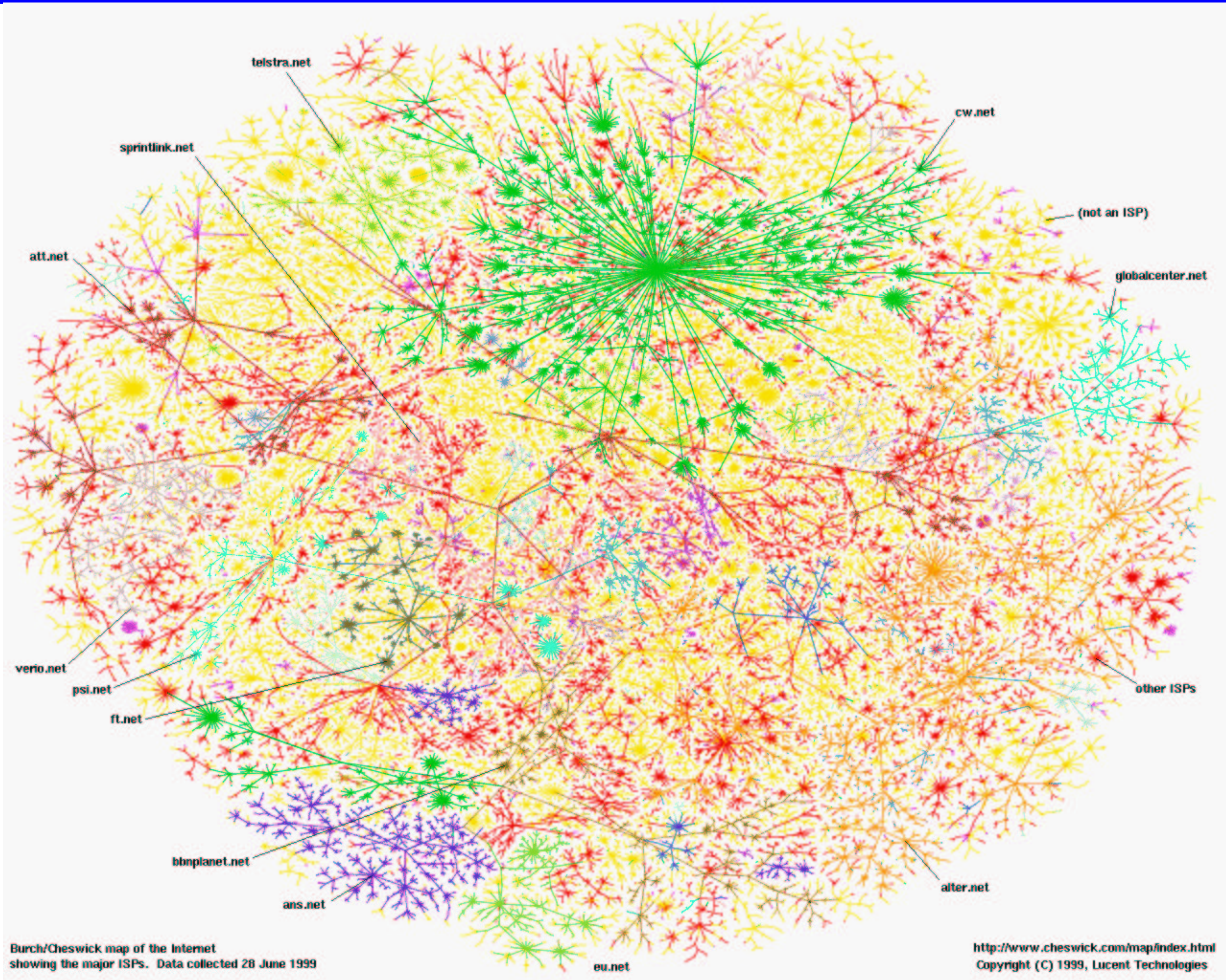
...given a medium that transmits an analog signal...



...how to build an Internet-compatible host.

But how do you build the Internet?

You don't. It grows...

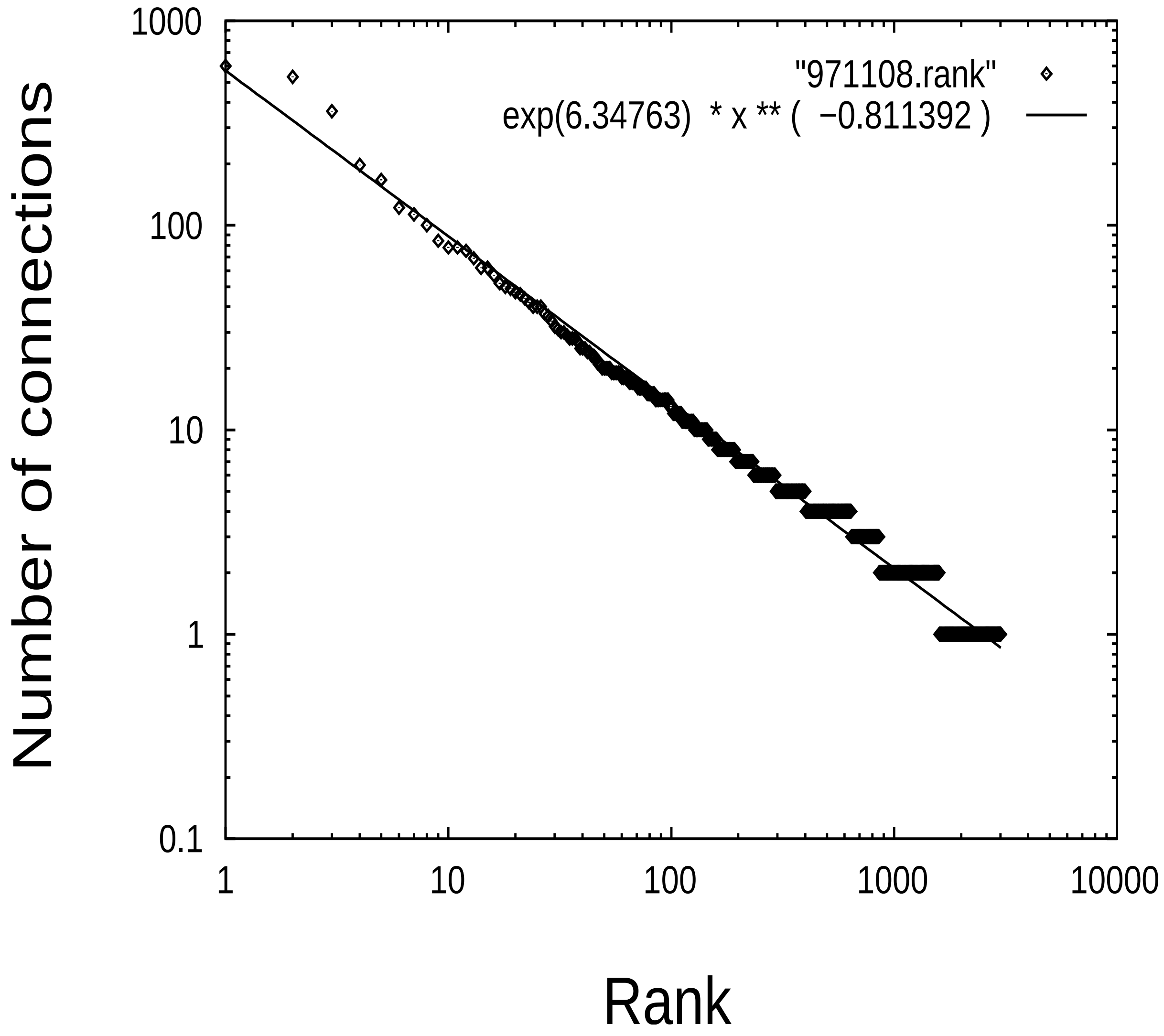


It grows, and grows...

Internet structure depends on:

- Technological details.
- Geographic details.
- Demographic details.
- Economic factors.
- Human factors.

It grows in mysterious ways



- Long-tailed distribution of connections (many small, some very large).
- Characteristic of rich-get-richer growth.

"On Power-Law Relationships of the Internet Topology" Faloutsos$\times 3$, 1999.

Zipf’s law

Zipf rank relationships in:

- Popularity of words (Zipf 1932).
- Sizes of cities (Zipf 1939).
- Individual income (Pareto 1906).
- Magnitude of earthquakes.
- Debt of bankrupt companies.

...and on and on.

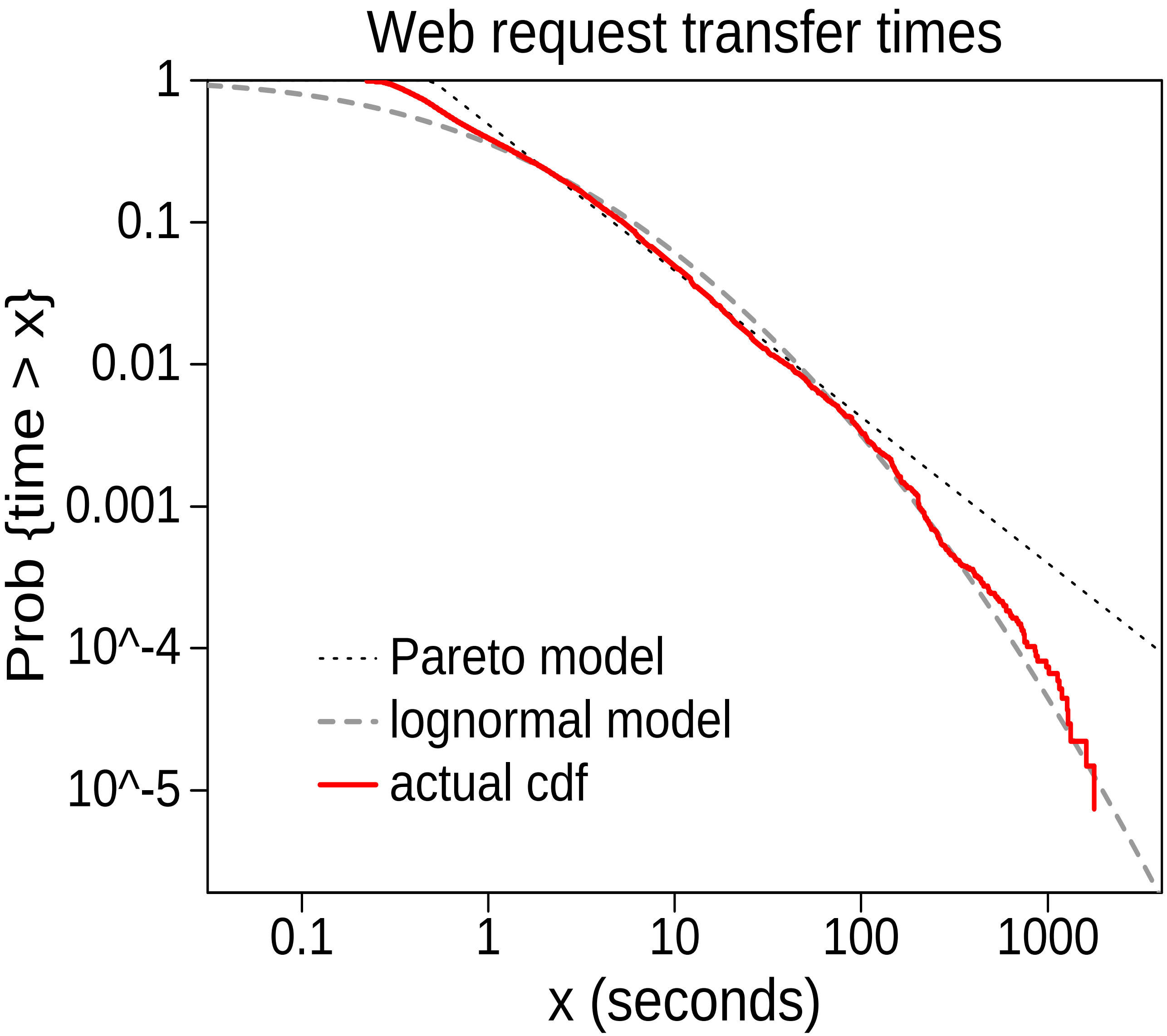
Zipf's law

And in the Internet...

- Degree of connectivity.
- Popularity of web pages.
- Size of web pages.
- Network transfer times.

...and on and on.

Long-tailed distributions



- Pareto distribution ⇒ straight line.
- Lognormal distribution ⇒ curved line.

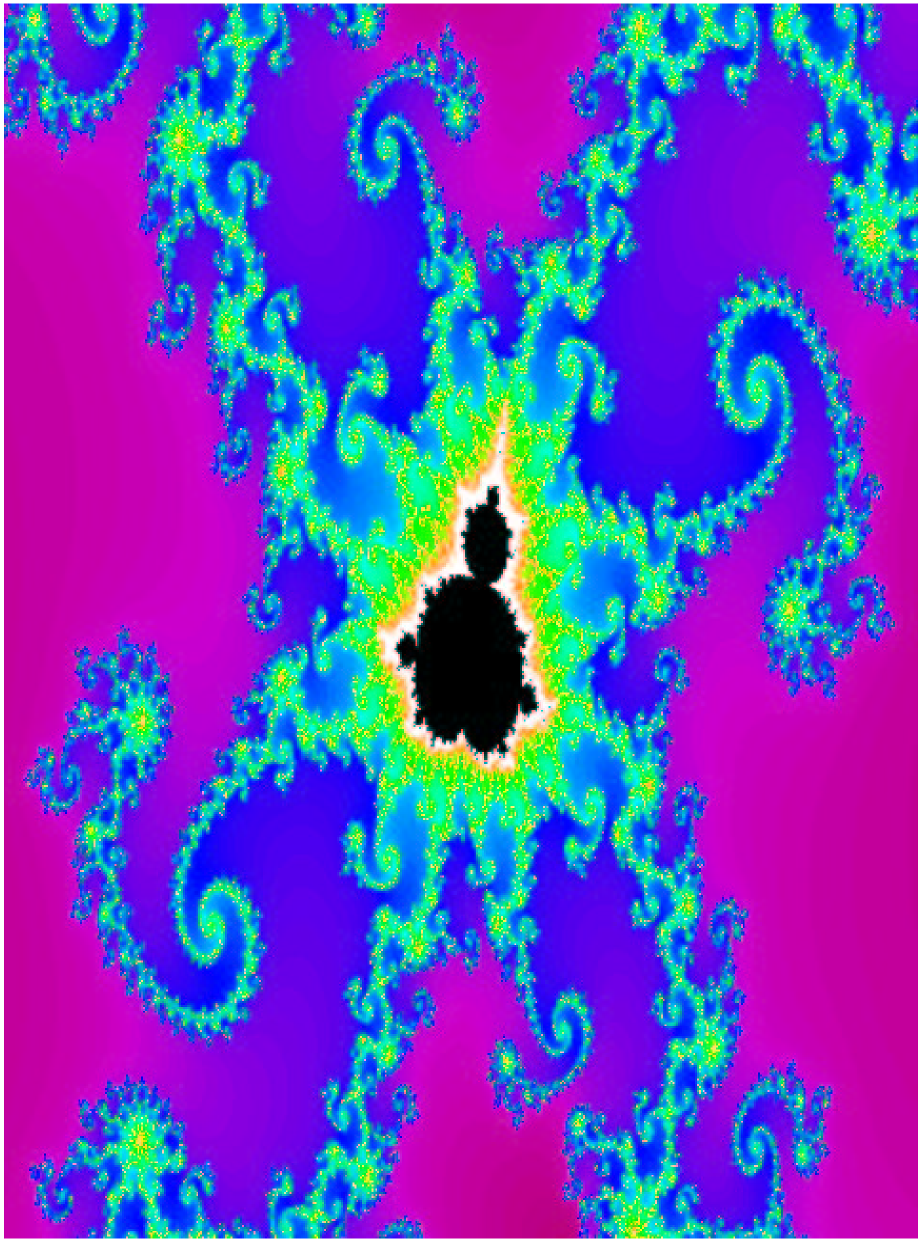
"Evidence for long-tailed distributions in the Internet", Downey, 2003.

So what?

Long-tailed distributions have several implications:

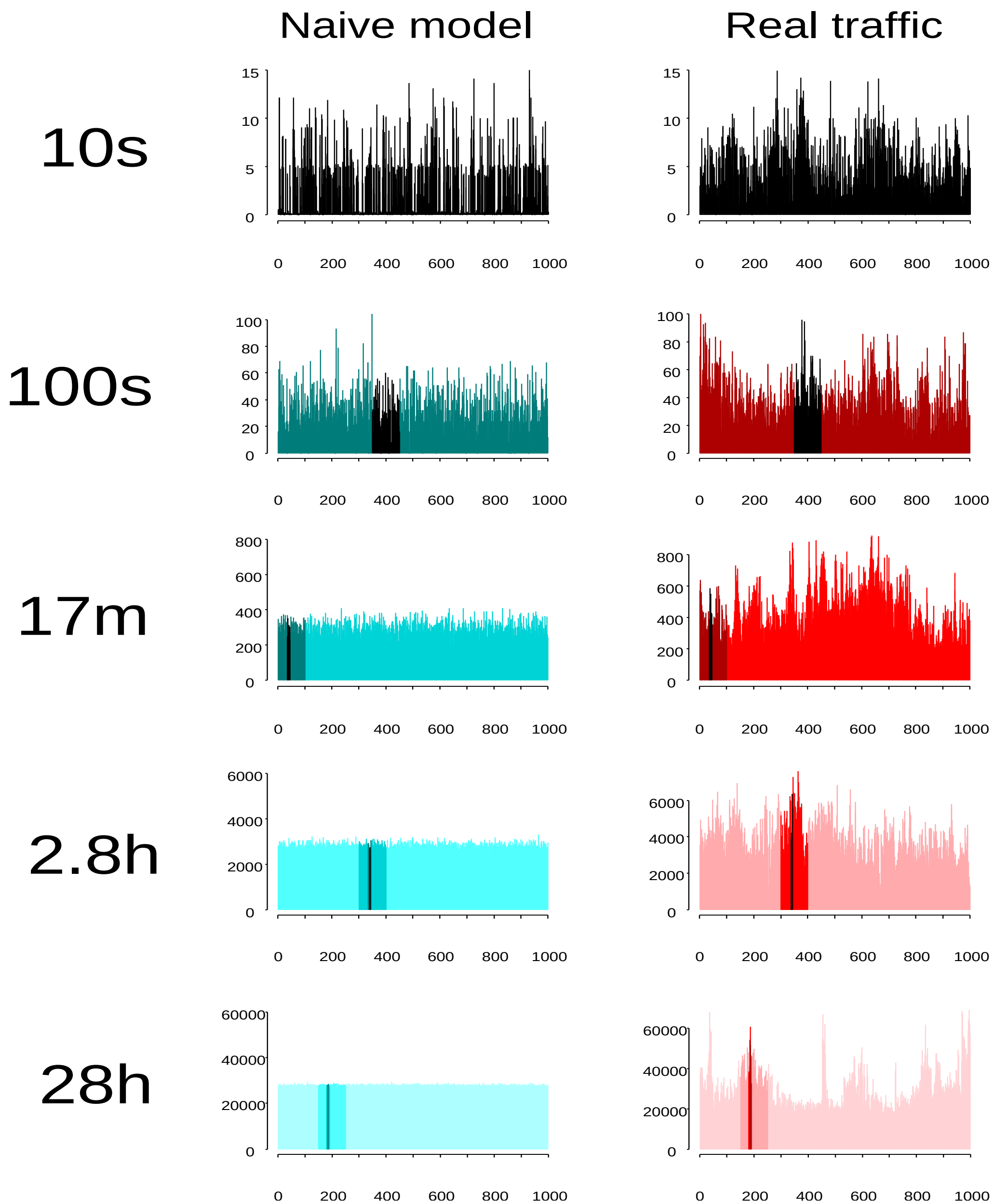
- Remaining lifetime increases with age.
- Network traffic is bursty.
- Network traffic is self-similar.

Self-similarity



- Real world: visually similar over a range of spatial scales.
- Fractals: geometrically similar over all spatial scales.
- Time-series: statistically similar over a range of time scales.

Self-similarity



- Normally, large scales are more predictable than small.
- Networks are bursty even at large time scales.

"On the Self-Similar Nature of Ethernet Traffic," Leland, Taqqu, Willinger and Wilson, 1993.

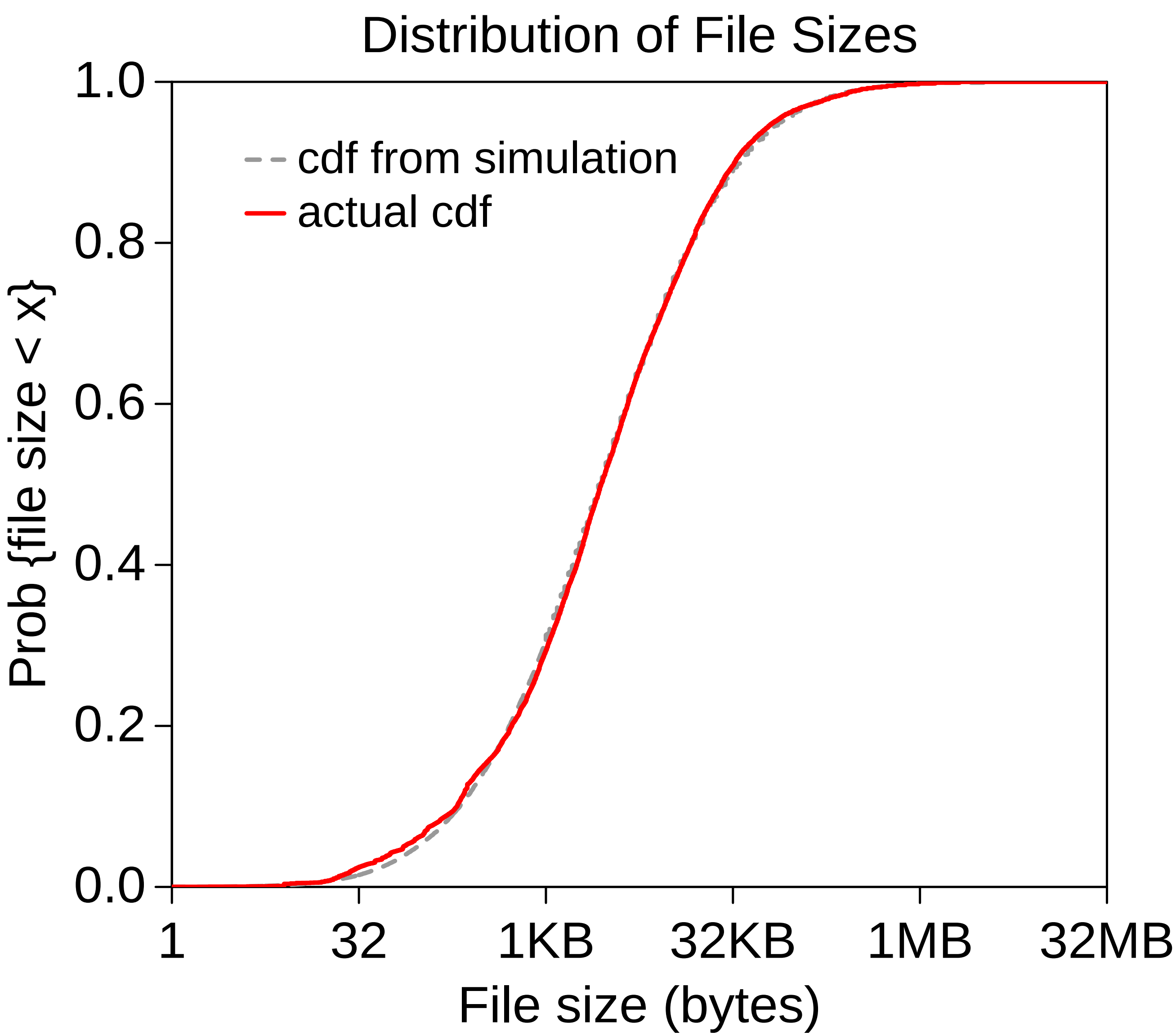
Self-similarity

Best current theory:

- Long-tailed transfer sizes cause self-similarity.

So what causes long-tailed transfer sizes?

File sizes



- File sizes are lognormal.
- Explained by simple model of user behavior.

"The structural cause of file size distributions," Downey, 2001.

User model

Most new files are a

1. copy,
2. modification, or
3. translation

of an existing file.

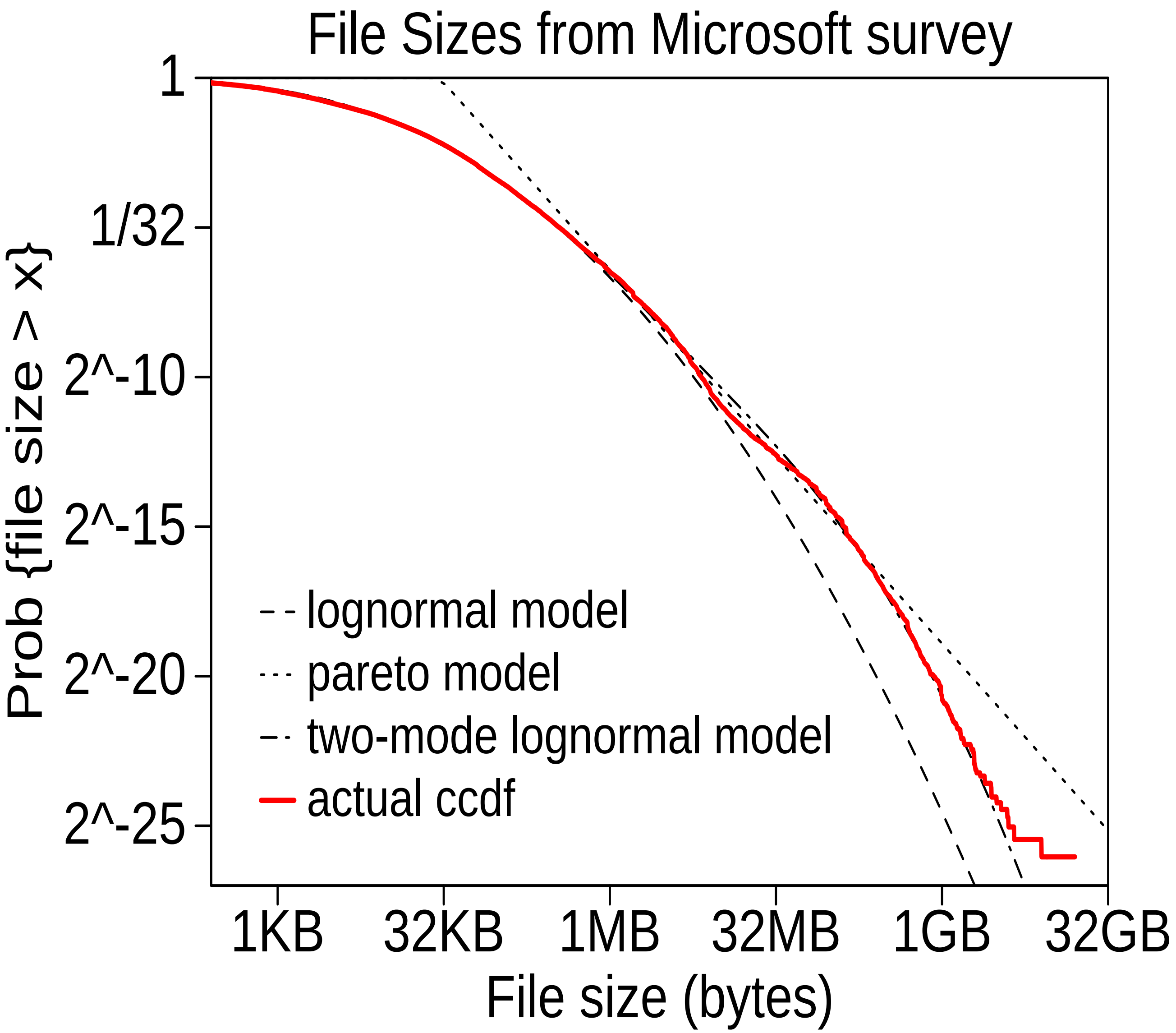
User model

Every day the user:

1. Chooses an existing file at random.
2. Chooses a small multiplier at random.
3. Creates a new file with size = old file size * multiplier.
4. Repeat.

Distribution of file sizes tends toward lognormal.

File sizes



- Mixture of lognormal distributions approaches a Pareto distribution.
- 40 million files from 10,000 machines.

Self-similarity in the Internet

As we understand it now...

- File systems evolve toward long-tailed size distributions.
- Long tailed-distributions induce self-similar traffic patterns.

Self-similarity in the Internet

These patterns are the result of interactions among:

- Technology.
- People.
- Information.

Summary

- Understand networking technology in terms of layers of abstraction.
- Each layer solves a set of problems, provides a set of capabilities.
- Understand the Internet in terms of underlying technology and emergent properties.

Hierarchical addresses

ICANN hands out blocks of addresses.

Class A	Class B	Class C
1...126	128...195	196...254
126 networks 16 million hosts per network	16,000 networks 65,000 hosts per network	2 million networks 254 hosts per network

A	18	181	0	131
	NETWORK	HOST		
B	128	197	10	145
	NETWORK		HOST	
C	216	34	94	186
	NETWORK			HOST

Hierarchical addresses

Addresses encode history.

- MIT has a Class A network (18.x.x.x)
- BU has a Class B network (128.197.x.x)
- Jeff Elkner has a Class C network (216.34.94.x)
- How did Olin get a Class A address??? (4.x.x.x)

Hierarchical addresses

The pros and cons.

- Inefficient use of addresses (IPv6).
- Common prefix $\Rightarrow$ colocation.