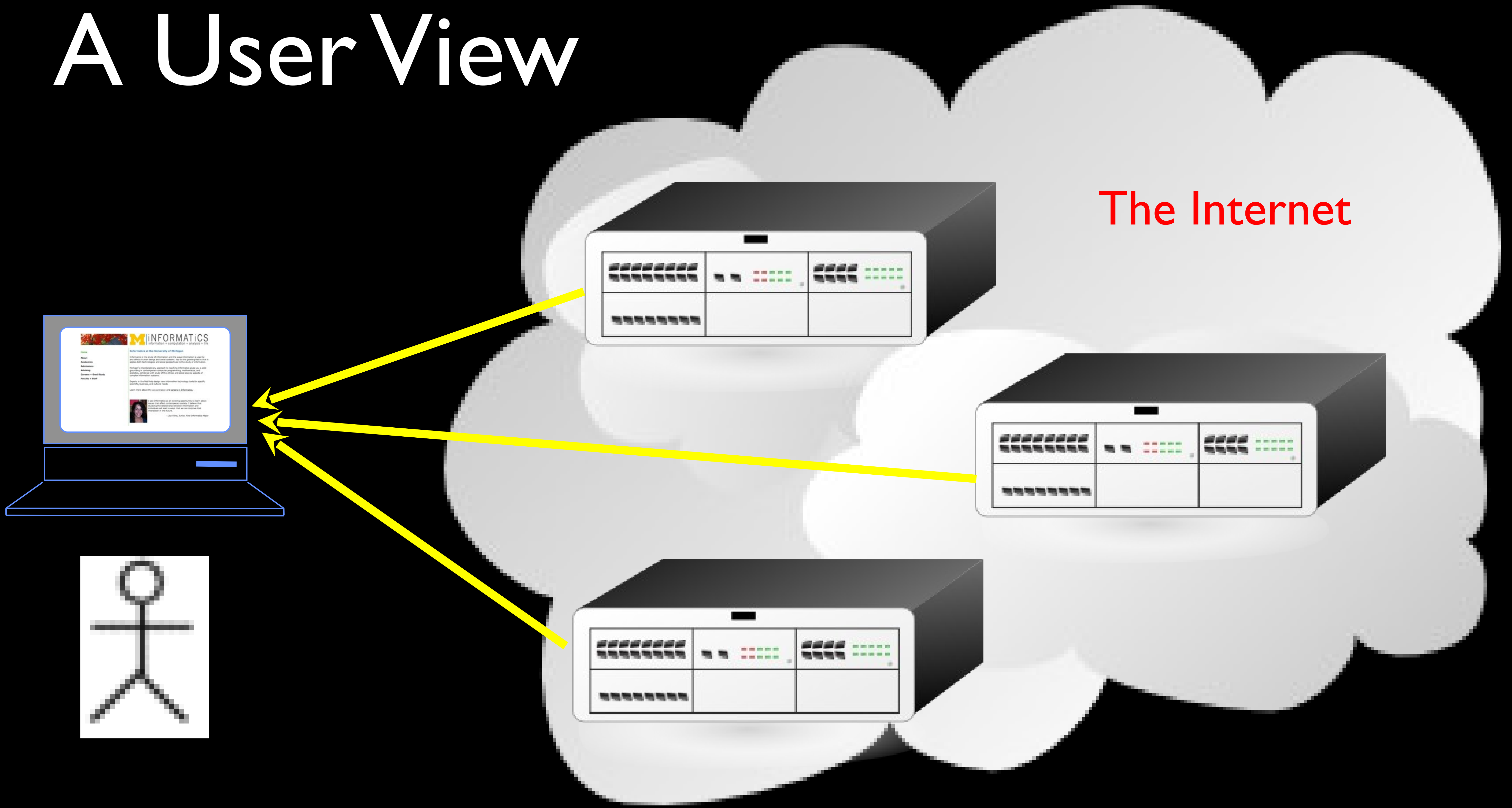
Internet Technologies

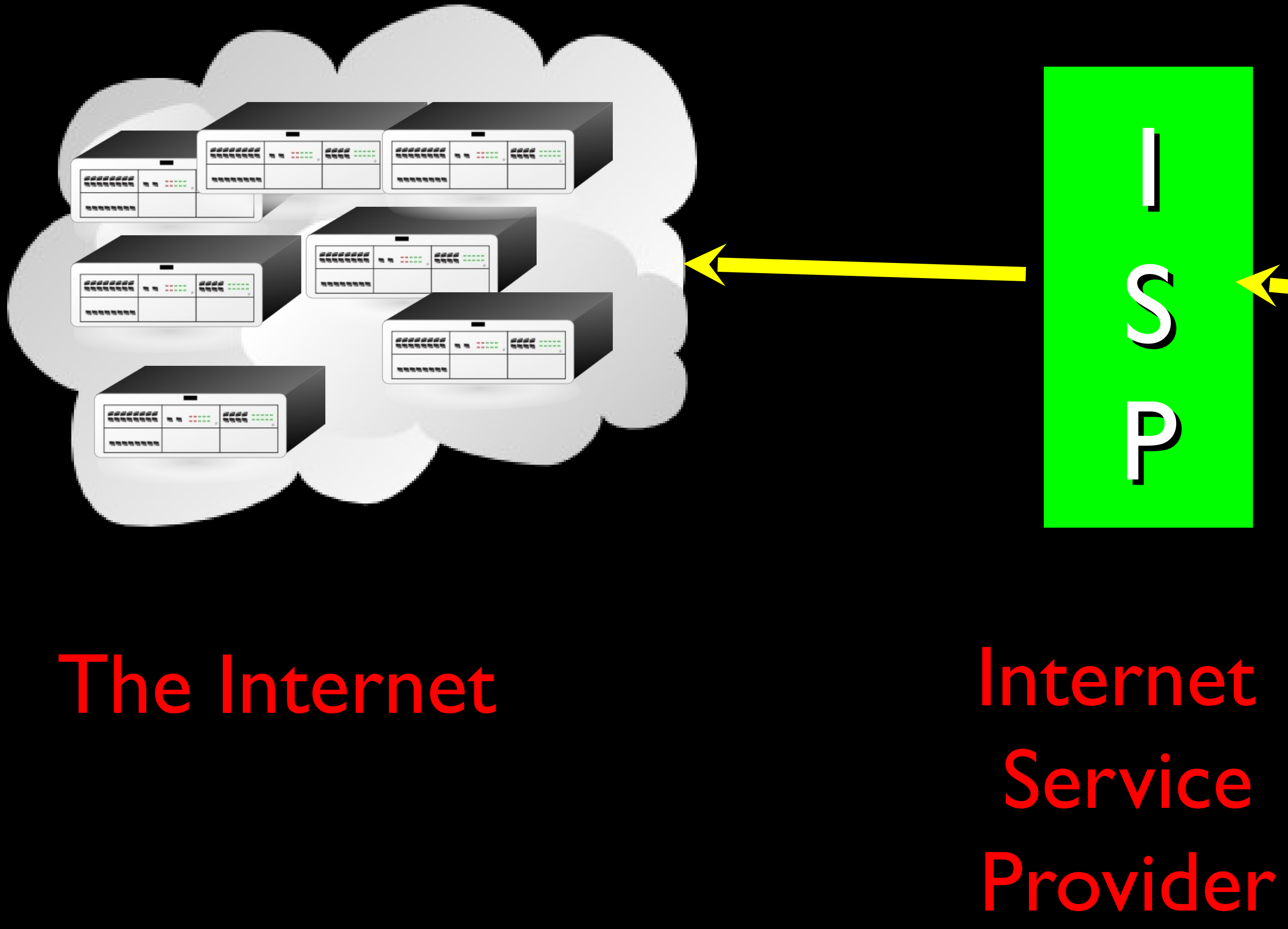
SI502 - Charles Severance

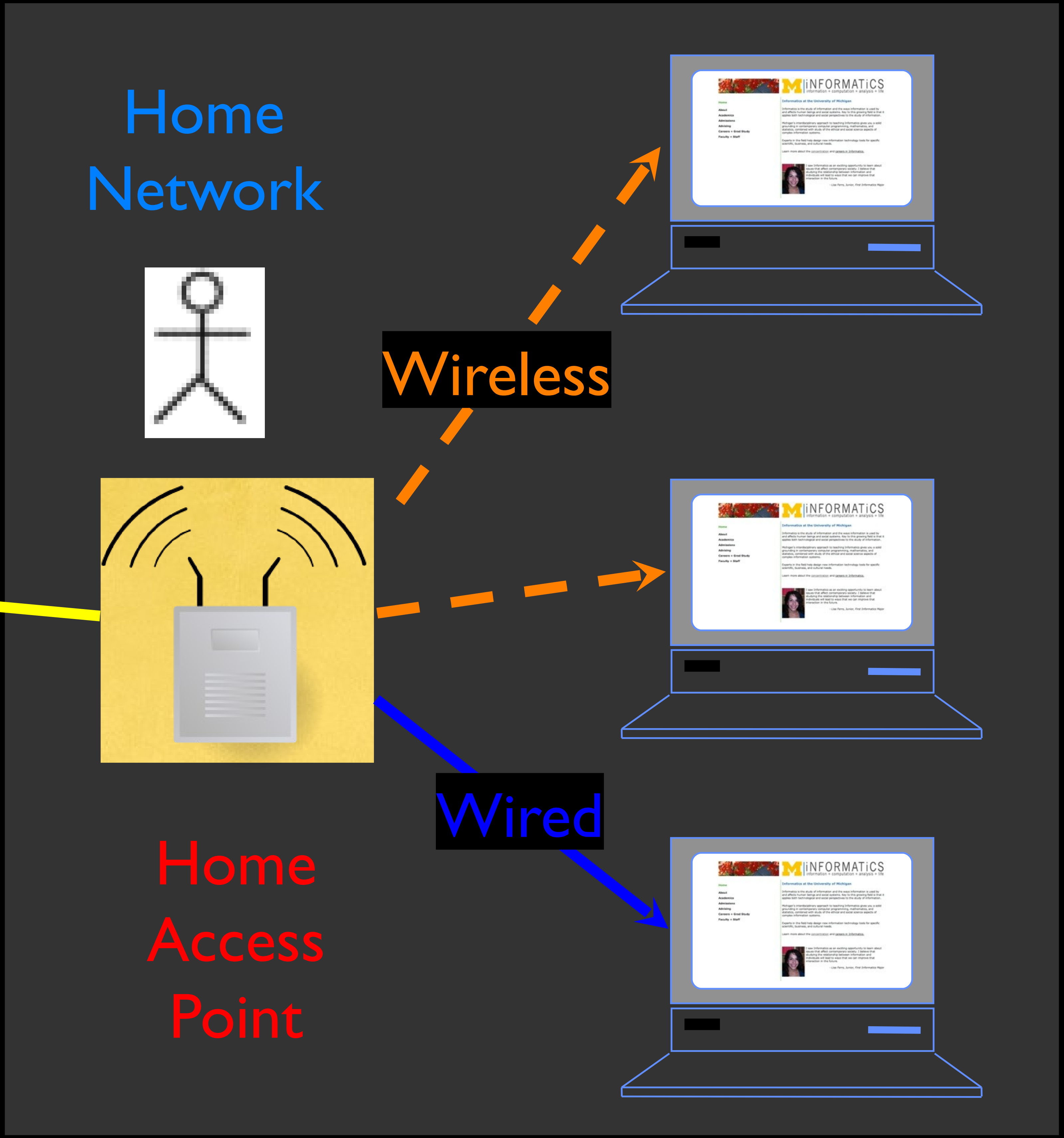
A User View



Clipart: http://www.clker.com/search/networksym/1

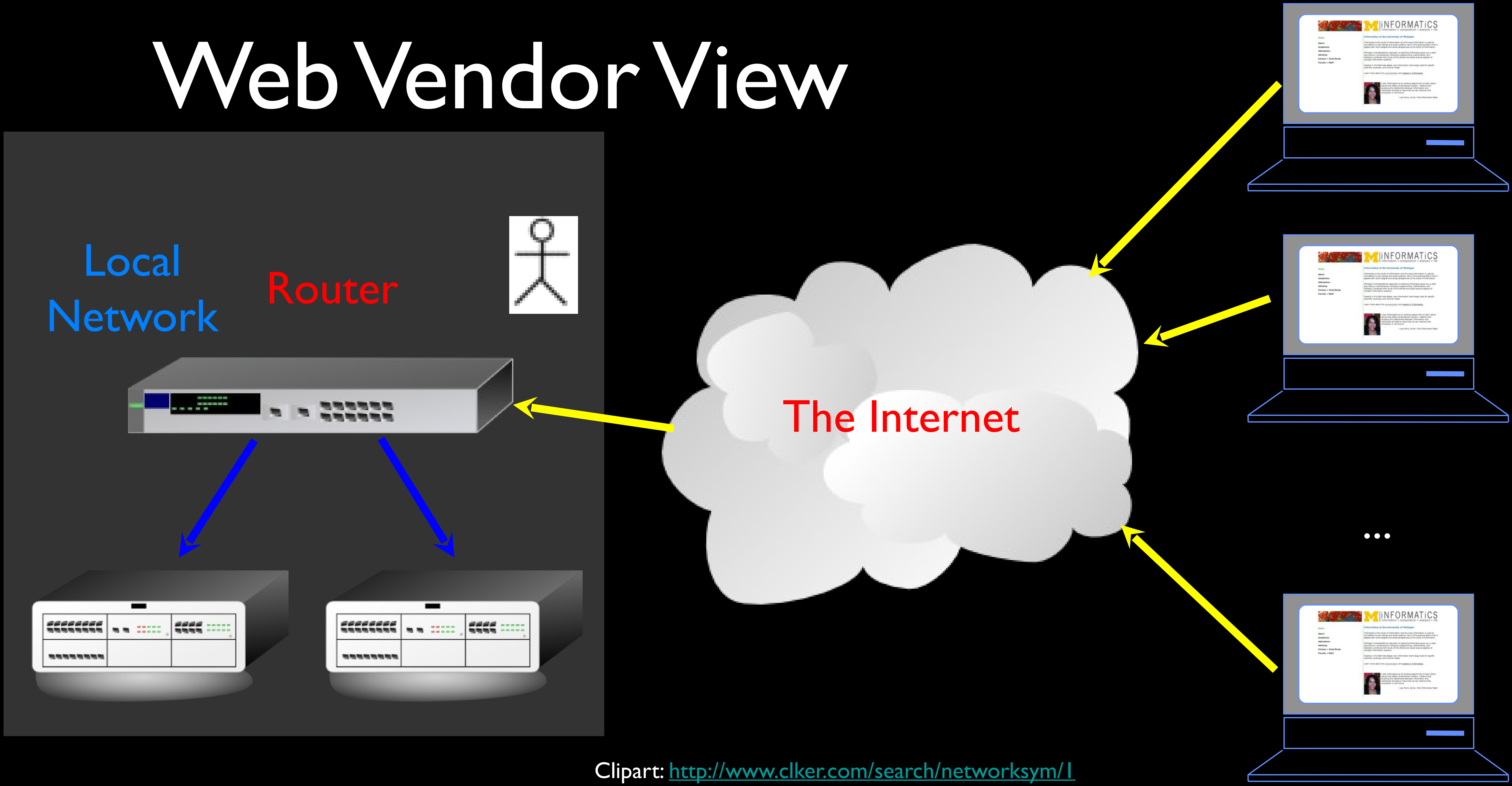
Home Tech Support View





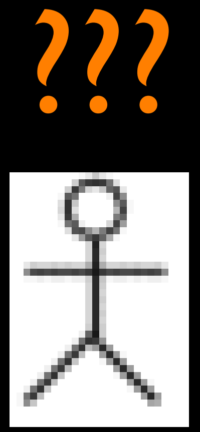
Clipart: http://www.clker.com/search/networksym/1

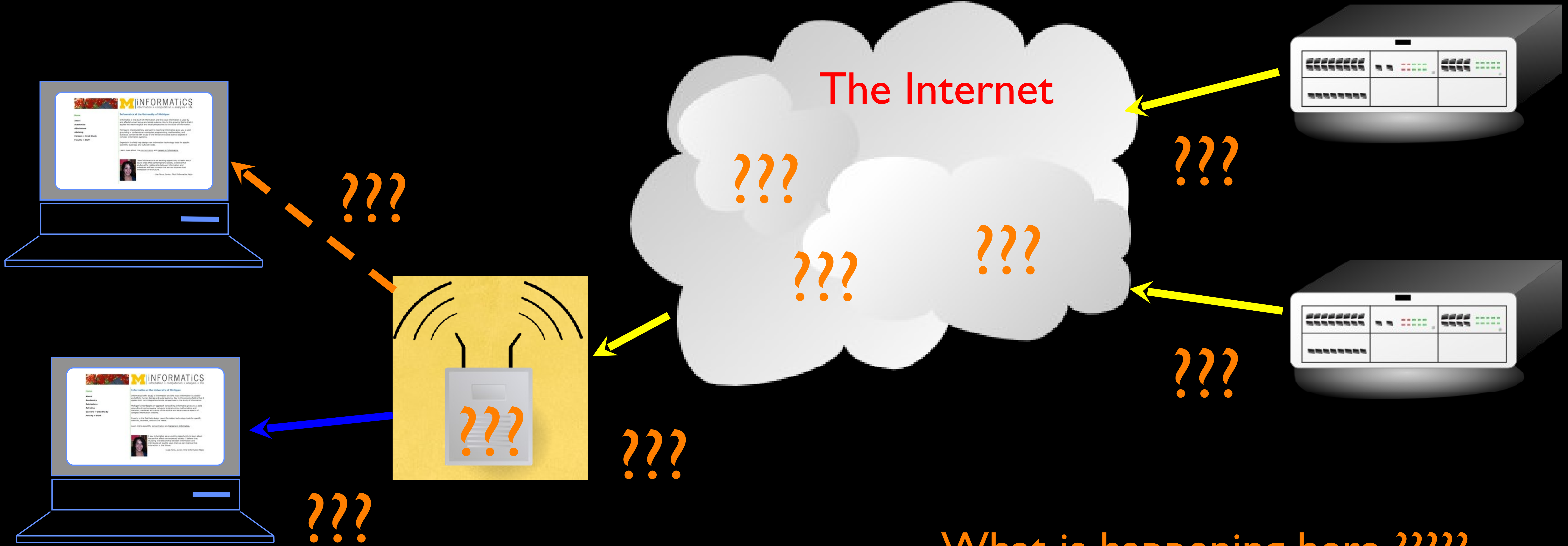
Web Vendor View



Clipart: http://www.clker.com/search/networksym/1

We are curious about many things....





What is happening here ?????

Clipart: http://www.clker.com/search/networksym/1

Outline (Part 1)

- Internet History
- Internet Overview
  - Internet architecture
  - Layered Architecture
- Link Layer
  - Wired
  - Wireless
- The Internet Protocol (IP)

Terminology

- Protocol - A set of rules that we agree to so computers can talk to each other - a language if you will
- LAN - Local area network
- WAN - Wide area network
- Router - Something that connects one network to another network
- Access Point - A home router - often with wired and wireless support
- Packet - Chunk of data - large amounts of data are broken up into packets

Internet History

http://en.wikipedia.org/wiki/Internet_history
http://en.wikipedia.org/wiki/NSFNet

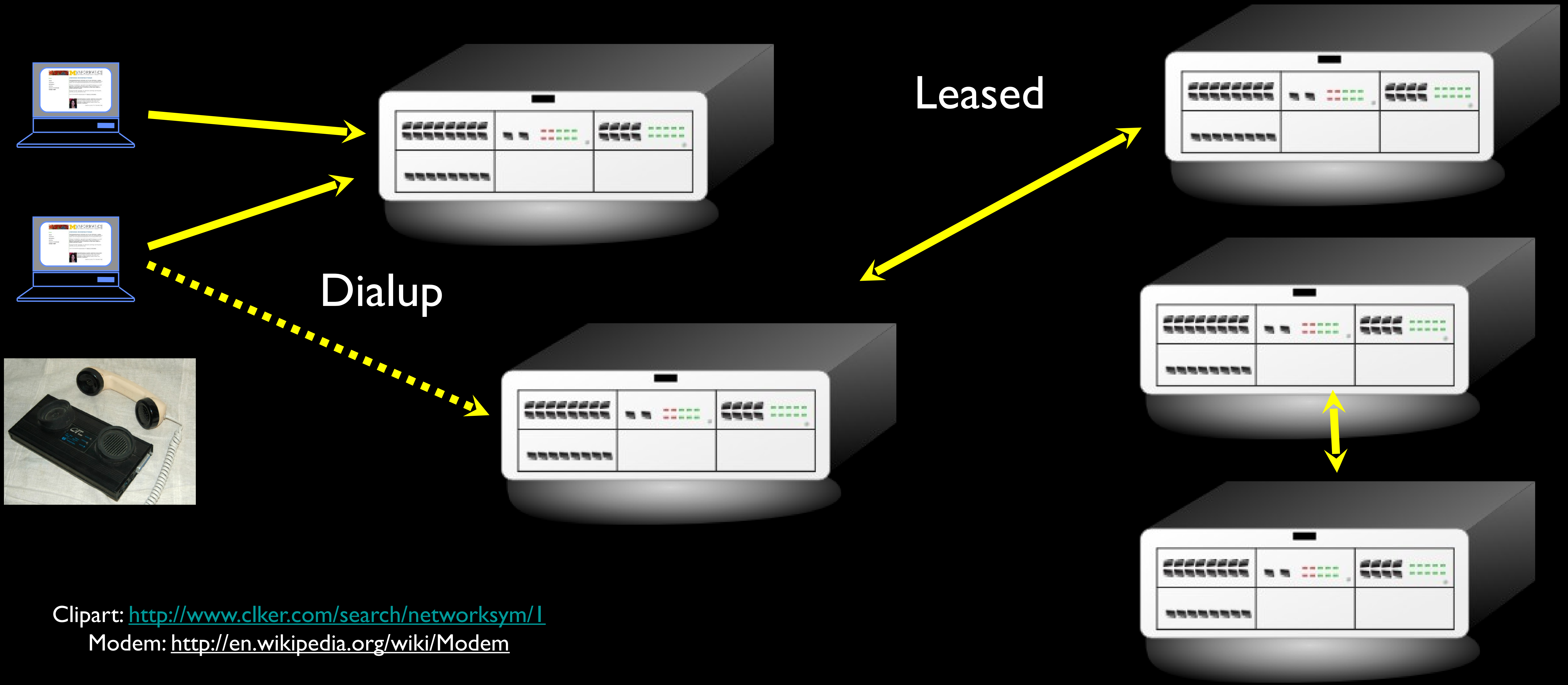
High Level Phases

- Research Networks - 1960s - 1970’s
- The First “Internet” - Mid1980’s
- Commercialization of the Internet - early 1990’s
- Uniquity of the Internet - 1996 and beyond

Before the Internet

- We connected computers directly to one another using leased phone lines
- These were very expensive - and the longer the connection the more expensive it was
- The phone companies made the rules

Phone Line Networking



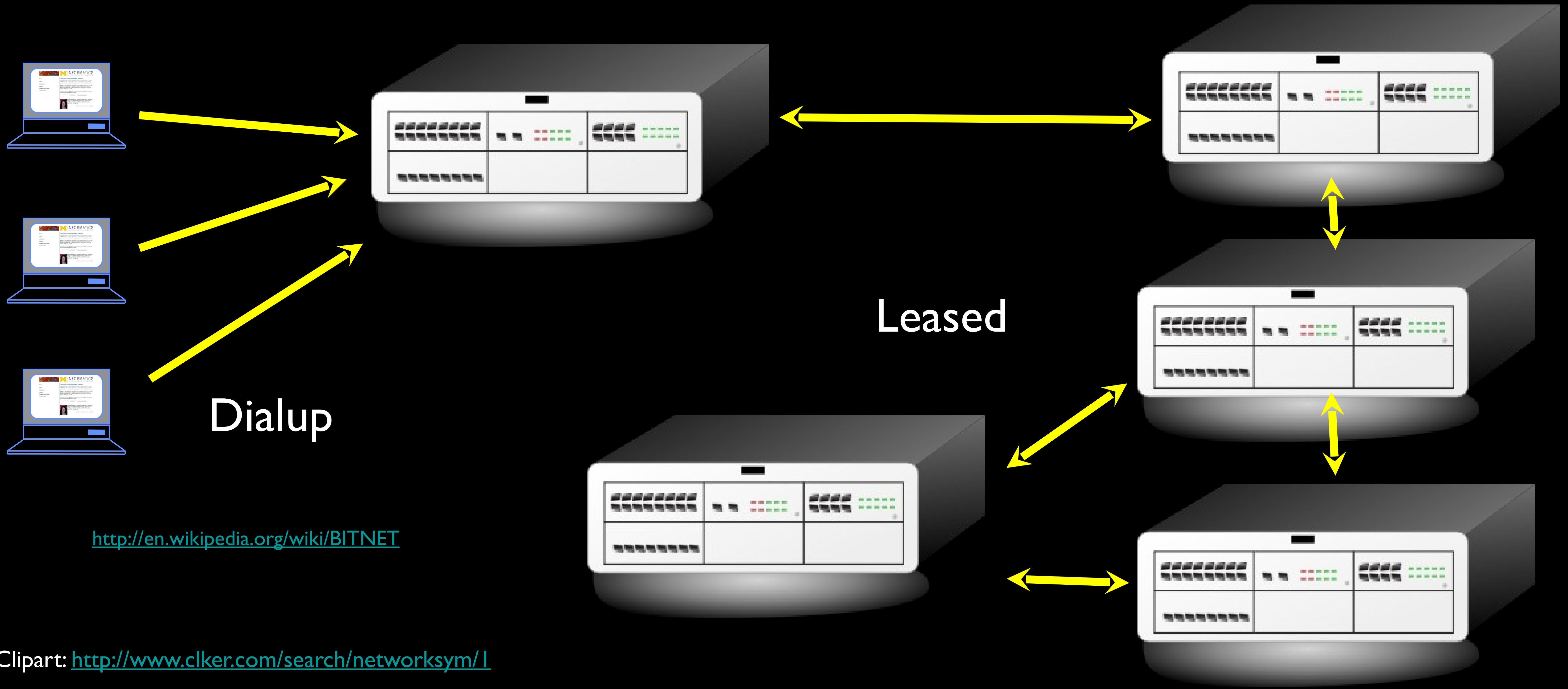
Clipart: http://www.clker.com/search/networksym/1
Modem: http://en.wikipedia.org/wiki/Modem

Phone Line Networking

- You were happy to connect to one computer without having to walk across campus
- You could call other computers long distance
- Pretty Common in the 1970's

http://www.youtube.com/watch?v=E4IztV7M3jI

Store and Forward Networking



http://en.wikipedia.org/wiki/BITNET

Clipart: http://www.clker.com/search/networksym/1

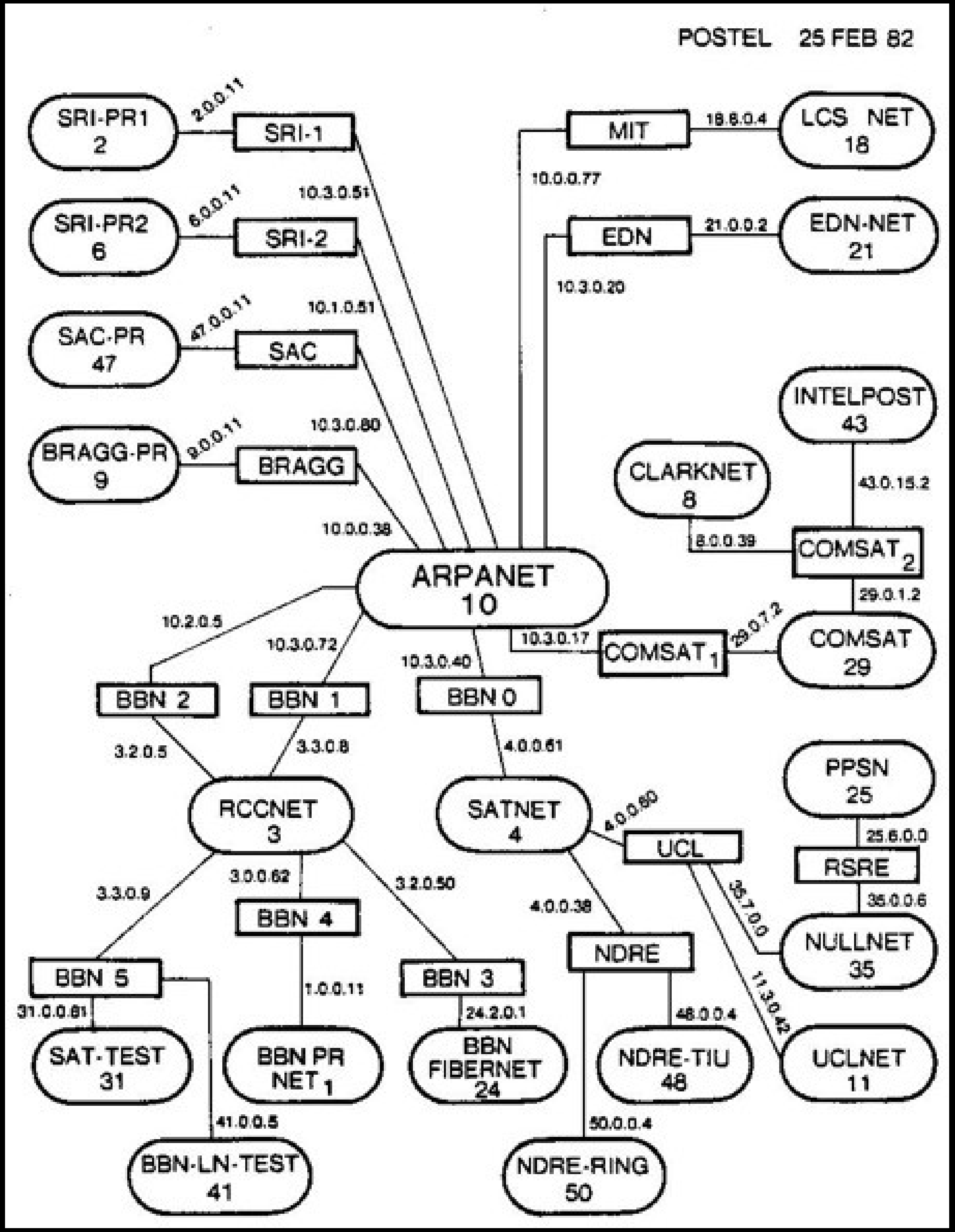
Store and Forward Networking

- Typically specialized in Mail
- E-Mail could make it across the country in 6-hours to about 2 days
- You generally focused your life on one computer
- Early 1980's

http://en.wikipedia.org/wiki/IBM_3270

Research Networks 1960-1980’s

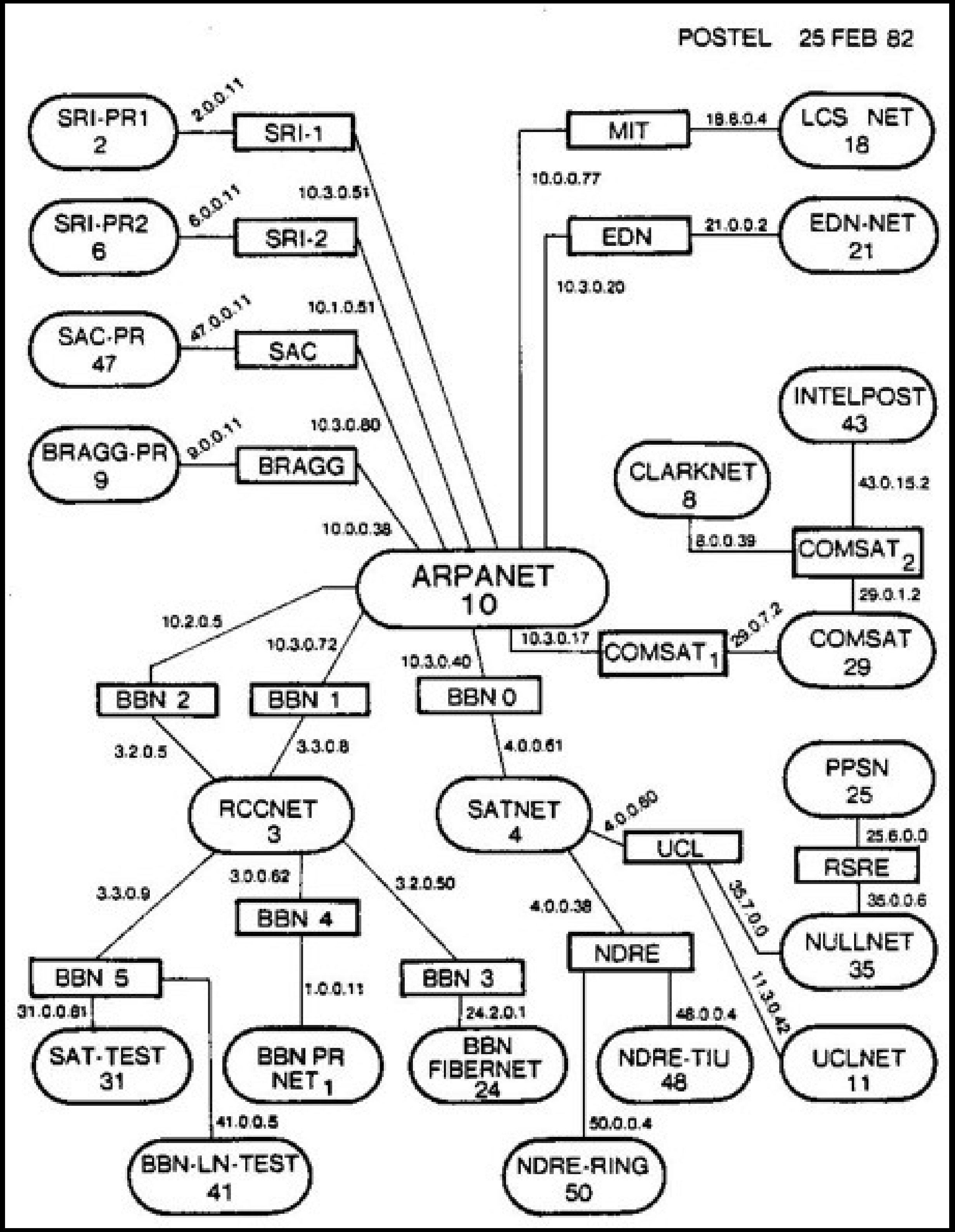
- How can we avoid having a direct connection between all pairs of computers? How can we share expensive long distance leased phone lines?
- The question was how can many computers share a common network?
- How can we dynamically handle outages?



Source: Unknown

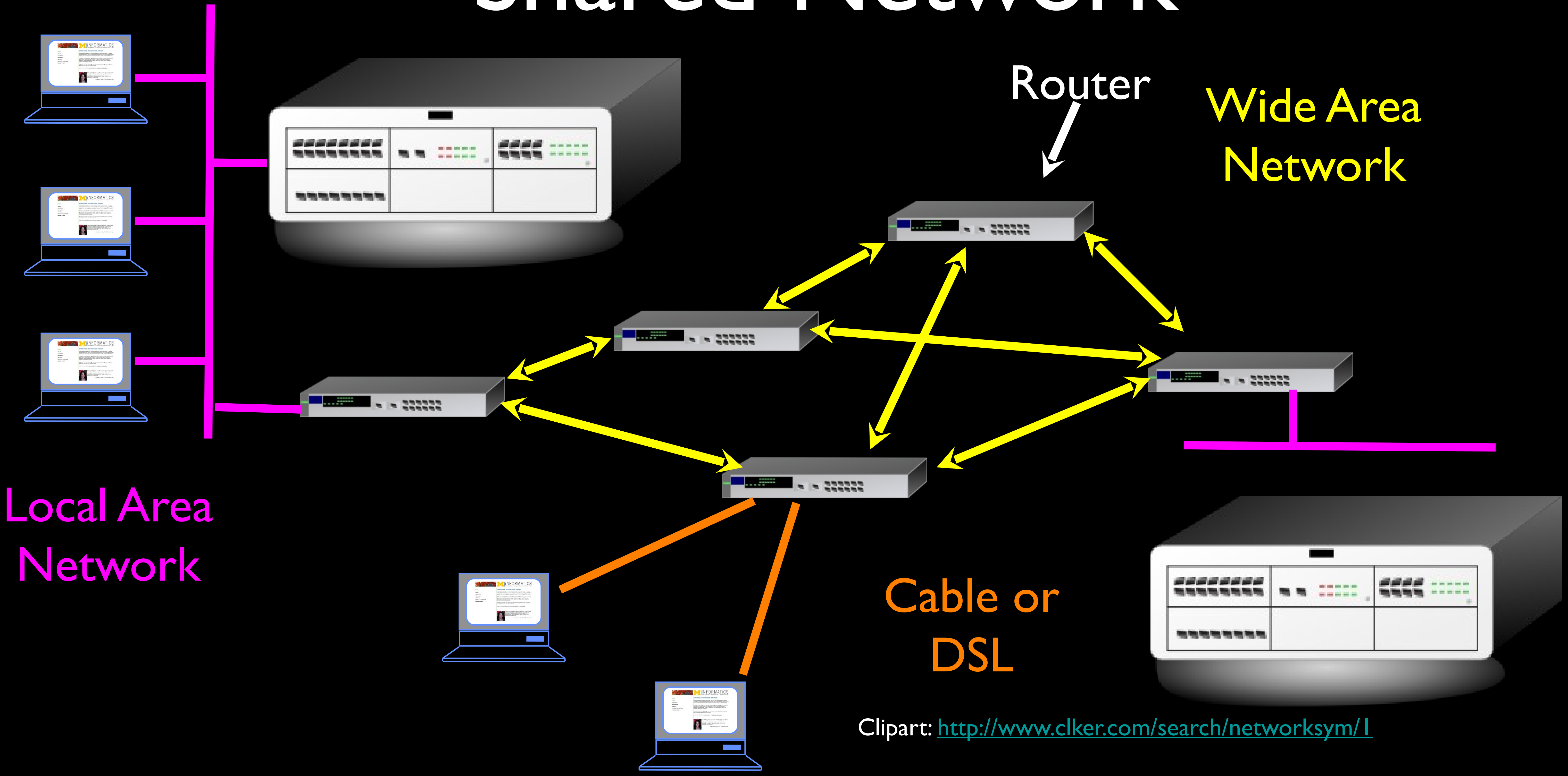
Research Networks

- What is the best protocol to solve all of the competing needs?
- Should we go with commercial solutions from IBM or Digital Equipment? Or should we build something “open”?
- ARPANET was an exclusive Club



Source: Unknown

Shared Network



Clipart: http://www.clker.com/search/networksym/1

Supercomputers...

- As science needed faster and faster computers, more universities asked for their own Multimillion dollar supercomputer
- The National Science Foundation asked, "Why not buy a few supercomputers, and build up a national shared network?"

NCSA - Innovation

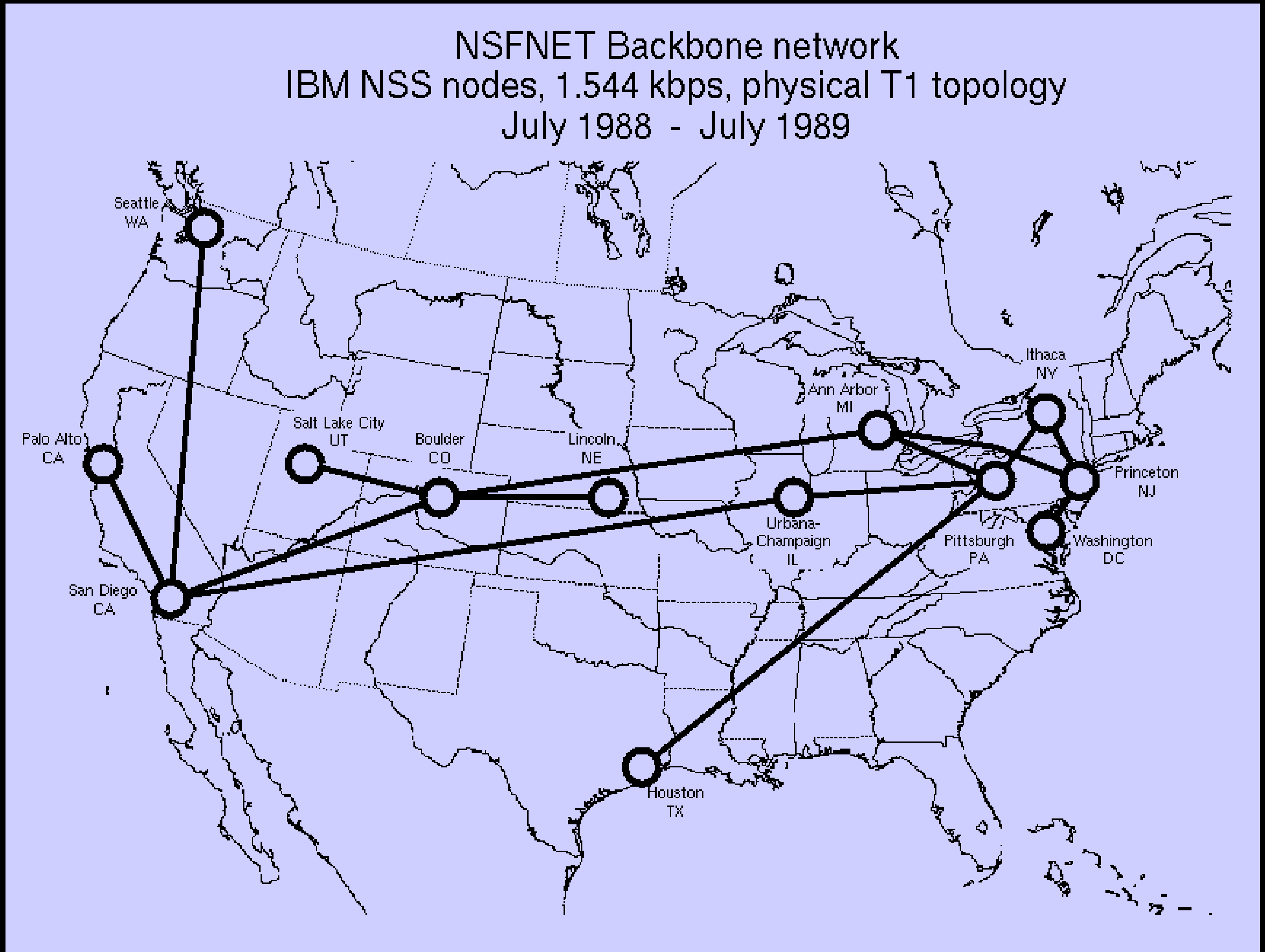
- We now “assume” the Internet and the Web - it was not so easy...
- A number of breakthrough innovations came from the National Center for Supercomputing Applications at Urbana-Champaign, Illinois
- High Performance Computing and the Internet were deeply linked

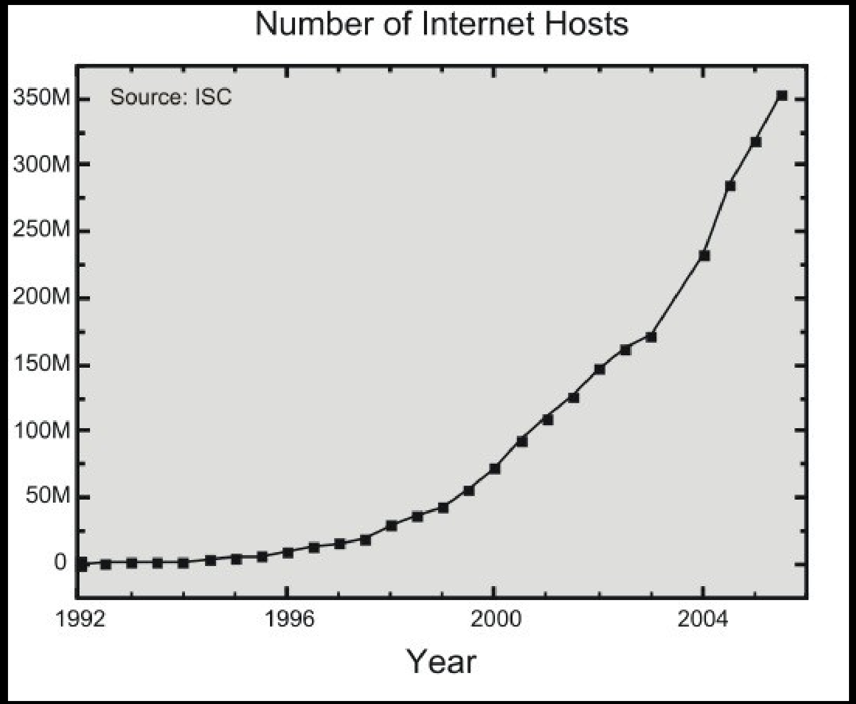
NSF Net

- NSFNet was the first network that was "inclusive"
- Standardized on TCP/IP
- Initially the goal was all research universities
- In the early 1990's commercial companies (Internet Service Providers) could join and resell service

NSFNET T1 Backbone and Regional Networks, 1991

http://virdir.ncsa.uiuc.edu/virdir/raw-material/networking/nsfnet/NSFNET_1.htm

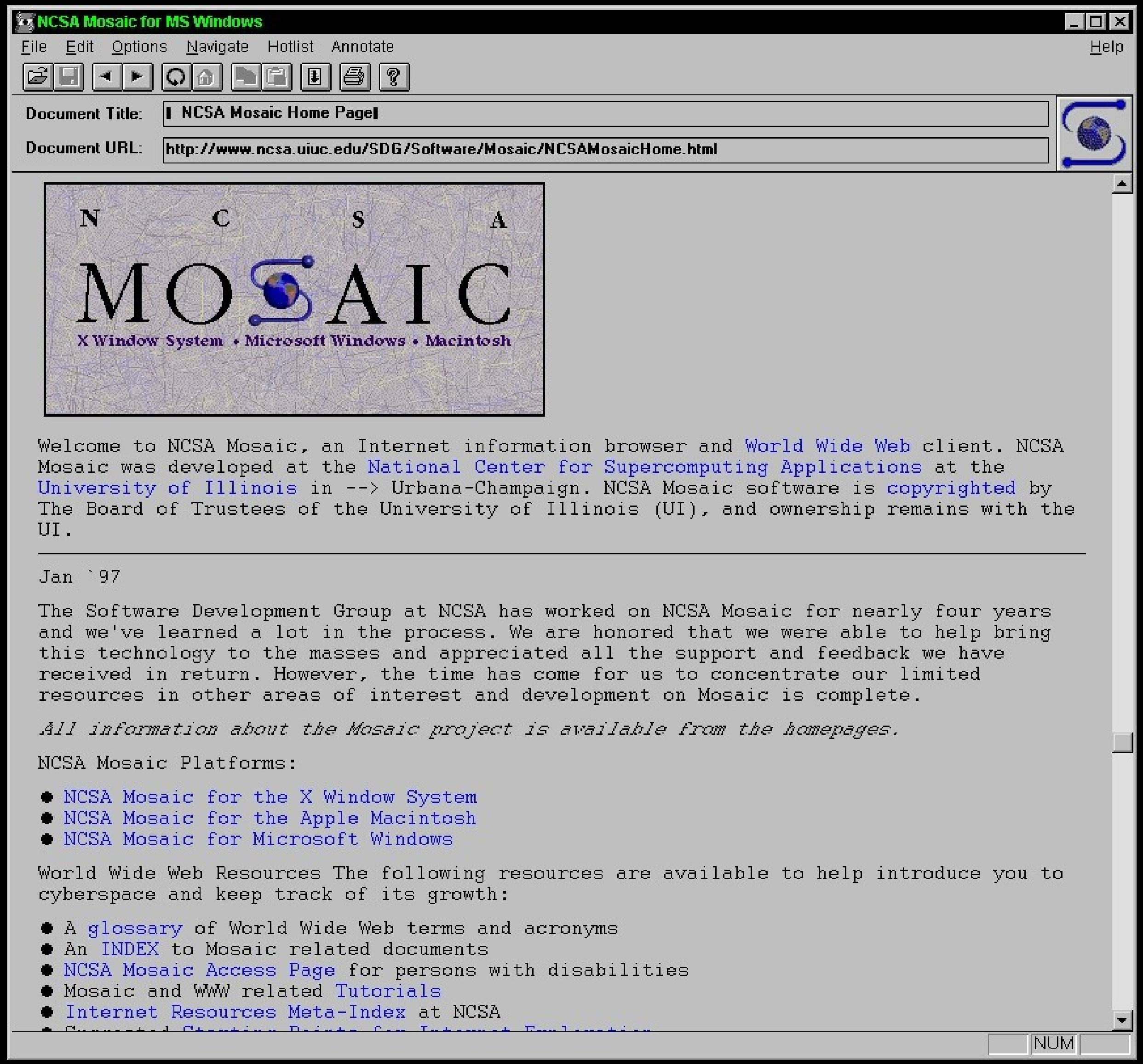




Source: http://hpwren.ucsd.edu/~hwb/NSFNET/NSFNET-200711Summary/

The Explosive Growth of the Web

- The web was invented in the early 1990's
- Growing in Academia 1993
- Growing everywhere 1994 - 1995
- Cable Modems to the home started in the mid 1990's



http://gladiator.ncsa.uiuc.edu/Images/press-images/mosaic.1.0.tif

Mosaic - Netscape - Firefox

- The Internet was infrastructure - the web gave the Internet a "user interface
- The Web was invented at CERN by Tim Berners-Lee and Robert Cailliau
- Mosaic was the first "consumer" web browser developed at NCSA

http://www.youtube.com/watch?v=x2GylLq59rI

http://www.youtube.com/watch?v=An7koSYfEpI

The Modern Internet

- In the late 1990's in the boom there was a great deal of Fiber optic that was installed in the US
- High speed and long distance were cheap and common
- Many national backbone networks emerged - commercial, government, academic, etc
- These networks swap data at "peering points" so we see one seamless Internet - after about 1999 - this was all pretty boring - it just worked

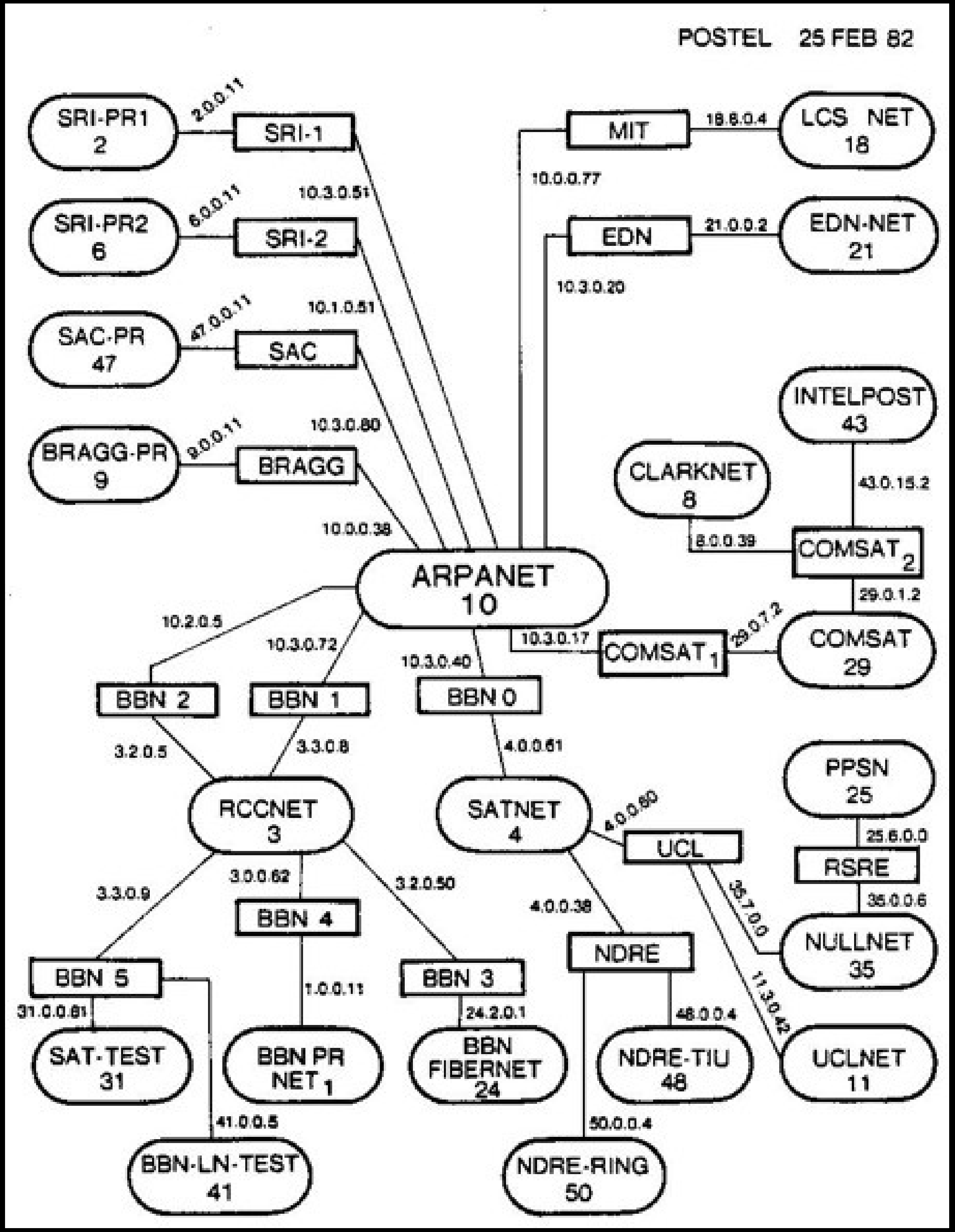
http://en.wikipedia.org/wiki/Internet_Exchange_Point

Internet Protocol Suite

http://en.wikipedia.org/wiki/Internet_Protocol_Suite

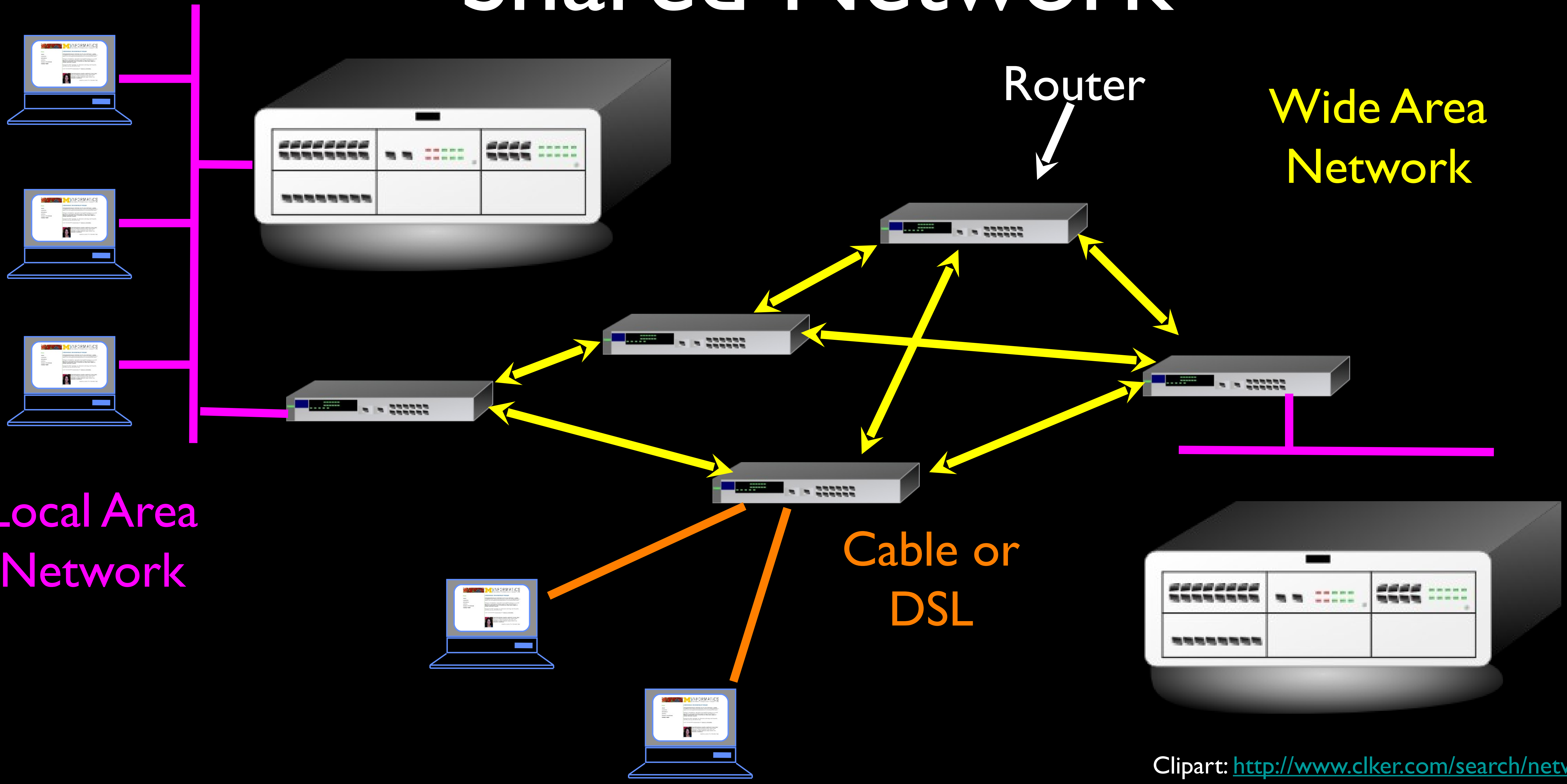
Research Networks 1960-1980's

- How can we avoid having a direct connection between all pairs of computers? How can we share expensive long distance leased phone lines?
- The question was how can many computers share a common network?
- How can we dynamically handle outages?



Source: Unknown

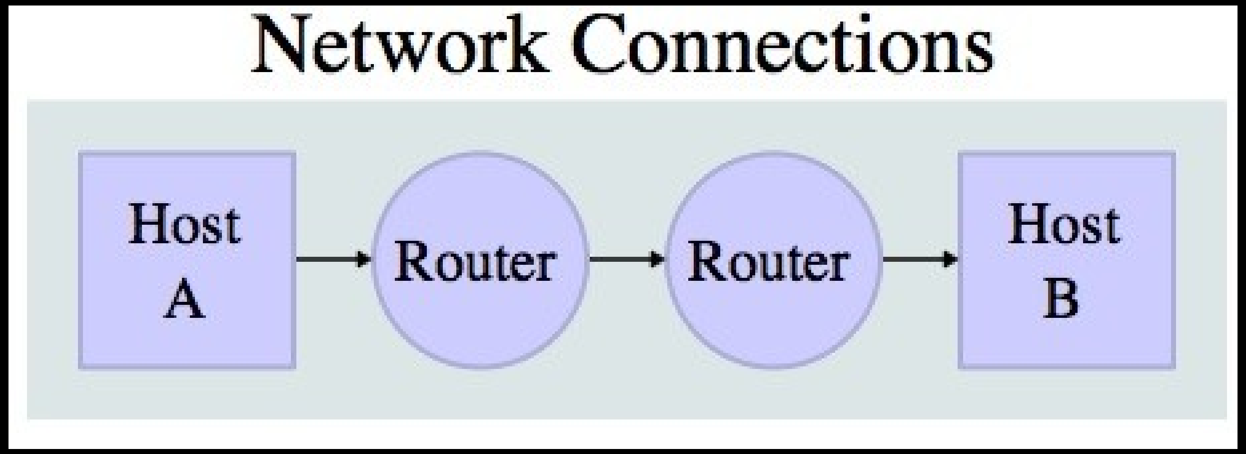
Shared Network



Clipart: http://www.clker.com/search/networksym/1

Shared Networks

- In order to keep cost low and the connections short geographically - data would be forwarded through several routers.
- Getting across the country usually takes about 10 "hops"
- Network designers continually add and remove links to "tune" their networks



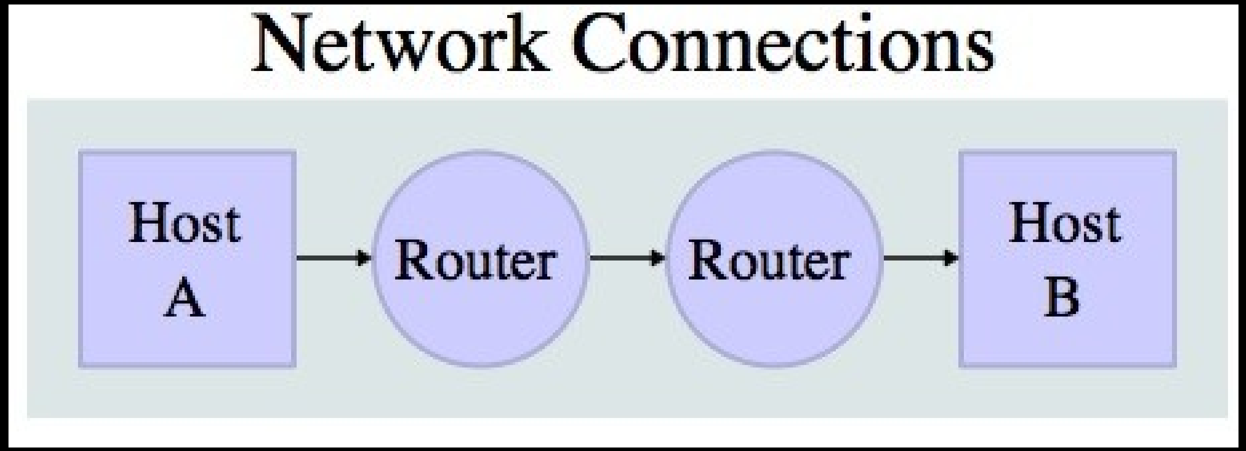
Source: http://en.wikipedia.org/wiki/Internet_Protocol_Suite

Layered Network Model

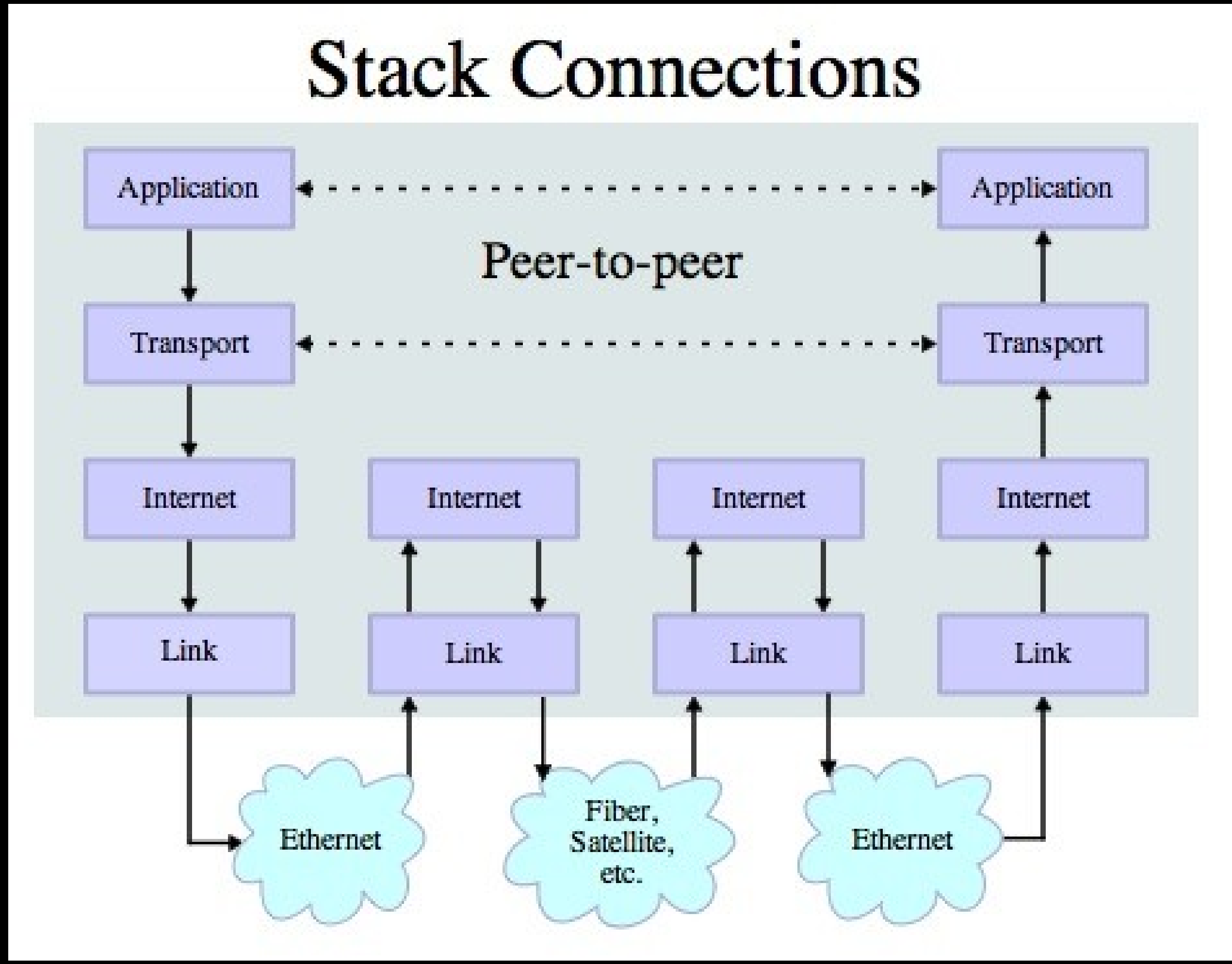
- A layered approach allows the problem of implementing a network to be broken into more manageable sub problems
- For example the IP layer is allowed to lose a packet if things go bad
- It is TCP's Responsibility to store and retransmit data.

Application Layer Web, E-Mail, File Transfer
Transport Layer (TCP) Reliable Connections
Internetwork Layer (IP) Simple, Unreliable
Link Layer (Ethernet, WiFi) Physical Connections

Layered Architecture



- The Physical and Internet Layers are like trucks - they haul stuff and get it to the right loading dock - it takes multiple steps
- The Transport layer checks to see if the trucks made it and send the stuff again ig necessary



Source: http://en.wikipedia.org/wiki/Internet_Protocol_Suite

Internet Standards

- The standards for all of the Internet protocols (inner workings) are developed by an organization
- Internet Engineering Task Force (IETF)
- www.ietf.org
- Standards are called “RFCs” - “Request for Comments”

INTERNET PROTOCOL

DARPA INTERNET PROGRAM

PROTOCOL SPECIFICATION

September 1981

The internet protocol treats each internet datagram as an independent entity unrelated to any other internet datagram. There are no connections or logical circuits (virtual or otherwise).

The internet protocol uses four key mechanisms in providing its service: Type of Service, Time to Live, Options, and Header Checksum.

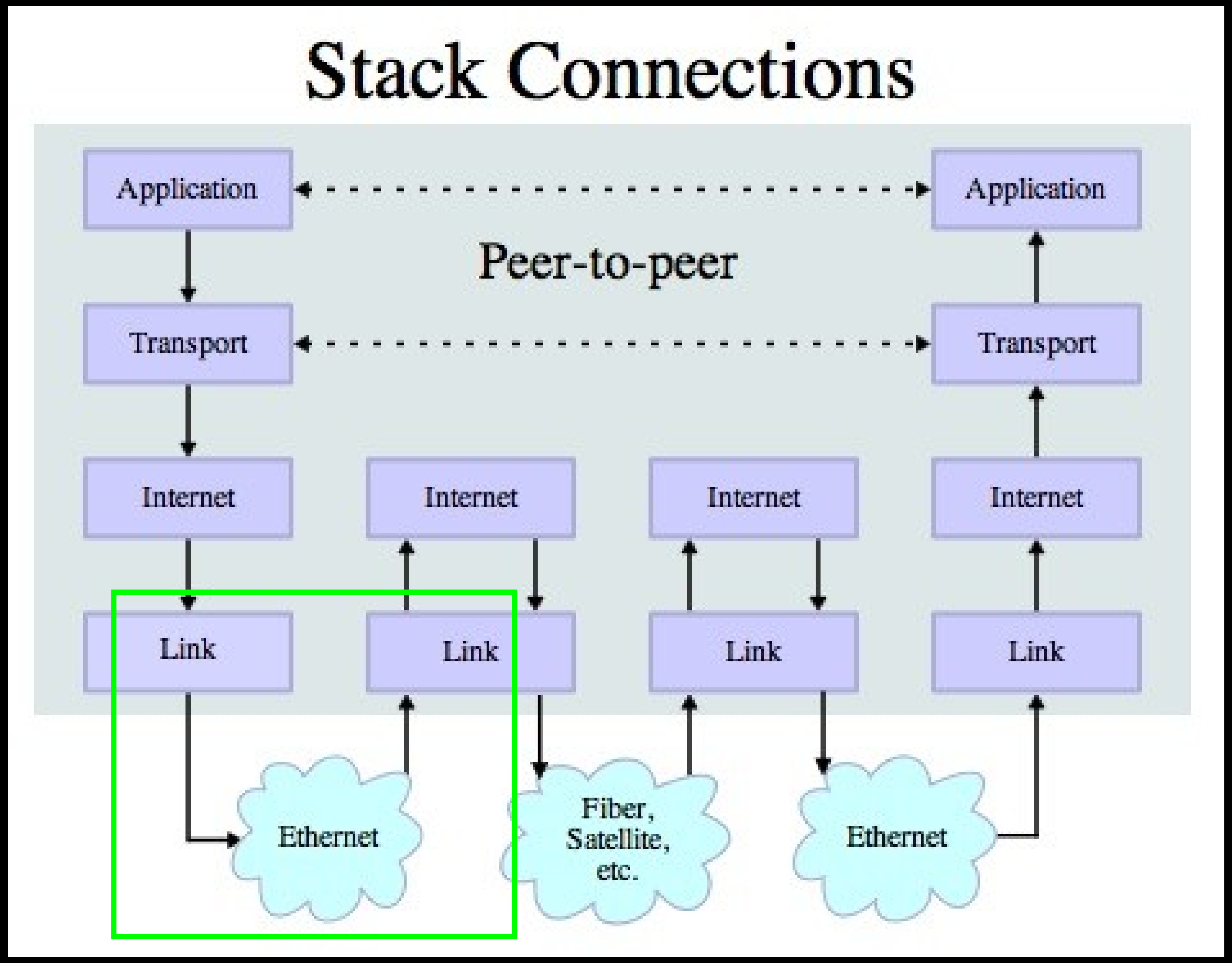
Source: http://tools.ietf.org/html/rfc791

Link Layer

http://en.wikipedia.org/wiki/Ethernet
http://en.wikipedia.org/wiki/Wireless_LAN
http://en.wikipedia.org/wiki/Category:Link_protocols

Link Layer (aka Physical Layer)

- As your data crosses the country may use a different physical medium for each "hop"
- Wire, Wireless, Fiber Optic, etc.
- The link is "one hop" - Is it up or down? Connected or not?
- Very narrow focus - no view at all of the "whole Internet"



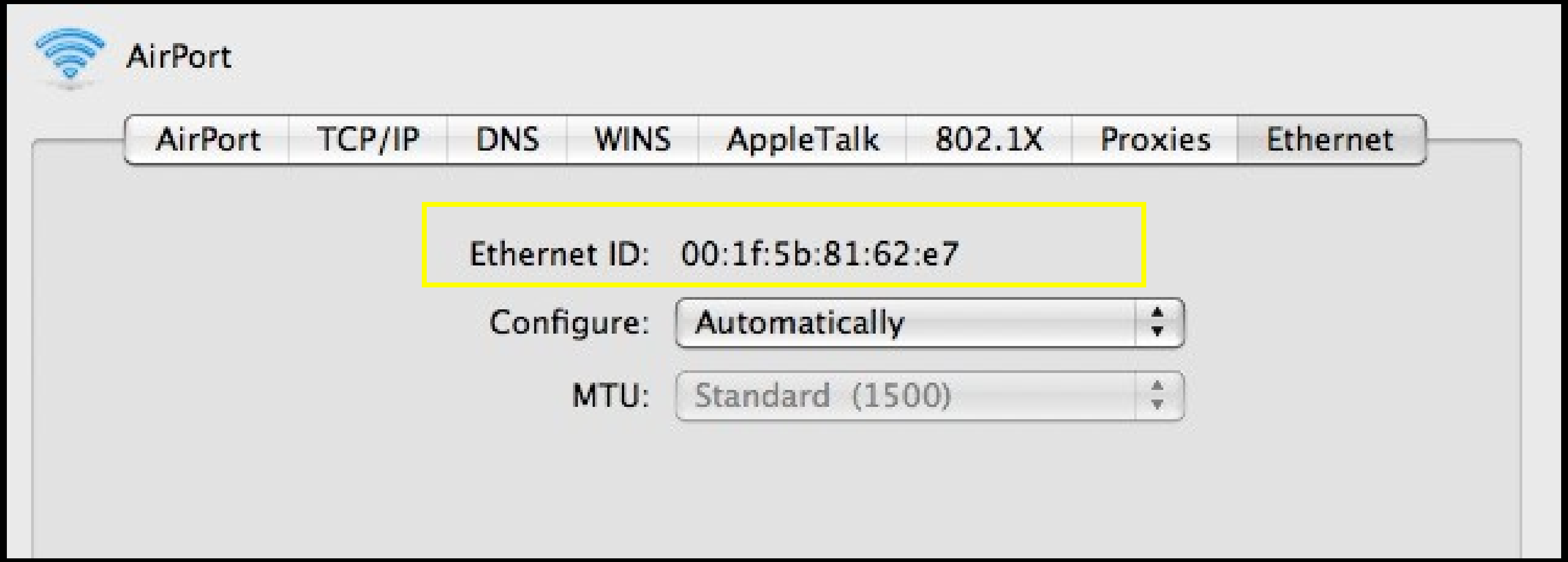
Source: http://en.wikipedia.org/wiki/Internet_Protocol_Suite

Common Link Technologies

- Ethernet
- WiFi
- Cable modem
- DSL
- Satellite
- Optical fiber

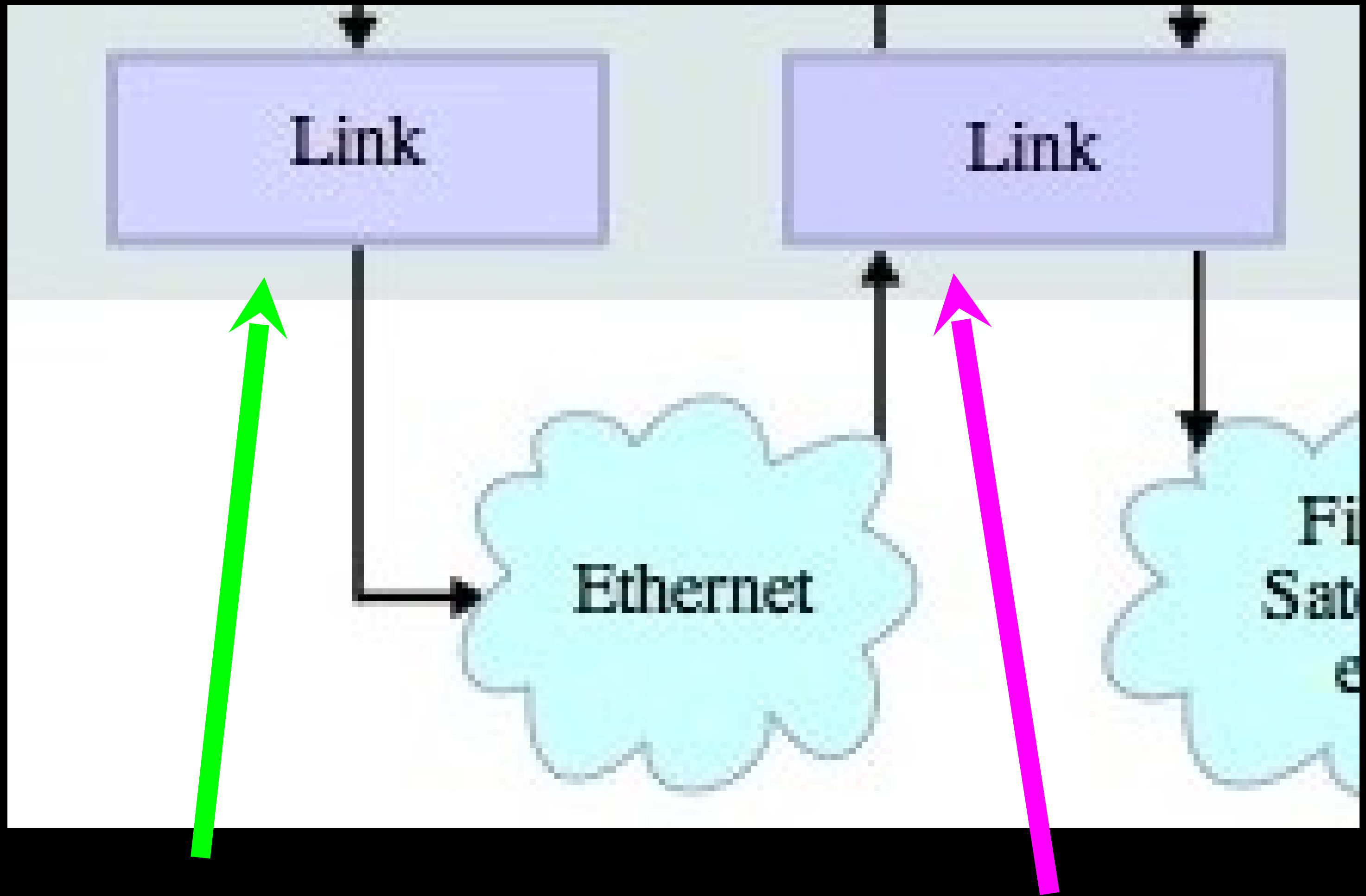
Link Layer Addresses

- Many physical layer devices have addresses built in to them by the manufacturer
  - Ethernet
  - Wireless Ethernet (Wifi)



Link Layer

- Physical addresses are to allow systems to identify themselves on the ends of a single link
- Physical addresses go no farther than one link
- Sometimes links like Wifi and Wired Ethernet are shared with multiple computers



0f:21:63:12:b3:1a 98:2f:4e:78:01:b4

Source: http://en.wikipedia.org/wiki/Internet_Protocol_Suite

CSMA/CD - Ethernet Protocol

- Carrier Sense Media Access with Collision Detection
- Because the Ethernet network is shared, each station “hears” all traffic to avoid garbled messages, systems must observe “rules”
- Ethernet rules are simple
- Listen for any traffic, wait until “silent”
- Begin transmitting data
- Listen for your own data
- If you cannot hear your own data clearly, assume a collision, stop and wait before trying again
- Each system waits a different amount of time to avoid “too much politeness”

Internet Protocol Layer

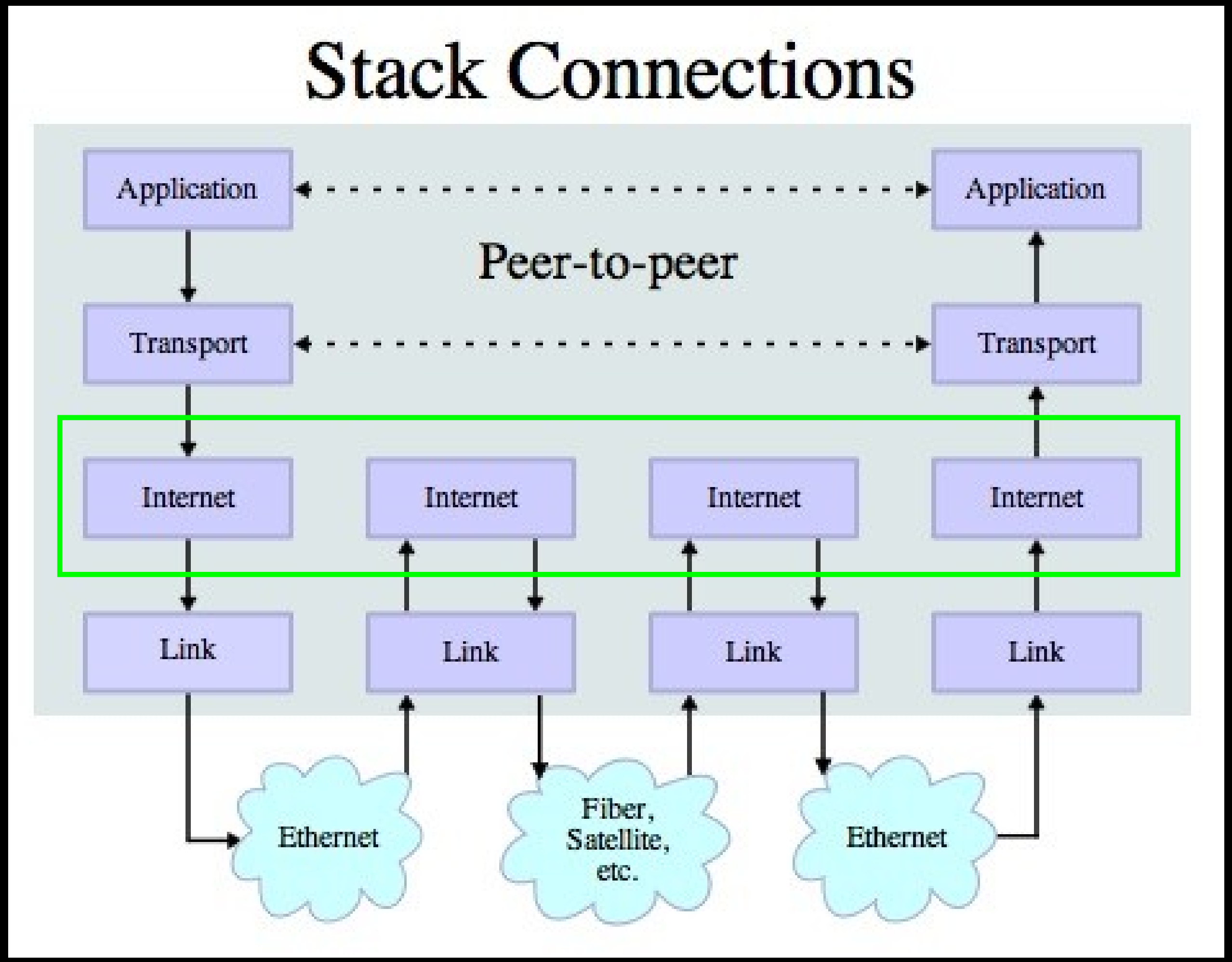
http://en.wikipedia.org/wiki/Internet_Protocol

http://en.wikipedia.org/wiki/Traceroute

http://en.wikipedia.org/wiki/Ping

Internet Protocol Layer

- IP Gets your data from this computer to the other computer half way across the world without knowing anything about the purpose of the data
- IP Is best effort - it is OK to drop data if things go bad...



Source: http://en.wikipedia.org/wiki/Internet_Protocol_Suite

Internetwork Layer (IP)

- The goal of the Internetwork layer is to transport data from one end-user system to another end-user systems hopping across as many physical connections as necessary.
- Internetwork Layer provides a mechanism to connect many LANs together effectively
- IP is an unreliable datagram protocol
- Packets may arrive out of order or not at all
- IP is responsible for making connections between millions of computers worldwide without using broadcast packets

IP Addresses

- The IP address is the worldwide number which is associated with one particular workstation or server
- Every system which will send packets directly out across the Internet must have a unique IP address
- IP addresses are based on where station is connected
- IP addresses are not controlled by a single organization - address ranges are assigned
- They are like phone numbers – they get reorganized once in a great while

IP Address Format

- Four numbers with dots - each number 1-255 (32 bits)
- Kind of like phone numbers with an "area code"
- The prefix of the address is "which network"
- WHile the data is traversing the Internet - all that matters is the network number

(734) 764 1855

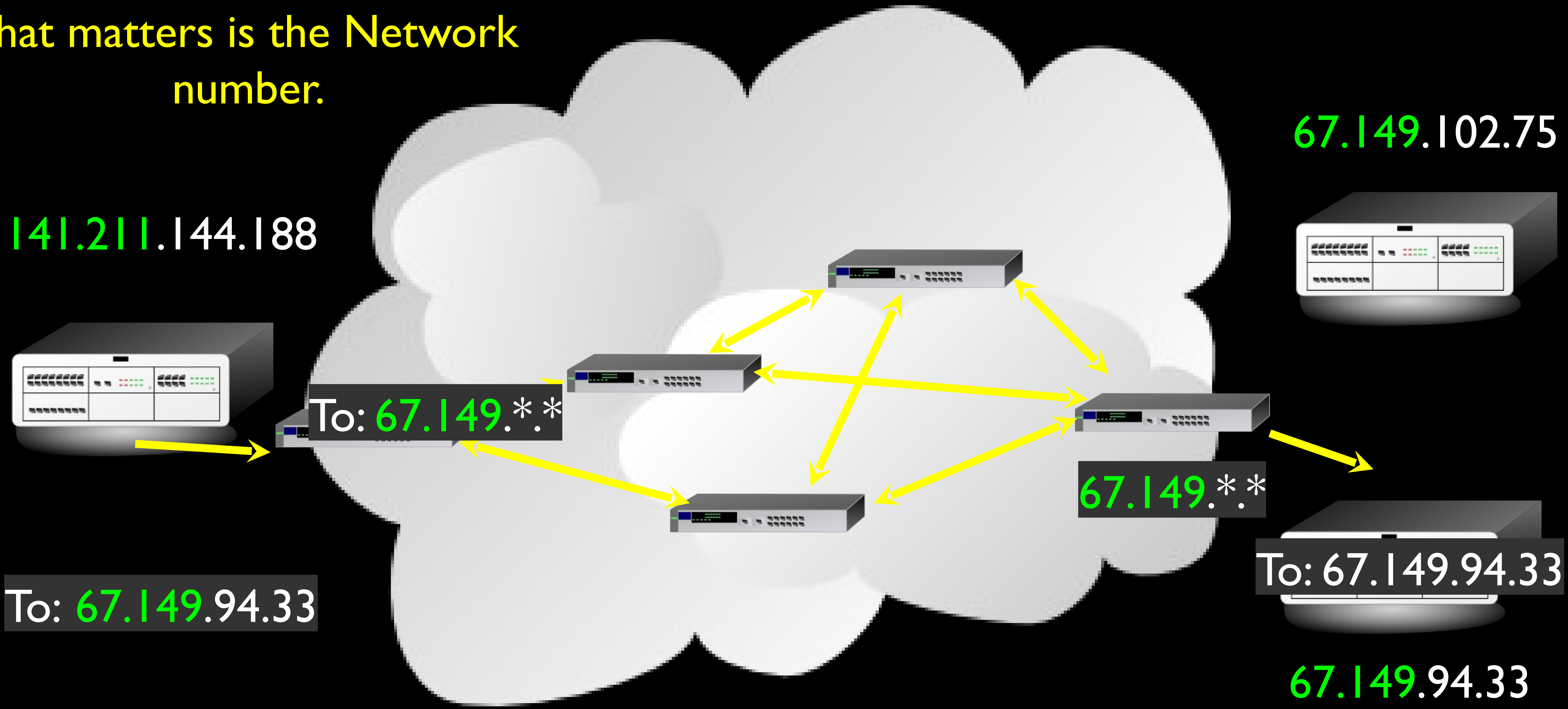
Area code

141.211.144.188

Network Number

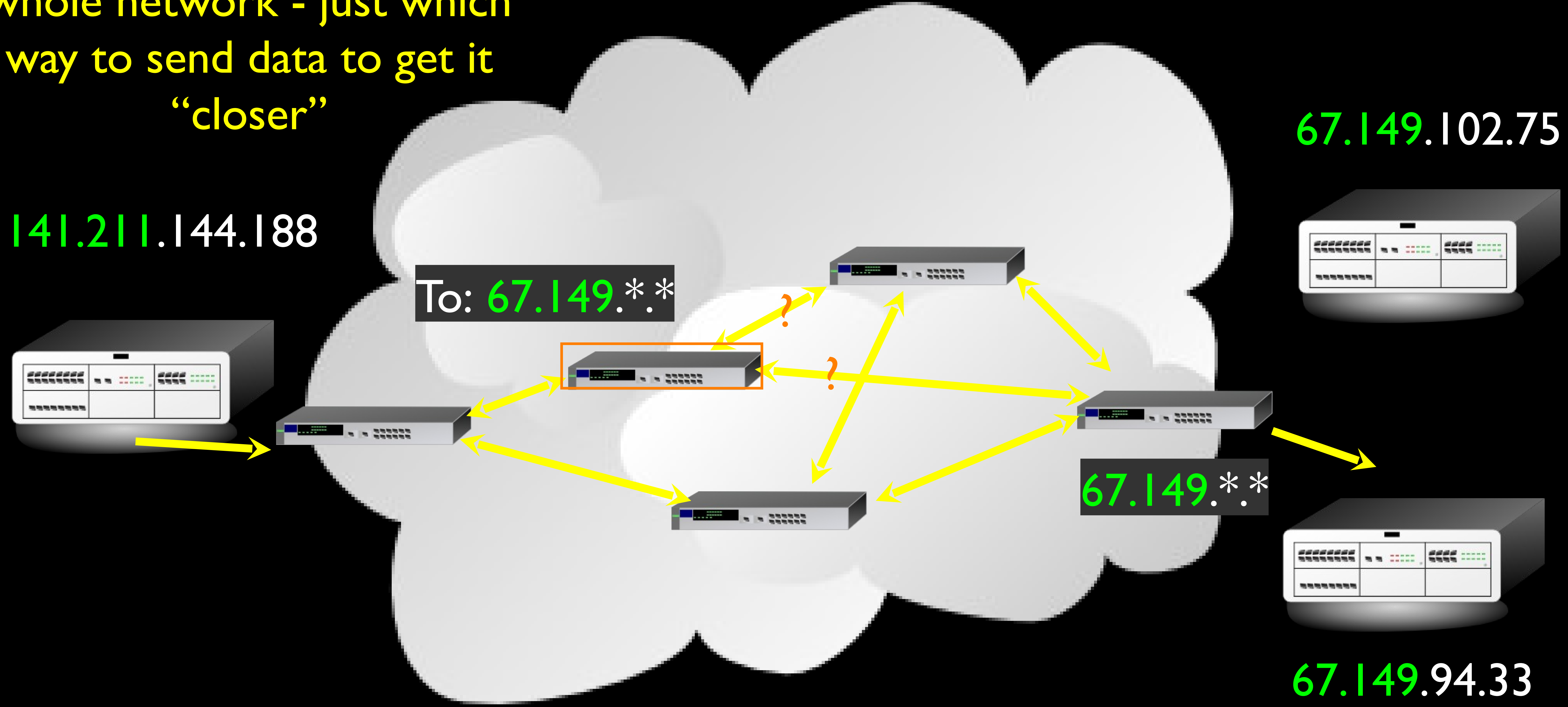
141.211.*.*

While in the network, all that matters is the Network number.



Clipart: http://www.clker.com/search/networksym/1

No single router knows the whole network - just which way to send data to get it “closer”

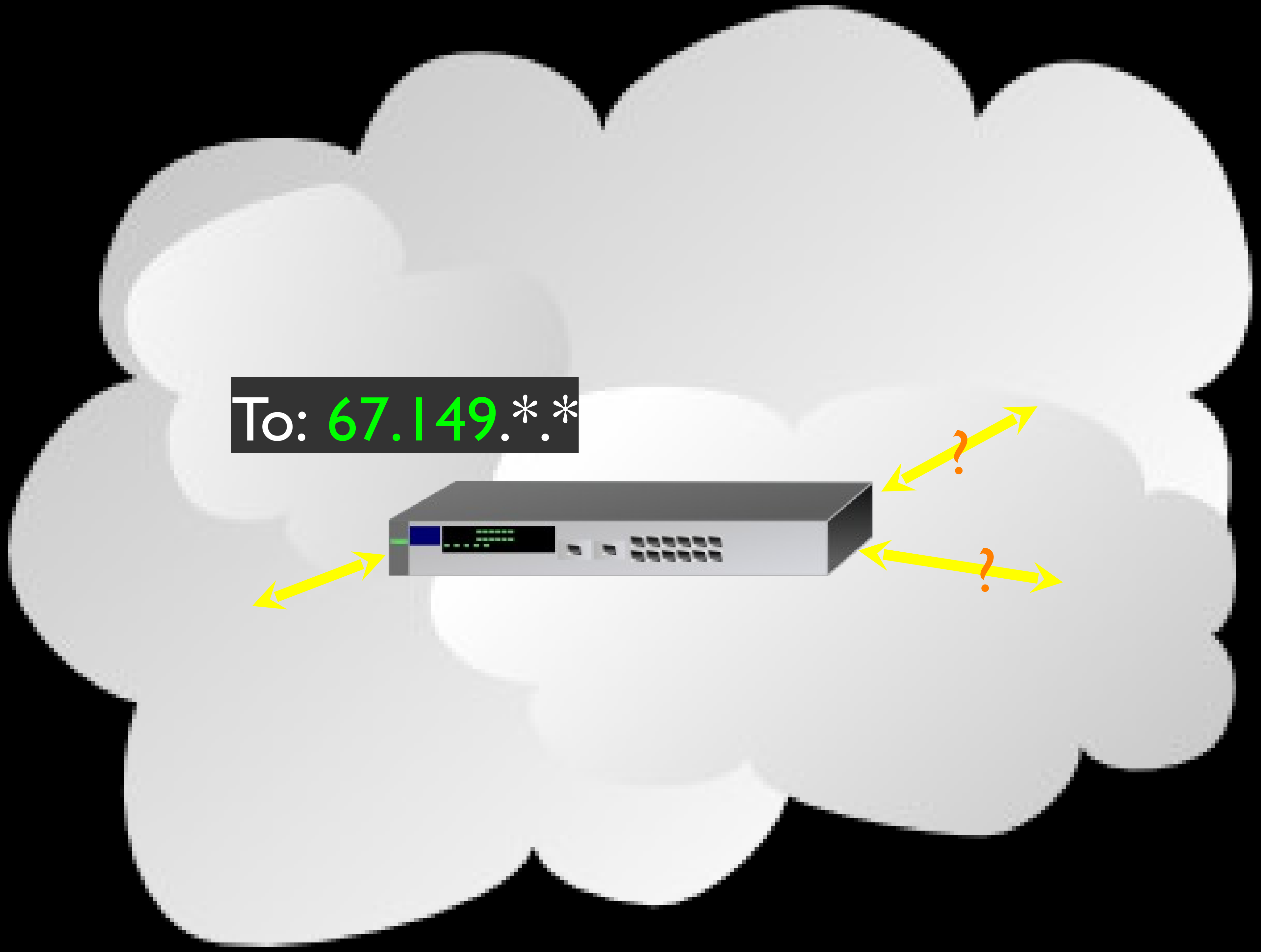


Clipart: http://www.clker.com/search/networksym/1

No single router needs to know the whole network - just which way to send data to get it "closer"

Actually either path will work - but one path is better.

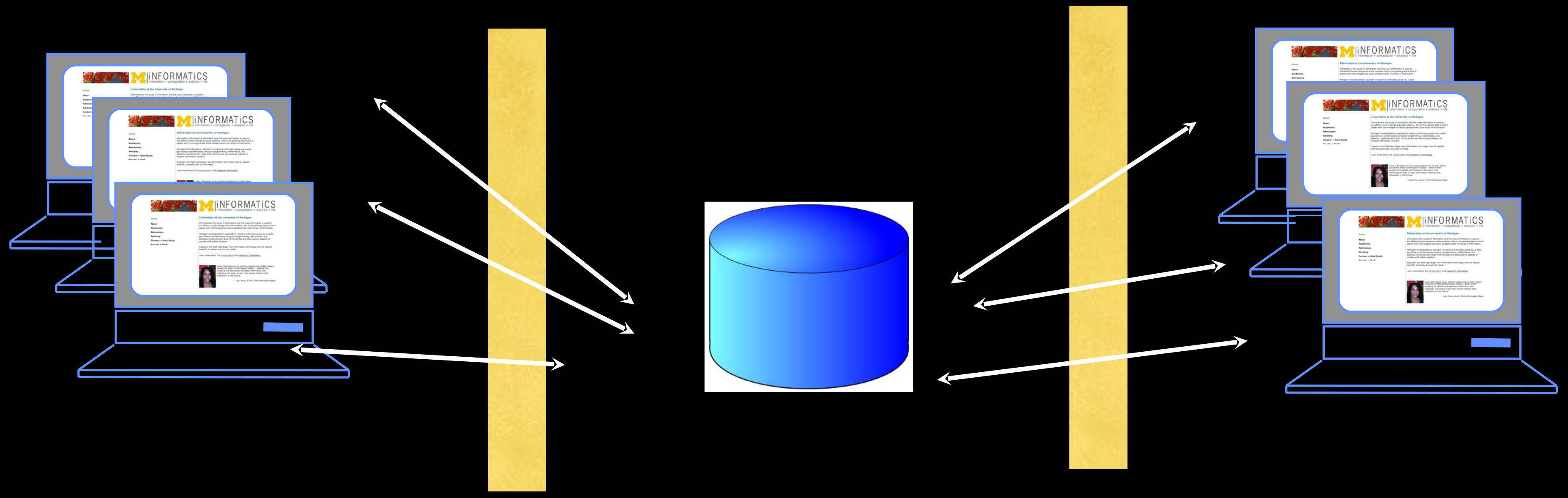
Routers keep dynamic tables of where the traffic they are processing needs to go...



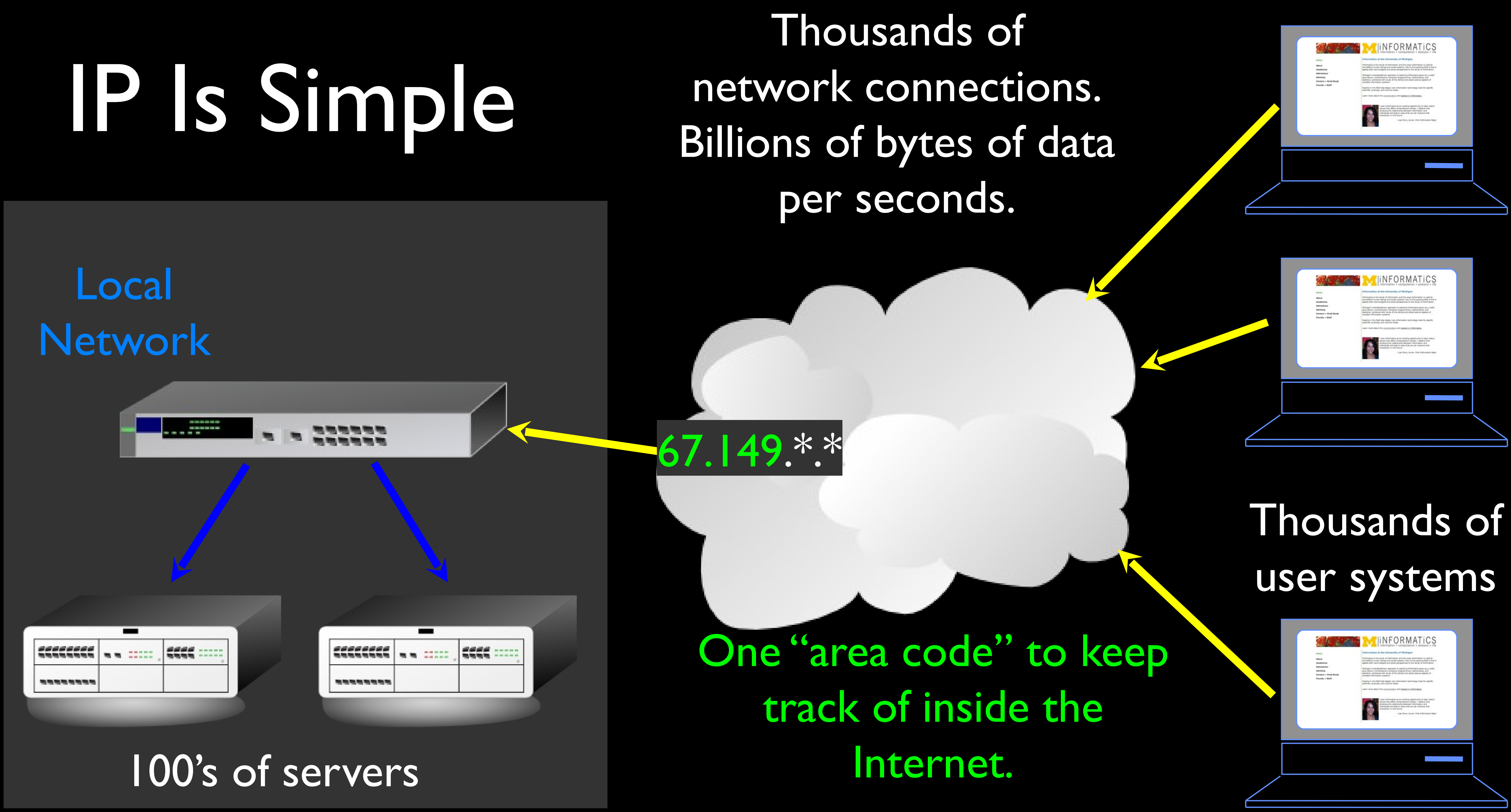
Clipart: http://www.clker.com/search/networksym/1

System to System IP

- Regardless of the number of connections between two systems, the traffic is transported across the internet as a single IP address - It is the responsibility of TCP to separate (de-multiplex) each stream on each system



IP Is Simple



Clipart: http://www.clker.com/search/networksym/1

Peering into the Internet

- Most systems have a command that will reveal the route taken across the internet (traceroute on Mac and tracert on Windows)
- Each IP packet has a field called "Time to Live" - TTL
- The TTL is used to deal with loops in the network - normally if routers got confused and ended up with a loop - the network would clog up rapidly.

Clipart: http://www.clker.com/search/networksym/1

How Traceroute Works

- Normal packets are sent with a Time to Live (TTL) of 255 hops
- Trace route sends a packet with TTL=1, TTL=2, ...
- So each packet gets part-way there and then gets dropped and traceroute gets a notification of where the drop happens
- This builds a map of the nodes that a packet visits when crossing the Internet.

Traceroute

$ traceroute www.stanford.edutraceroute to www5.stanford.edu (171.67.20.37), 64 hops max, 40 byte packets 1 141.211.203.252 (141.211.203.252) 1.390 ms 0.534 ms 0.490 ms 2 v-bin-seb.r-bin-seb.umnet.umich.edu (192.122.183.61) 0.591 ms 0.558 ms 0.570 ms 3 v-bin-seb-i2-aa.merit-aa2.umnet.umich.edu (192.12.80.33) 6.610 ms 6.545 ms 6.654 ms 4 192.122.183.30 (192.122.183.30) 7.919 ms 7.209 ms 7.122 ms 5 so-4-3-0.0.rtr.kans.net.internet2.edu (64.57.28.3 17.672 ms 17.836 ms 17.673 ms 6 so-0-1-0.0.rtr.hous.net.internet2.edu (64.57.28.57) 31.800 ms 41.967 ms 31.787 ms 7 so-3-0-0.0.rtr.losa.net.internet2.edu (64.57.28.44) 63.478 ms 63.704 ms 63.710 ms 8 hpr-lax-hpr--i2-newnet.cenic.net (137.164.26.132) 63.093 ms 63.026 ms 63.384 ms svl-hpr--lax-hpr-10ge.cenic.net (137.164.25.13) 71.242 ms 71.542 ms 76.282 ms10 oak-hpr--svl-h 10ge.cenic.net (137.164.25.9) 72.744 ms 72.243 ms 72.556 ms11 hpr-stan-ge--oak-hpr.cenic.net (137.164.27.158) 73.763 ms 73.396 ms 73.665 ms12 bbra-rtr.Stanford.EDU (171.64.1.134) 73.57 ms 73.682 ms 73.492 ms13 * * *14 www5.Stanford.EDU (171.67.20.37) 77.317 ms 77.128 ms 77.648 ms

Traceroute

$ traceroute www.msu.edutraceroute to www.msu.edu (35.8.10.30), 64 hops max, 40 byte packets 141.211.203.252 (141.211.203.252) 2.644 ms 0.973 ms 14.162 ms 2 v-bin-seb.r-bin-seb.umnet.umich.edu (192.122.183.61) 1.847 ms 0.561 ms 0.496 ms 3 v-bin-seb-i2-aa.merit-aa2.umnet.umich.edu (192.12.80.33) 6.490 ms 6.499 ms 6.529 ms 4 lt-0-3-0x1.eq-chi2.mich.net (198.108.23.121) 8.096 ms 8.113 ms 8.103 ms 5 xe-0-0-0x23.msu6.mich.net (198.108.23.213) 7.831 ms 7.962 ms 7.965 ms 6 192.122.183.227 (192.122.183.227) 12.953 ms 12.339 ms 10.322 ms 7 cc-t1-ge1-23.net.msu.edu (35.9.101.209) 9.522 ms 9.406 ms 9.817 ms 8 * * *

Traceroute

$ traceroute www.pku.edu.cntraceroute: Warning: www.pku.edu.cn has multiple addresses; using 162.105.129.104traceroute to www.pku.edu.cn (162.105.129.104), 64 hops max, 40 byte packets 1 141.211.203.252 (141.211.203.252) 1.228 ms 0.584 ms 0.592 ms 2 v-bin-seb.r-bin-seb.umnet.umich.edu (192.122.183.61) 0.604 ms 0.565 ms 0.466 ms 3 v-bin-seb-i2-aa.merit-aa2.umnet.umich.edu (192.12.80.33) 7.511 ms 6.641 ms 6.588 ms 4 192.122.183.30 (192.122.183.30) 12.078 ms 6.989 ms 7.619 ms 5 192.31.99.133 (192.31.99.133) 7.666 ms 8.95 ms 17.861 ms 6 192.31.99.170 (192.31.99.170) 59.275 ms 59.273 ms 59.108 ms 7 134.75.108.2 (134.75.108.209) 173.614 ms 173.552 ms 173.333 ms 8 134.75.107.10 (134.75.107.10) 256.760 134.75.107.18 (134.75.107.18) 256.574 ms 256.530 ms 9 202.112.53.17 (202.112.53.17) 256.761 ms 256.801 ms 256.688 ms10 202.112.61.157 (202.112.61.157) 257.416 ms 257.960 ms 257.74 ms11 202.112.53.194 (202.112.53.194) 256.827 ms 257.068 ms 256.962 ms12 202.112.41.202 (202.112.41.202) 256.800 ms 257.053 ms 256.933 ms

Internet Protocol is the Magic

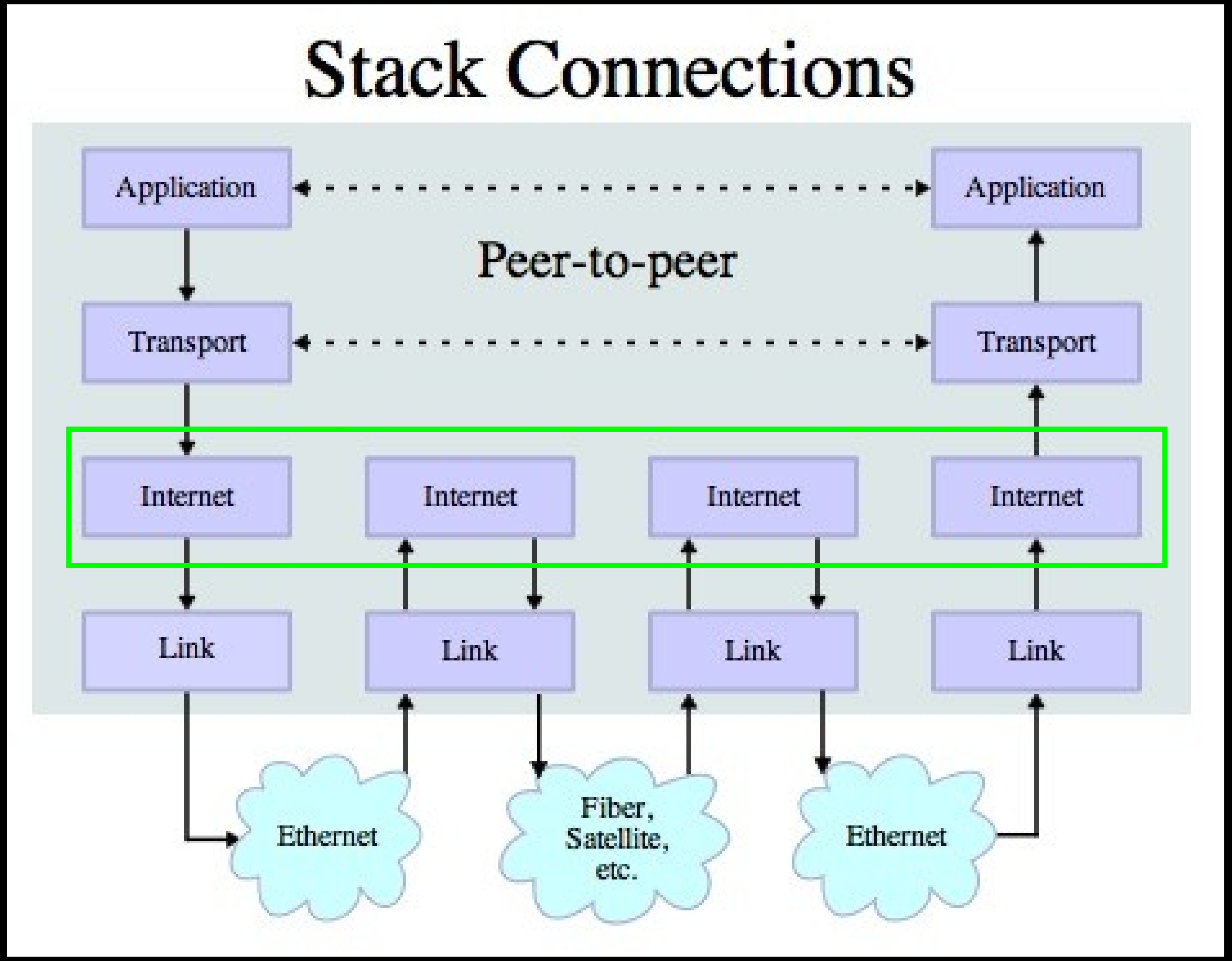
- What it does do: Best effort to get data across bunch of hops from one network to another network
- Best effort to keep track of the good and bad paths for traffic - tries to pick better paths when possible
- What it does not do: No guarantee of delivery - if things go bad - the data vanishes (see "Best Effort")
- This makes it fast and scalable to very large networks - and ultimately "reliable" because it does not try to do too much

Domain Name System (DNS)

http://en.wikipedia.org/wiki/Domain_Name_System
http://en.wikipedia.org/wiki/Nslookup

Domain Name System

- The Domain Name System maps more reabible names like
  - www.umich.edu
- To the more efficiently routable IP address
  - 141.211.32.166



Source: http://en.wikipedia.org/wiki/Internet_Protocol_Suite

Domain Name System

- Numeric addresses like 141.211.63.45 are great for Internet routers but lousy for people
- Each campus ends up with a lot of networks (141.211.*.*, 65.43.21.*)
- Sometimes (rarely) the numbers get reorganized
- For things like home addresses - the numbers change all the time
- When servers physically move they need new IP addresses

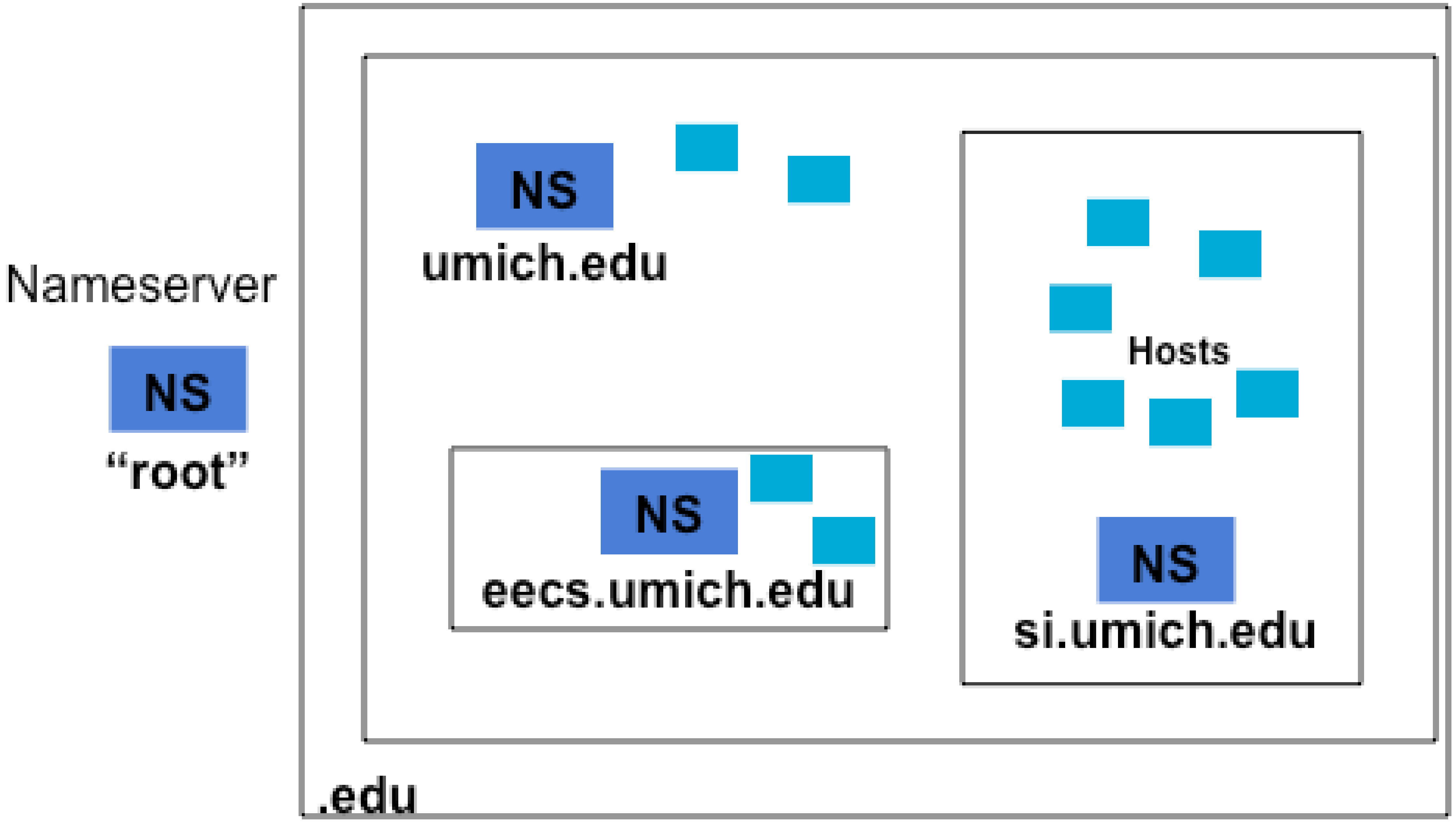
Domain name servers

- Domain name (DNS) servers store an index that maps names like "projects.si.umich.edu" to IP addresses.
- Computers can do lookup operations on DNS servers
- No single server stores all computer names and addresses.
- Instead, a hierarchical system of domains and lookups is used.
- Each suffix forms a domain; ".edu", ".com" etc are called the *top-level domains* (TLD).

DNS: Internet Address Book

- The Domain Name System is a big fast distributed database of Internet names to Internet "phone numbers"
- IP Addresses reflect technical "geography"
  - 141.211.63.44 - read left to right like a phone number
- Domain names reflect organizational structure
  - www.si.umich.edu - read right to left like postal address
  - 311 West Hall, Ann Arbor, MI 440109, USA, Earth

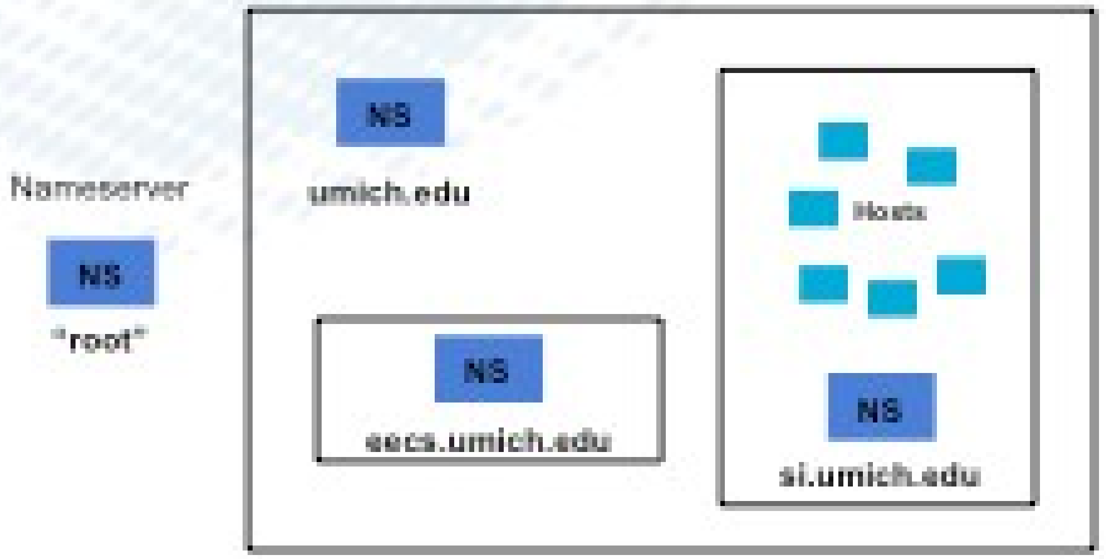
Global Hierarchy of Local Namespaces



BY: Paul Resnick and Rahul Sami.
Slide used with permission.

Namespace Administration

- If SI adds a server or changes a server's IP address
  - SI makes a change to the SI nameserver
- All names in the si.umich.edu domain are resolved by the SI nameserver
  - As a result, name allocation/use requires permission from SI



Caching and Updating

- At any point of time, DNS entries (mappings from a name to an IP address) may be cached in millions of DNS servers and computers.
- A change is made only in the authoritative DNS server (e.g., the si.umich.edu DNS server for bart.si.umich.edu)
- To limit inconsistencies, DNS entries specify an “expiry time” after which they must be thrown out of any cache they are in.
  - – but there still may be inconsistencies for a while..

Registering a domain name

- The root name server is maintained by the Internet Assigned Numbers Authority.
- There is a fixed set of top-level domains
- Responsibility for assigning subdomains of TLDs is delegated to
  - a small number of corporations (e.g., VeriSign) for .com, .net, etc.
  - national Internet authorities (for domains ending in .us, .ca,.uk,.in, etc.)
- You can register a domain name like *brilliantideas.com* for a small fee per year (currently about $20) *if it has not yet been assigned.*

SCHOOL OF INFORMATION
UNIVERSITY OF MICHIGAN

Transmission Control Protocol

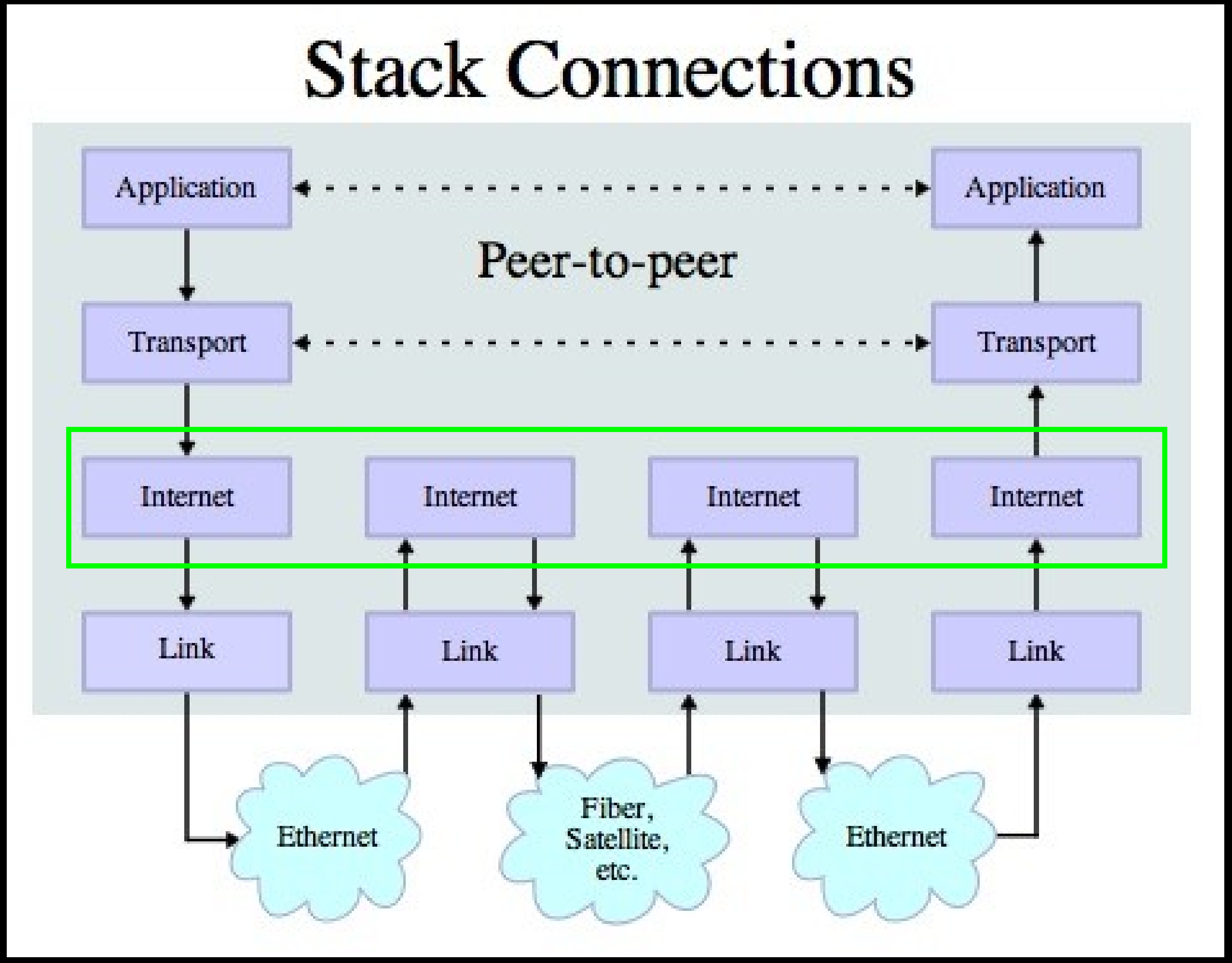
http://en.wikipedia.org/wiki/Transmission_Control_Protocol

Review: Magic IP

- What it does - Tries to get data a bunch of hops from one network to another network
- Keeps track of the good and bad paths for traffic - tries to pik better paths when possible
- But no guarantee of delivery - if things go bad - the data vanishes
- This makes it fast and big - and ultimately "reliable" because it does not try to do too much

Internet Protocol

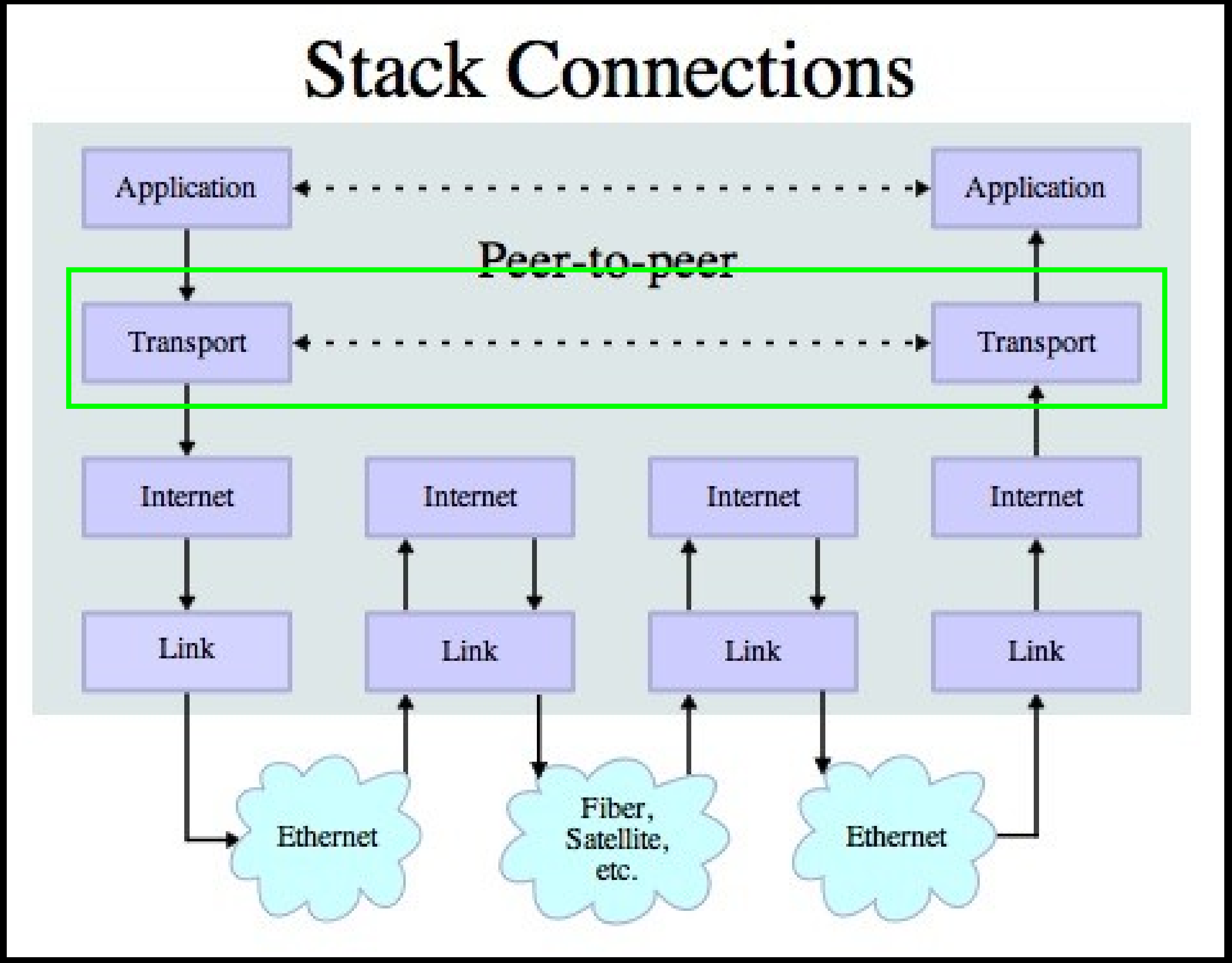
- So many links / hops
- So many routes
- Thinks can change dynamically and IP has to react (links up/down)
- IP can drop packets



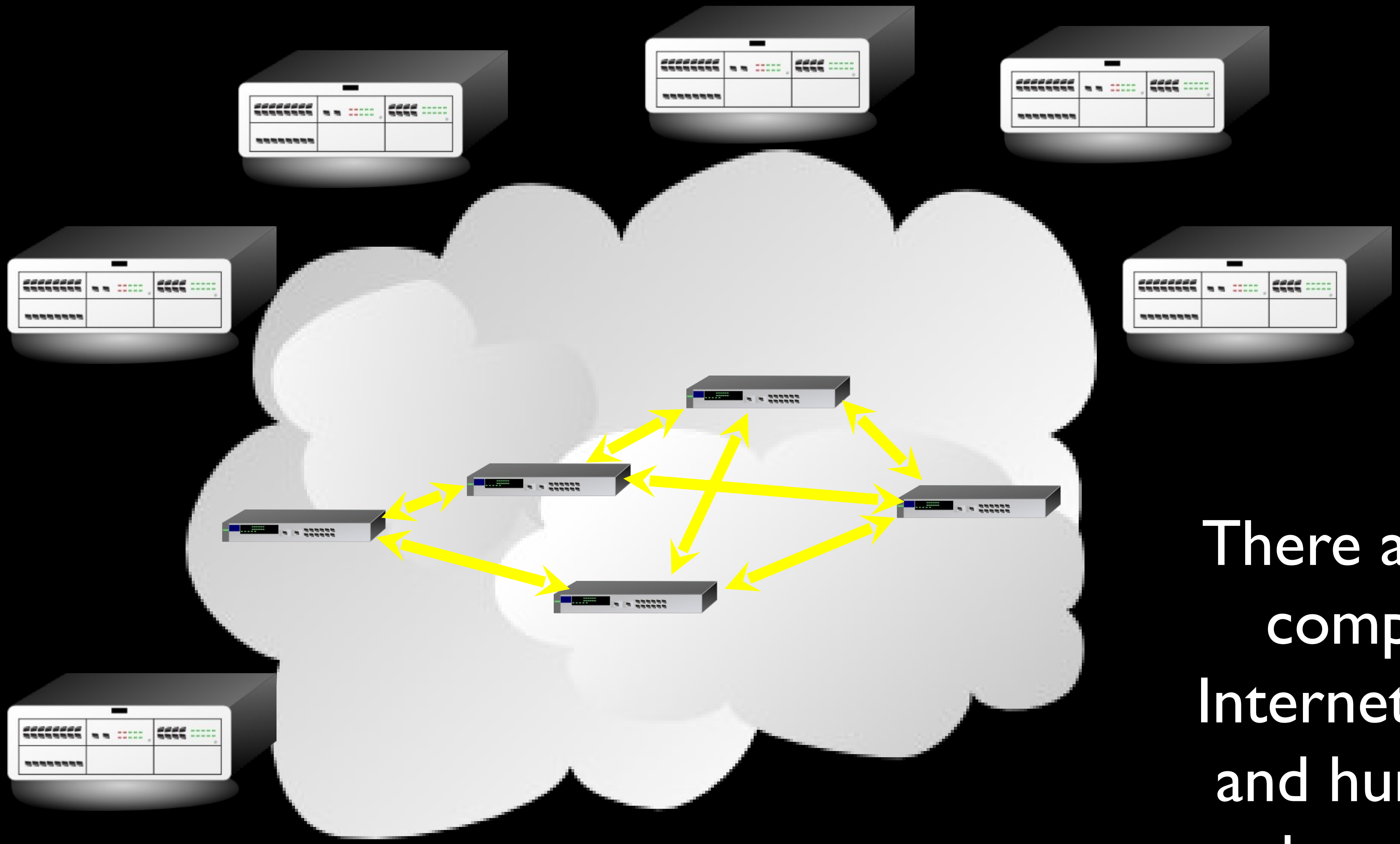
Source: http://en.wikipedia.org/wiki/Internet_Protocol_Suite

Transport Protocol (TCP)

- Built on top of IP
- Assumes IP might lose some data
- In case data gets lost - we keep a copy of the data a we send until we get an acknowledgement
- If it takes "too long" - just send it again



Source: http://en.wikipedia.org/wiki/Internet_Protocol_Suite



There are hundreds of millions of computers connected to the Internet and thousands of routers and hundreds of billions bytes of data enroute at any moment.

Who should store it in case it is lost? Routers? Computers?

Clipart: http://www.clker.com/search/networksym/1

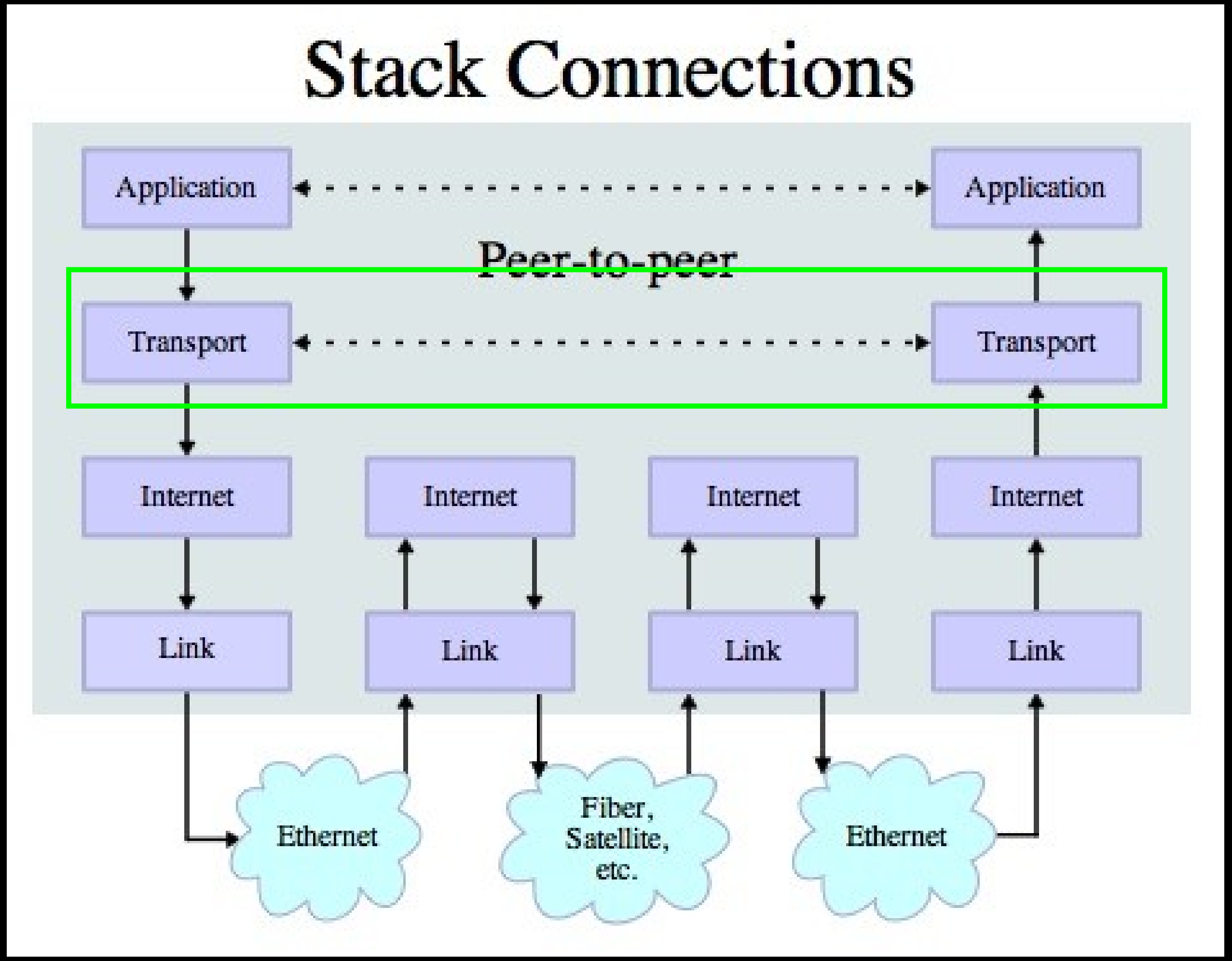
Transport Protocol (TCP)

- The responsibility of the transport layer is to present a reliable end-to-end pipe to the application
- Data either arrives in the proper order or the connection is closed
- TCP keeps buffers in the sending and destination system to keep data which has arrived out of order or to retransmit if necessary
- TCP provides individual connections between applications

Application Layer

Transport Protocol (TCP)

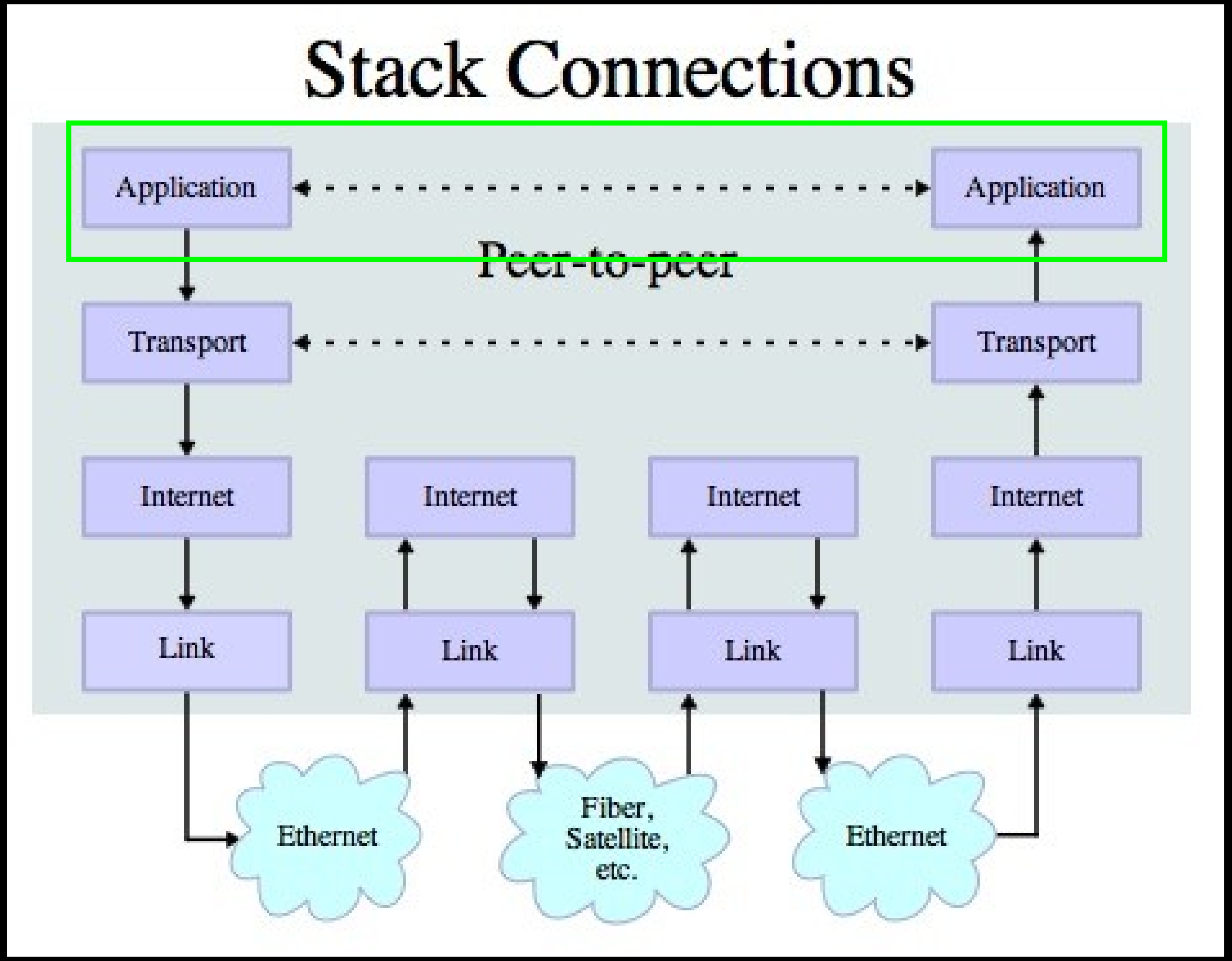
- Built on top of IP
- Assumes IP might lose some data - stores and retransmits data if it seems to be lost
- Handles "flow control" using a transmit window
- Provides a nice reliable pipe



Source: http://en.wikipedia.org/wiki/Internet_Protocol_Suite

Application Protocol

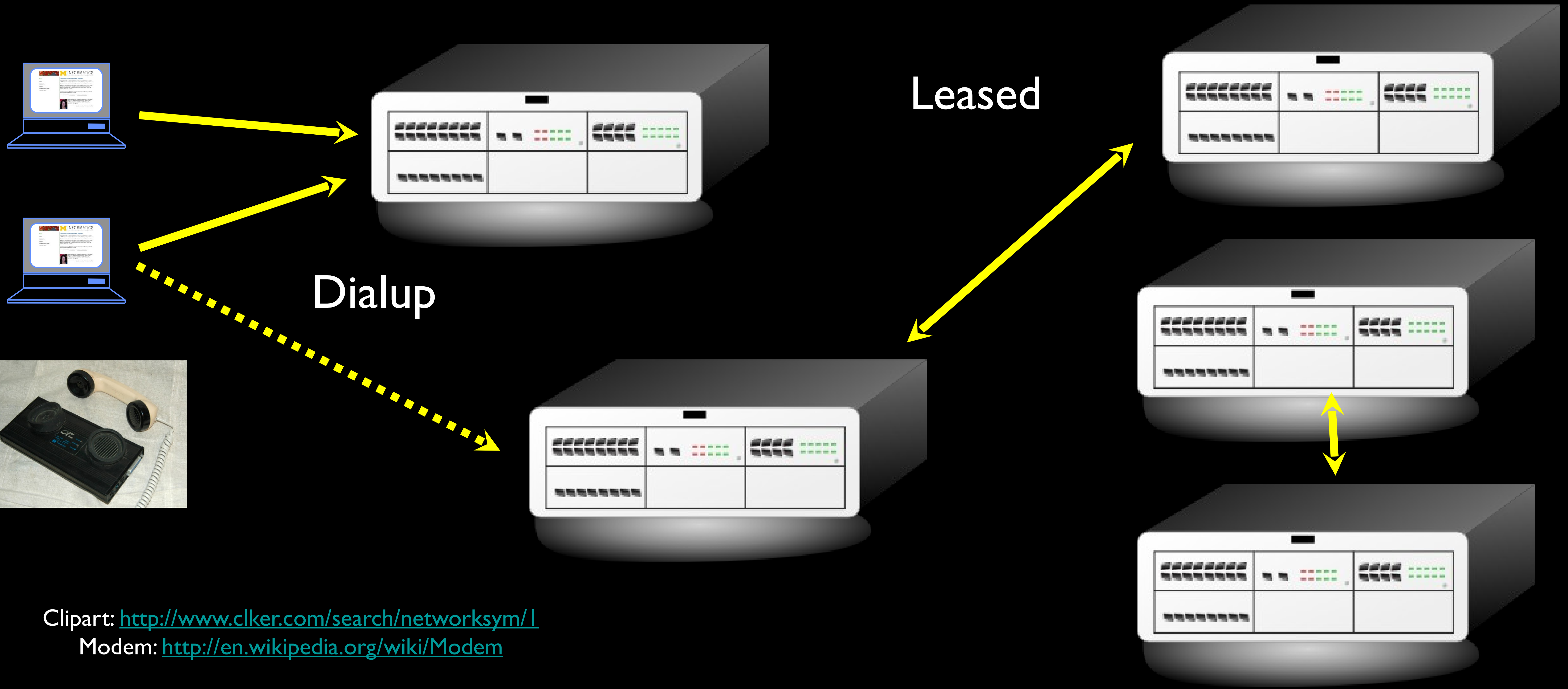
- Since TCP gives us a reliable pipe, what to we want to do with the pipe? What problem do we want to solve?
- Application Protocols
  - Mail
  - World Wide Web



Source: http://en.wikipedia.org/wiki/Internet_Protocol_Suite

Summary

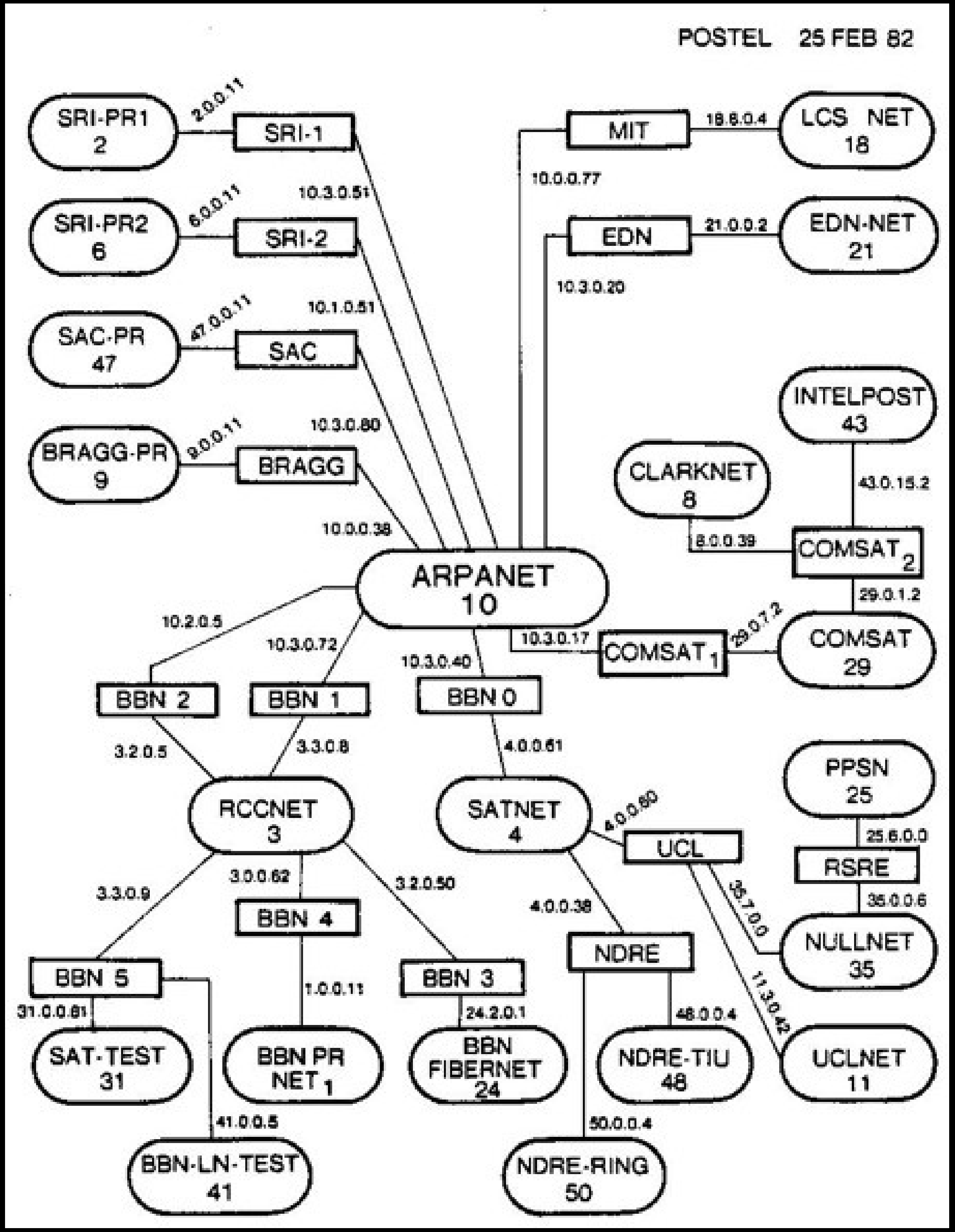
Phone Line Networking



Clipart: http://www.clker.com/search/networksym/1
Modem: http://en.wikipedia.org/wiki/Modem

Research Networks 1960-1980's

- How can we avoid having a direct connection between all pairs of computers? How can we share expensive long distance leased phone lines?
- The question was how can many computers share a common network?
- How can we dynamically handle outages?



Source: Unknown

NSF Net

- NSFNet was the first network that was “inclusive”
- Standardized on TCP/IP
- Initially the goal was all research universities
- In the early 1990’s commercial companies (Internet Service Providers) could join and resell service

Layered Network Model

- A layered approach allows the problem of implementing a network to be broken into more manageable sub problems
- For example the IP layer is allowed to lose a packet if things go bad
- It is TCP's Responsibility to store and retransmit data.

Application Layer
Web, E-Mail, File Transfer

Transport Layer (TCP)
Reliable Connections

Internetwork Layer (IP)
Simple, Unreliable

Link Layer (Ethernet/WiFi)
Physical Connections

An Amazing Design

- Hundreds of millions of computers
- Thousands of routers inside the Internet
- Hundreds of millions of simultanrous connections
- Trillions of bytes of data moved per second around the world
- And it works

The Internet

- It is said that "The Internet is the largest single engineering effort ever created by mankind"
- It was created to work in an organic way - to repair itself and automatically adjust when parts fail
- No one part of the Internet knows all of the Internet (like life)
- It is never 100% up - but it seems up all the time