

# The Internet and Sockets

Computer Security

Tom Chothia

# This Time:

- How the Internet works.
  - Some History
  - TCP/IP
- Network address translation
- Uniform resource locators (URL)

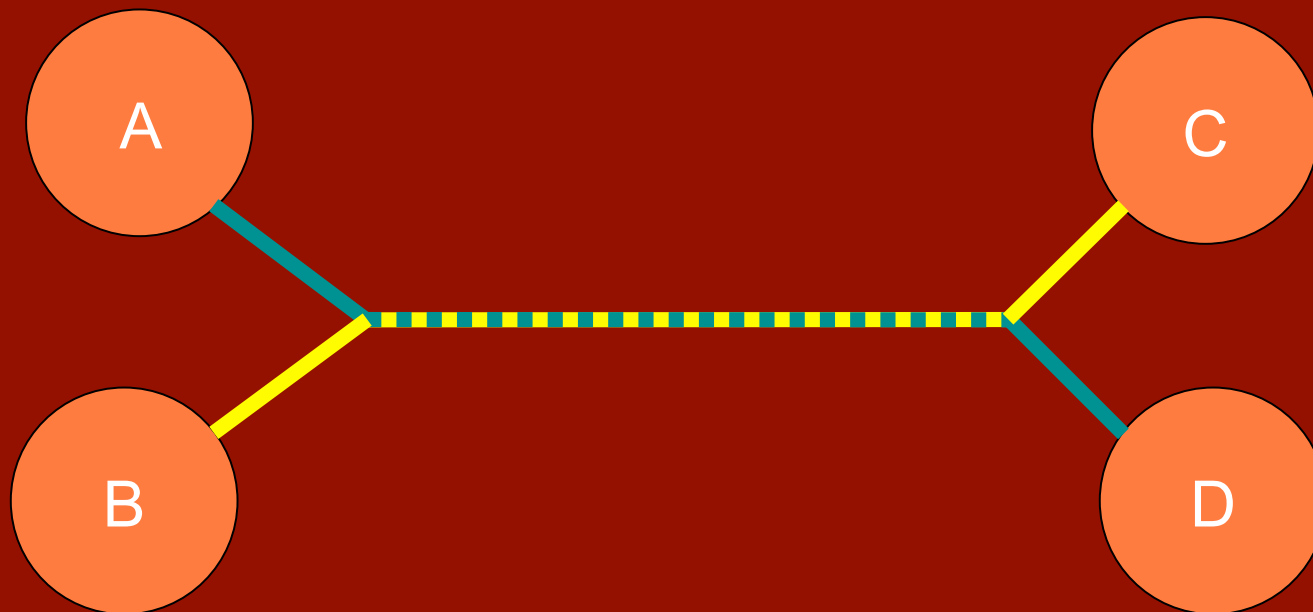
# Before the Internet

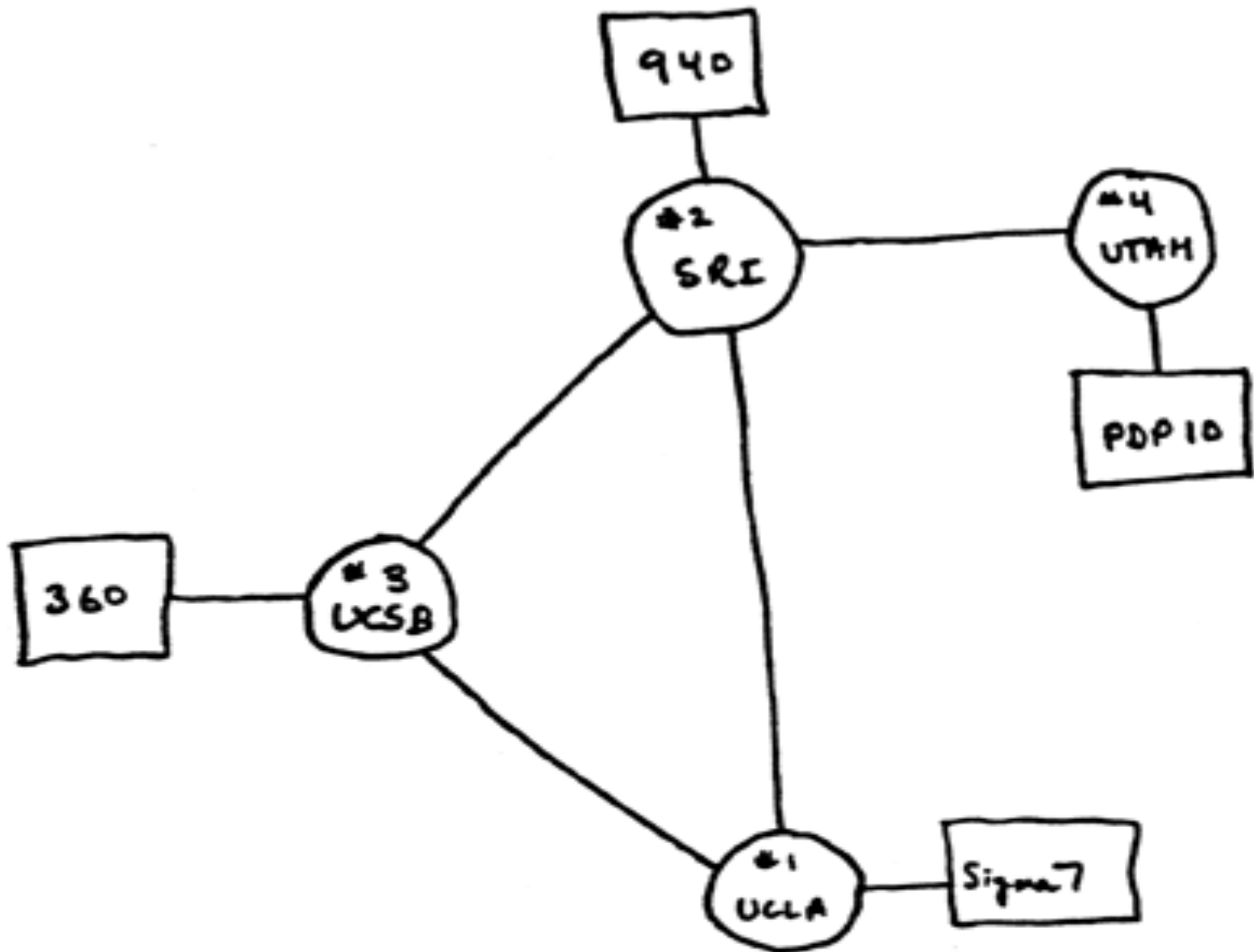
- Life was solitary, poor, nasty, brutish and short.
- Computer Networks:
  - local networks,
  - telephone line connections,
  - leased line.

# The Start 1969

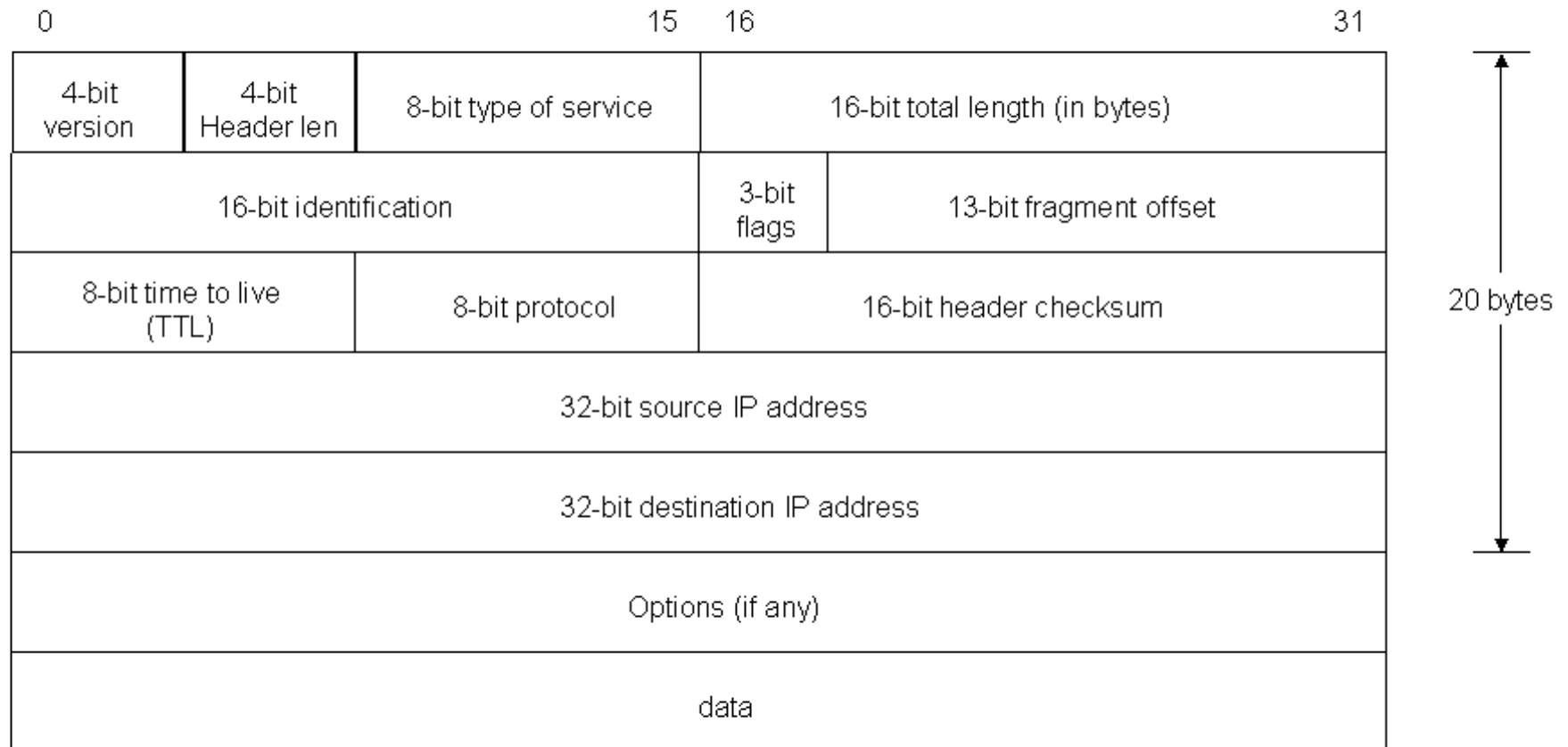
- The US Defense Advanced Research Projects Agency (then ARPA now DARPA) gives research grants to universities to buy computers.
- They decide to link their computers.
- But how?

# Packet Switching Lets all Users Share the Line



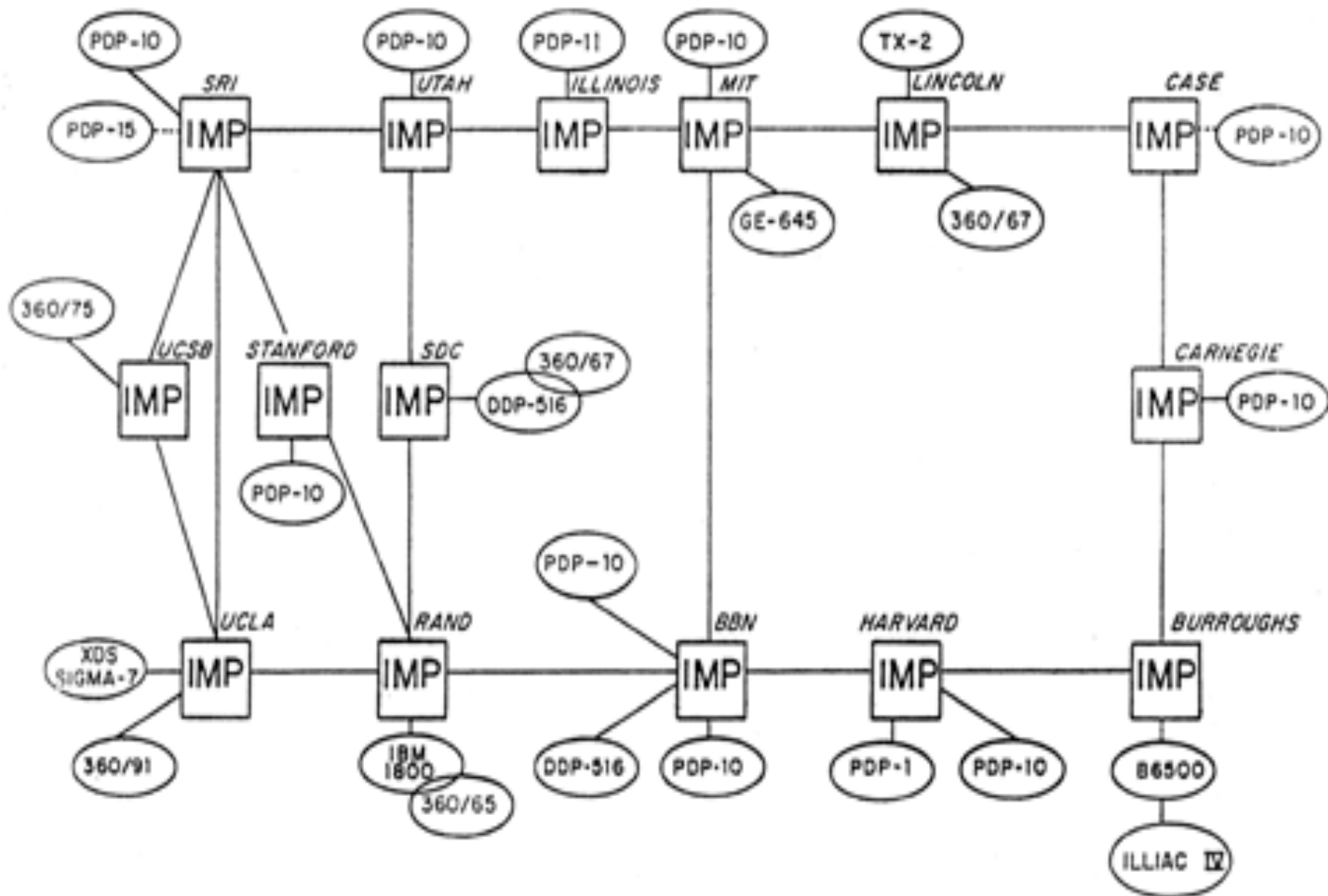


# IP PACKET HEADER



# IP Addresses

- Every computer has an IP address.  
e.g. 147.188.193.15
- Every router on the Internet keeps a list of which connection it should use for which addresses.

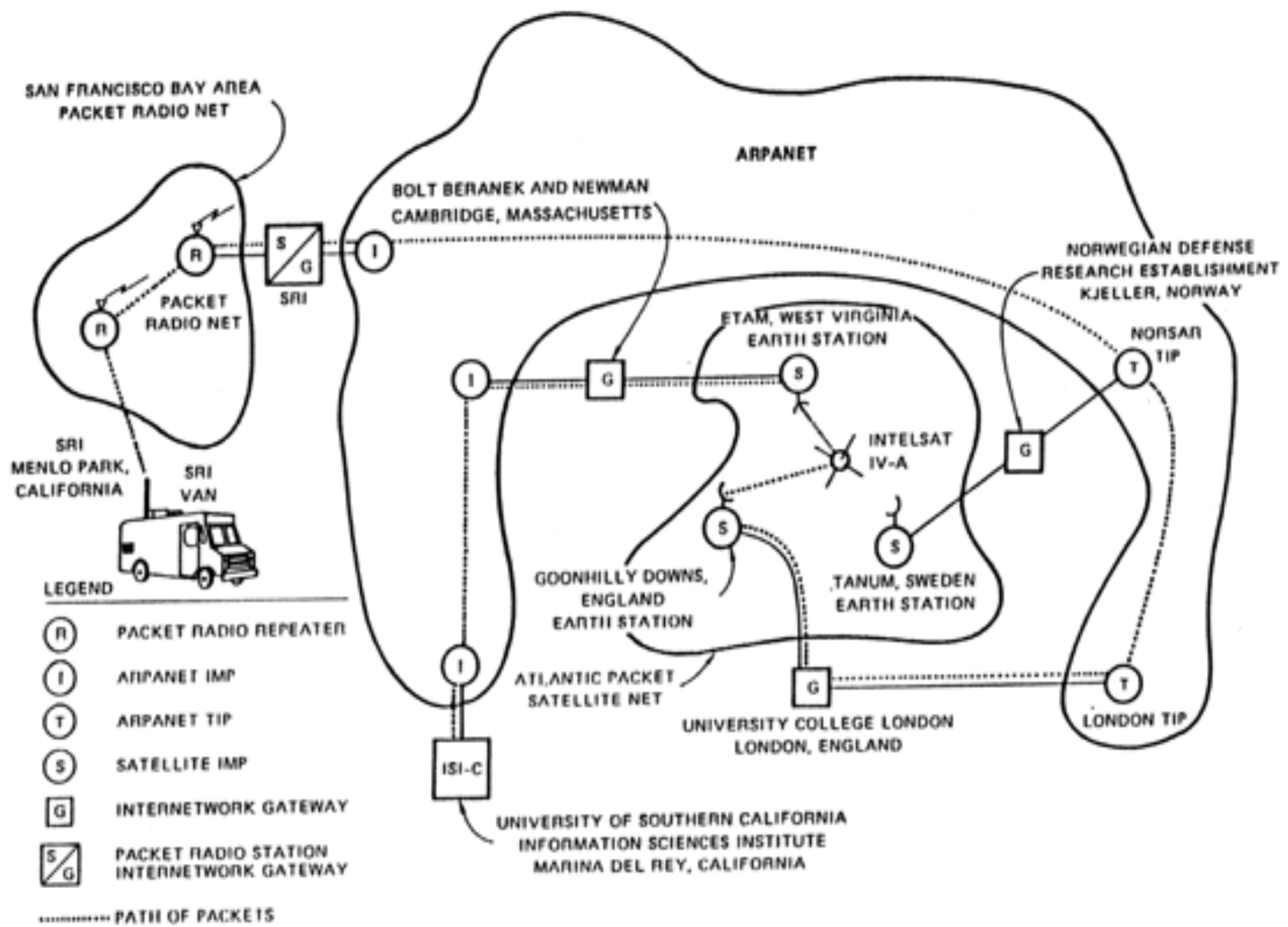


ARPA NET, APRIL 1971



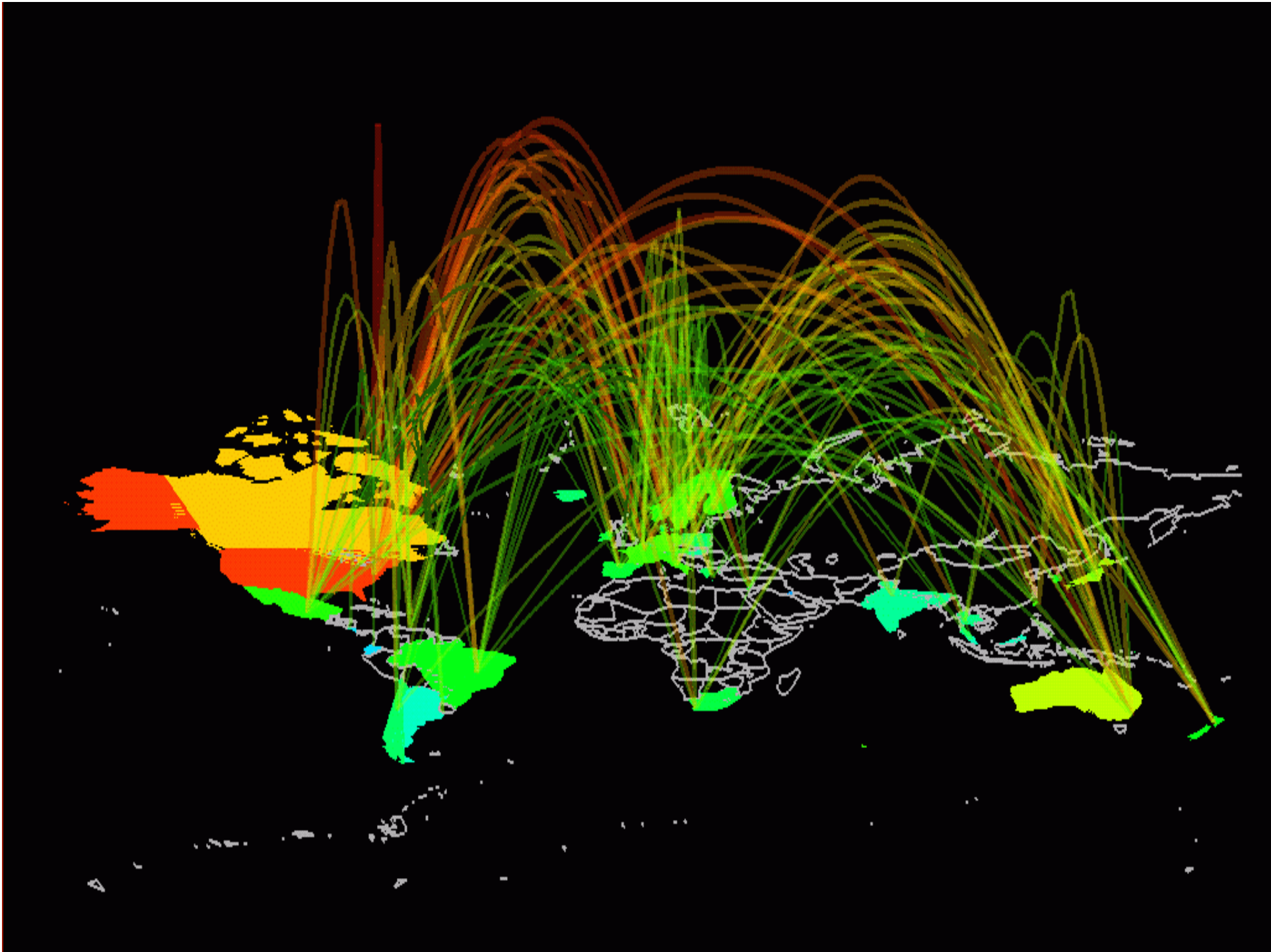
# Transmission Control Protocol

- 1974: daily traffic more than 3 million packets a day. Many are getting lost.
- TCP is a protocol than runs on top on IP, if an IP packet gets lost. It requests that it is resent.
- TCP/IP becomes allows Inter network connections. So in 1977:



# The ARPAnet

- 1969 the birth of the ARPAnet.
- 1971 E-mail @
- 1977 the ARPAnet becomes a Inter-network.
- 1978 first Spam message send.
  - Admin assures users this will never happen again
- 1983 TCP/IP becomes standard.

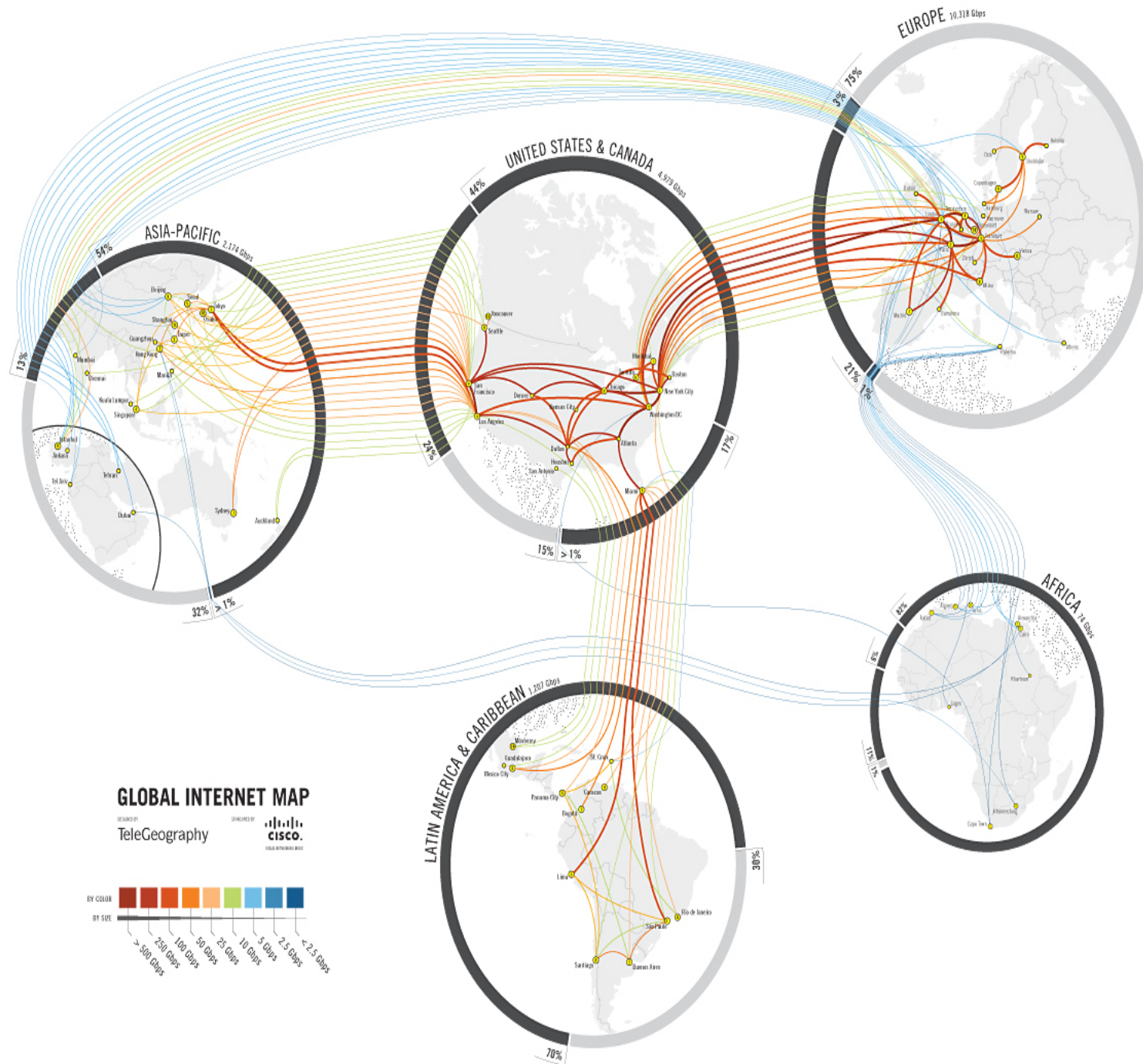


# Domain Name Servers (DNS)

- Remembering IP address is too hard.
- So people associate names with addresses.  
e.g. news.bbc.com → 212.58.226.141
- A hierarchy of servers list handle requests
- The route for most of Europe is RIPE based in Amsterdam.

# The Internet

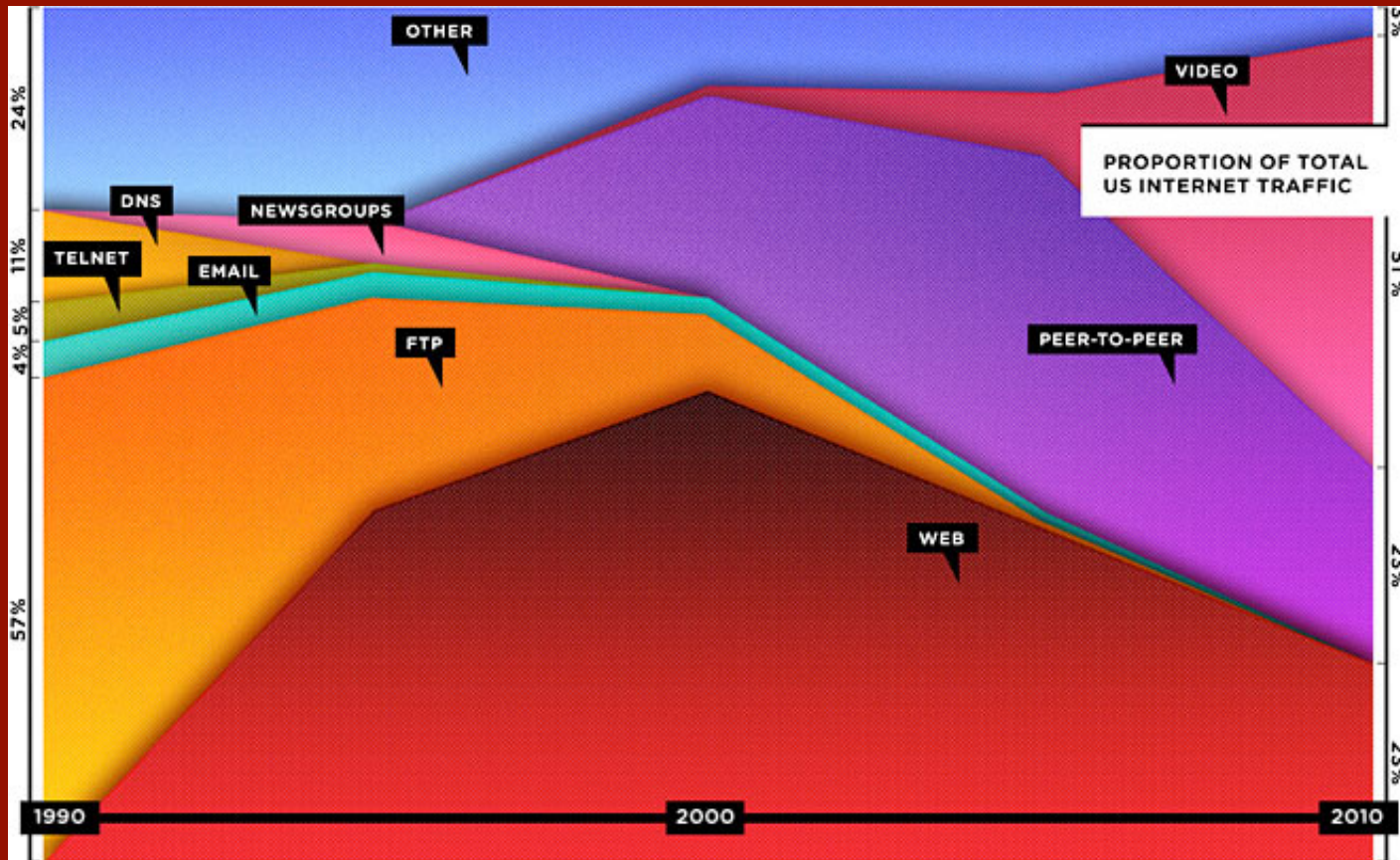
- 1988 The Morris Worm, CERT formed.
- 1989 The Web: HTTP and HTML.
- 1992 The first ISPs and *Mosaic* the first web browser.
- 1994 Yahoo



# Today

- All computers on the Internet have an IP address.
- Key Internet Protocols: TCP/IP
- DNS to turn names into numbers.

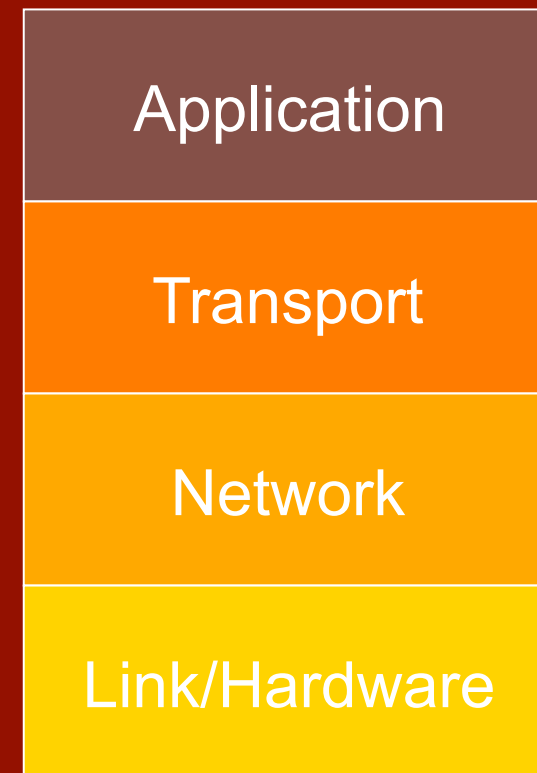
# Internet Traffic



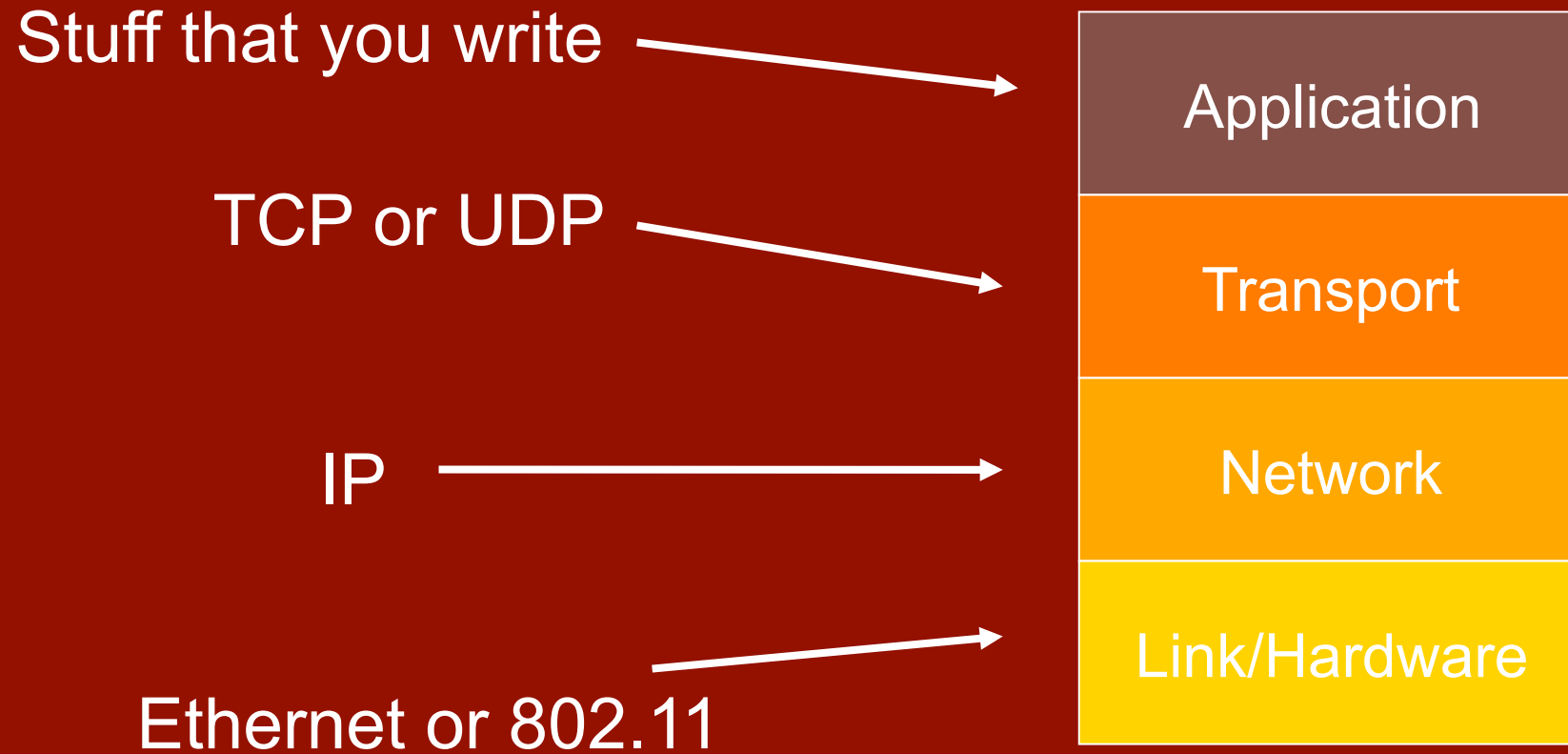
# The Internet Protocol Stack

Internet communication uses a stack of protocols.

Each protocol uses the protocol below it to send data.



# The Stack, Most of the Time:



# Sockets

- A computer is identified by an IP address.
- To allow multiple connections each computer has a port number.
- A simple connection between two computers is called a socket.
- Sockets are uniquely identified by (IP sender, port sender, IP receiver, port receiver).

# Port numbers

IANA registers port numbers. E.g. WWW runs on port 80, ssh on port 22, dns on 53, (free wifi).

0-1023: “well known ports” assigned by IANA  
(best not to use these)

1023-49151 “registered ports” IANA allows  
anyone to register and use these.

49151-65535: cannot be registered.

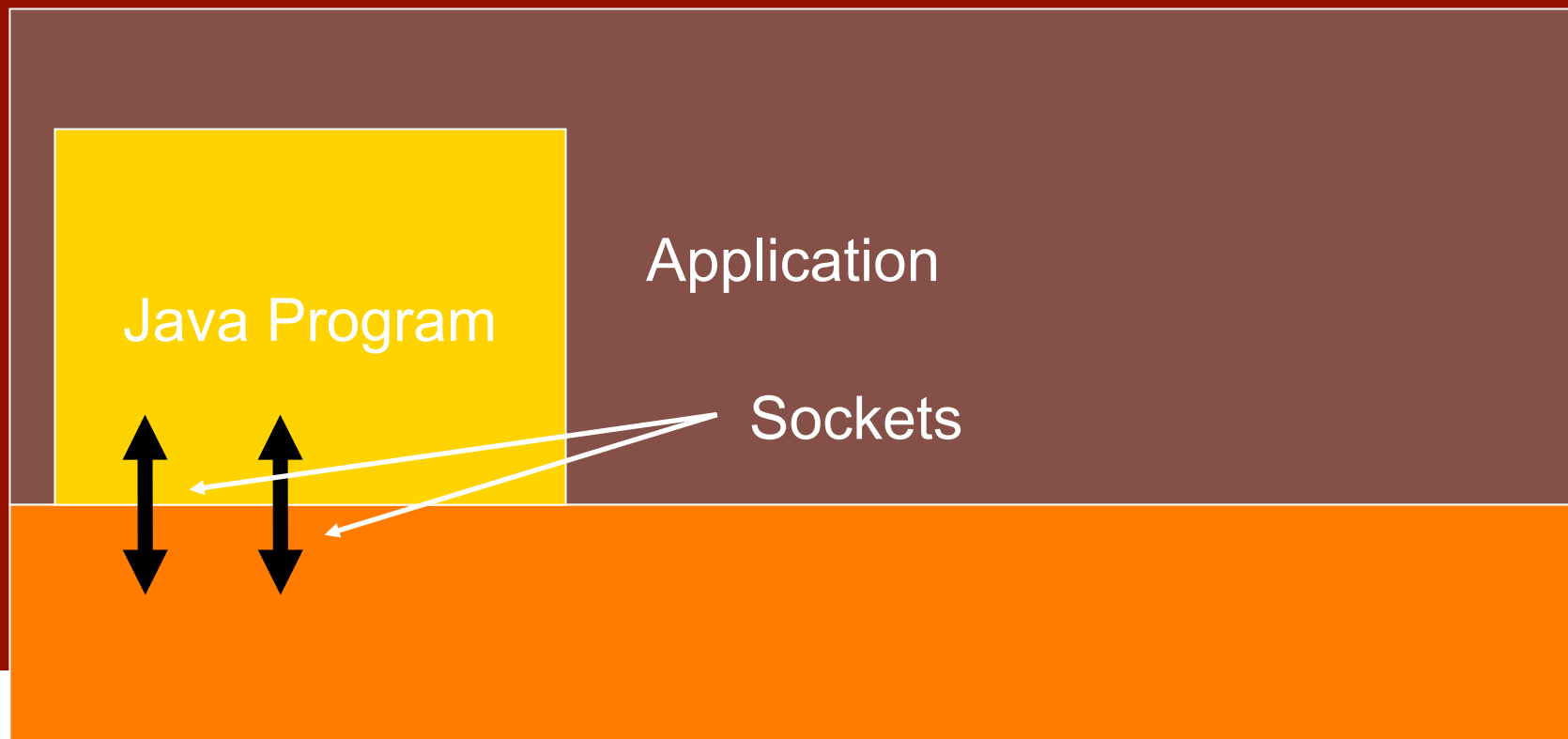
# Nmap: <http://nmap.org/>

A network scanning tool: tells you which ports are open.

```
> nmap 127.0.0.01
```

Many other functions (see `man nmap`)

# Our View of the Stack in Java



# Sockets in Java

- Use `java.io.Socket` to make a socket that connects to a remote computer.
- Use `java.io.ServerSocket` to make a socket that listens for an incoming connection.
- Sockets just give bytes, so then need a wrapper.

# WireShark [www.wireshark.org](http://www.wireshark.org)

A network protocol analyzer.

It records all Internet traffic, so it can then be views and analysed.

Excellent for debugging, protocols and finding out how Internet applications work.

# Network Address Translation

How many IP address are there?

0.0.0.0 to 255.255.255.255

Therefore

$256 * 256 * 256 * 256 = 4\ 294\ 967\ 296$  address

Not enough for 1 per person (I have 4)

# Network Address Translation

Answer IP6: Internet Protocol version 6

- $2^{128}$  addresses! Enough for everyone
- Not backwards compatible ☹️

Another answer: Network Address Translation (NAT), Share an address.

# NAT

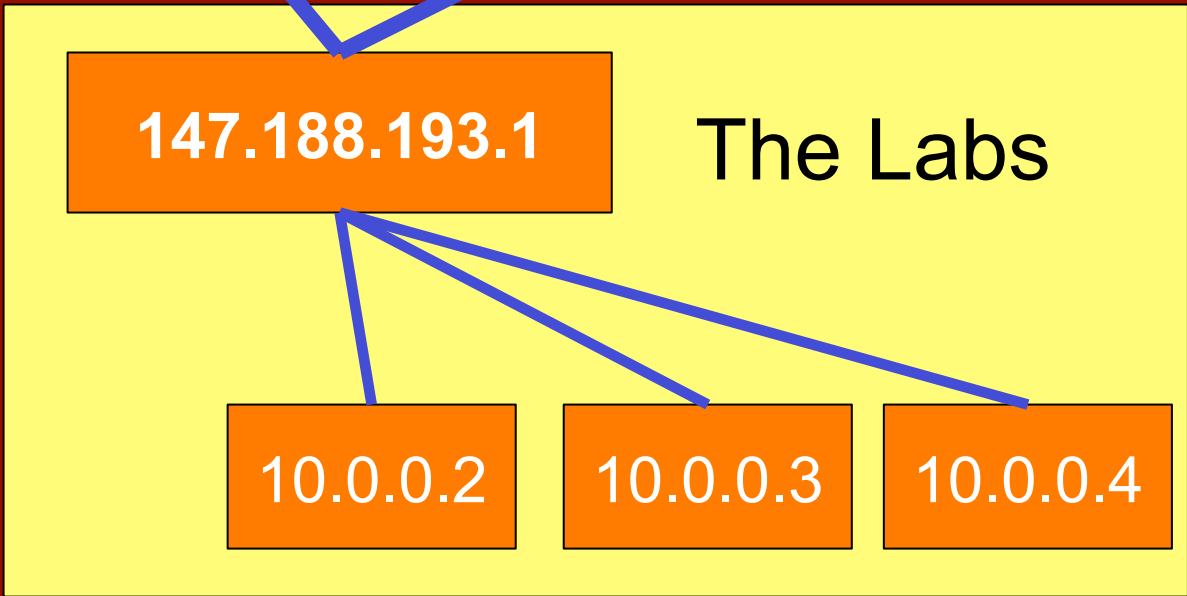
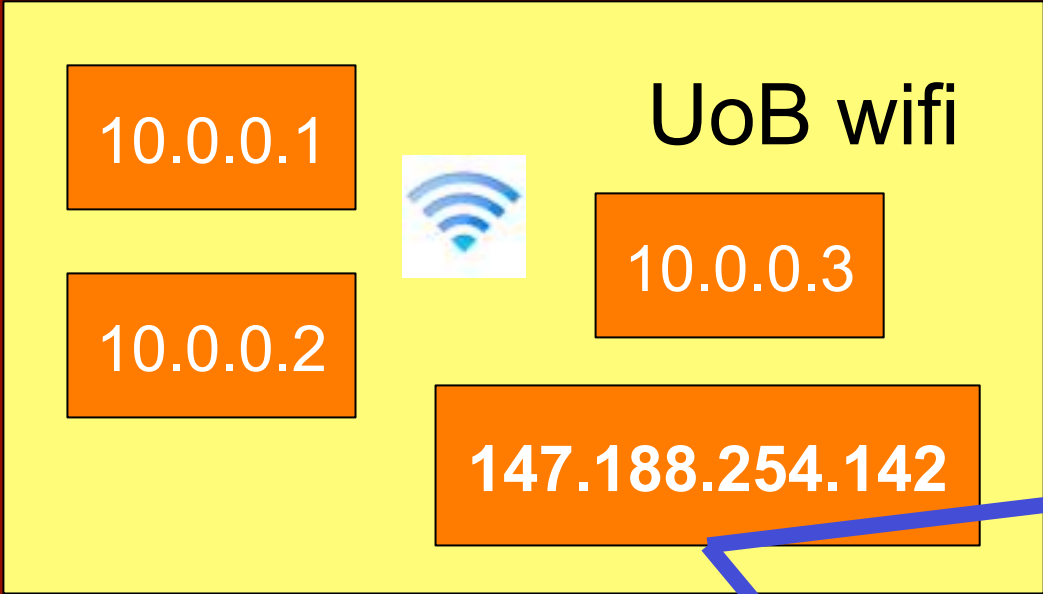
Address in the ranges:

192.168.\*.\* , 172.16.\*.\* 172.31.\*.\* , 10.\*.\*.\*

are local only.

Local machines communicate with the outside Internet use a NAT router that shares it's IP address.

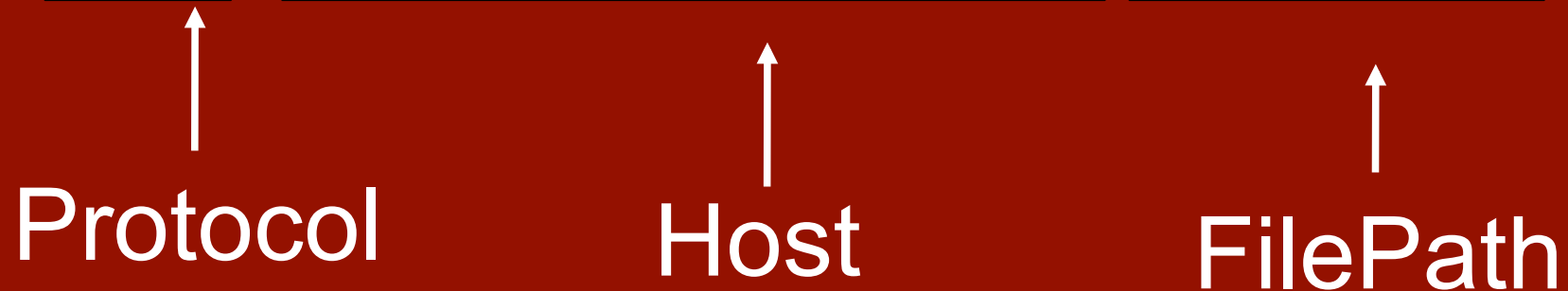
Ports used to tells the machines apart.



# Uniform Resource Locators

- Many Internet systems use URLs, e.g.

`http://www.cs.bham.ac.uk/index.html`



This URL refers to the file “index.html” on the computer “www.cs.bham.ac.uk” which should be accessed using http.

# URLs aren't just for HTTP

- <https://www.gmail.com/index.html>
- <ftp://ftp.funet.fi/pub/standards/RFC/rfc959.txt>
- <rmi://68.45.12.7/process>
- <svn+ssh://svn.cwi.nl/projects>

# More complex URLs

<http://www.cs.bham.ac.uk:3080/index.html>

↑  
Port number

<ftp://tpc@my.ftp.bham.ac.uk/myfiles.txt>

↑  
Username

Anchor

[http://en.wikipedia.org/wiki/Uniform\\_Resource\\_Locator#Syntax](http://en.wikipedia.org/wiki/Uniform_Resource_Locator#Syntax)

# URL syntax

- The most general URL has the syntax:
- `protocol://  
username:password@domain:port  
/filepathname?query_string#anchor`
- All parts are options!

# Conclusion

- The Internet uses a packet based routing system IP.
- Information sent in clear text and can be read and altered by any machine on its path.
- Network address translation
- URLs