Basic Network Organization

Portions Courtesy Ellen Liu

Outline

- Internet and Internet Standards
- Protocols and Protocol Layering
  - Packet-switching
  - Segment, packet, frame
  - TCP segment header and IP packet header
- Addressing in networks
- The IP protocol

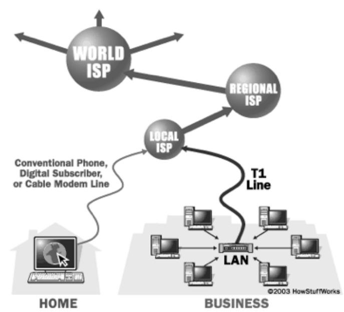
Internet

- Internet started as a research network called ARPANET in 1969. It became commercial in late 1980s
- Today's Internet is a collection of networks owned by various levels of ISPs (Internet service providers)
- It has now evolved into a public utility

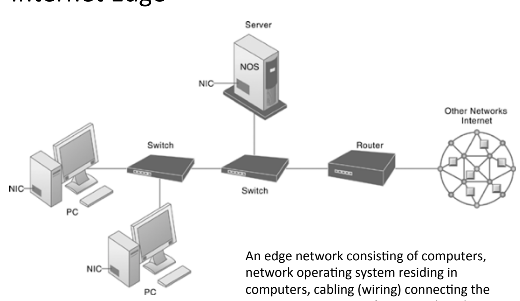
A map of the Internet: http://www.opte.org/maps/

Internet

- Backbone (tier1) ISPs
  International coverage. Are equals
- Regional (tier2) ISPs
  connect to 1 or more tier1 ISPs
- Local ISPs closest to end users



Internet Edge



An edge network consisting of computers, network operating system residing in computers, cabling (wiring) connecting the devices, network interface cards (NICs), switches, and a router

Internet Governance

No formal management. Policies established by professional and government organizations

- ICANN: Internet Corporation for Assigned Names and Numbers. Allocation of IP addresses, domain names, protocol port numbers, autonomous system numbers
- ISOC: Internet Society. Overlooking technical development
  - IETF: Internet Engineering Task Force. Produces Internet standards
  - IAB: Internet Architecture Board. Directly oversees IETF's work

Internet Standards and Documents

- RFC (request for comment) - a memorandum published by IETF describing methods, behaviors, research, or innovations applicable to the working of the Internet and Internet-connected systems
- There are 6921 RFCs as of today. See the rfc index at http://www.ietf.org/download/rfc-index.txt
- RFCs started as Internet Drafts. Each went through an intensive review process.
- There are many IETF working groups. Each is busy on the Internet drafts under the group charter. Everybody who is interested can join these groups and get involved.

Stony Brook University CSE/ISE 311: Systems Administration

RFCs

- Not all RFCs are standards. RFC status include:

  proposed standard (STD), informational, experimental, best current practice (BCP), historic, unknown
- Once an RFC is distributed, its contents never change
- Updates can extend, clarify, or supersede old RFCs, are distributed with a new RFC number
- RFC2026 and RFC5540 describe this process

Stony Brook University CSE/ISE 311: Systems Administration

Protocols

- Protocols: define type, format, and order of messages sent and received among network entities, and actions taken on message transmission and receipt, or other events (e.g., timer expires)
- Computers talk to each other in a way that is really not much different from how we humans talk to each other
  - Hi. Hi. Got the time? 5 o'clock. Thanks. Bye. Bye.
  - Connection request. Connection response. Got page index.html? Here you are. ACK. Connection teardown. Connection teardown.

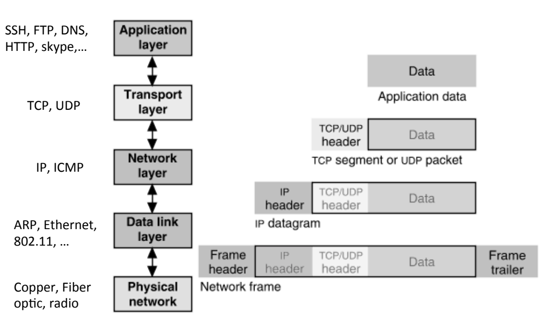
Stony Brook University CSE/ISE 311: Systems Administration

Protocol Layering

- Network protocols are arranged in a hierarchy or stack, with higher-level ones making use of the ones beneath them
- Five protocol layers in Internet Protocol Stack: Application, Transport, Network, Data Link, Physical layers
- There are other networking protocol stacks. E.g., ISO OSI 7-layer model, ATM, X.25, SNA
  - Not as widely used as Internet Protocol stack

Stony Brook University CSE/ISE 311: Systems Administration

Internet Protocol Layering



Stony Brook University CSE/ISE 311: Systems Administration

Physical Layer

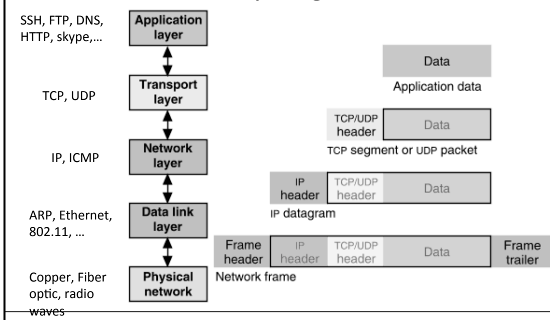
- Physical layer is the specification of low-level electrical signals (or waves or light beams) used to encode a message
  - Generally encapsulated from the administrator
    - Although there are some limits on signal length over a medium

Packets

- The basic unit of data transmission
- Most media specify a Maximum Transmission Unit (MTU)
  - Packets cannot be larger than the MTU
  - Higher-level protocol messages may have to be split across multiple packets

Stony Brook University CSE/ISE 311: Systems Administration

Internet Protocol Layering



Stony Brook University CSE/ISE 311: Systems Administration

Data Link Layer

- Software-level abstraction of the physical layer + some higher-level protocols
- Ethernet (most common wired network)
  - Older wired protocols include Token Ring
- Wireless (802.11)
  - 801.11a, b, g, n, etc all specify different radio wave specifications

Stony Brook University CSE/ISE 311: Systems Administration

MAC Address

- Unique identifier for a network device
  - E.g., 0e:d1:c3:db:e7:b3
  - First few bytes encode manufacturer and model
  - Others are supposed to be unique
- Used at the Data Link Layer to specify the destination for a message (packet)
- Note: Many NICs allow you to change the MAC address

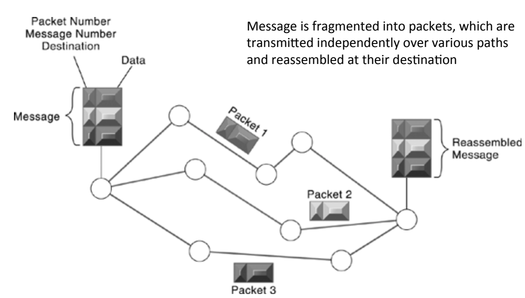
Stony Brook University CSE/ISE 311: Systems Administration

Packet Switching

- **Packets**: method of slicing digital messages into parcels. Each packet contains a *header* and *payload. The payload carries a parcel of message*
- As packets become available, they are sent along paths between a sender/receiver pair, then reassembled at the destination (see next slide)
- Store-and-forward: entire packet must arrive at a router before it can be transmitted onto next link
  - This introduces L/R seconds delay. L: packet size, R: link capacity (also called bandwidth, transmission rate)

Stony Brook University CSE/ISE 311: Systems Administration

Fragmentation and Reassembly



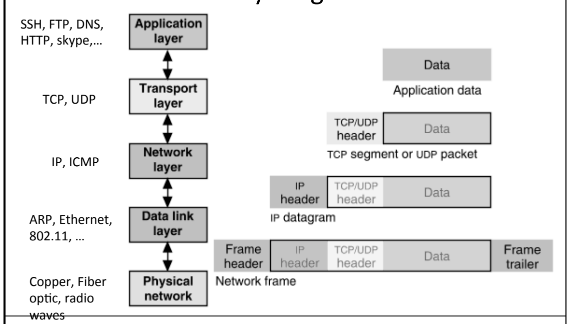
Segment, Packet, and Frame

- The name of the primitive data unit depends on the layer of the protocol in question
  - At the network layer, it is called a packet or datagram
  - At the transport layer above, it is called a segment
  - At the data link layer below, it is called a frame
- As the unit travels down the protocol stack in preparation for being sent, each protocol adds some header for doing its job
  - Thus, e.g., a packet is a segment plus a packet header, i.e., the segment becomes the packet payload

IP (Internet Protocol)

- Most common Network Layer protocol
- Routing packets from source machine to destination machines
  - Across networks (i.e., the Internet)
  - Data Link Layer is sufficient within a local network
    - E.g., among computers connected via a single wireless access point

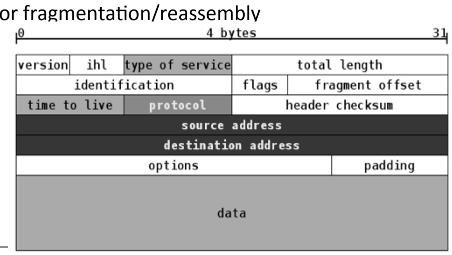
Internet Protocol Layering



IP Packet Header

- E.g., source, destination address: IP addresses of the source and destination; version: v4 or v6; entire second row for fragmentation/reassembly

IPv4 packet header (20 bytes)



IP Addresses

- IP address: also called network address. Used by software called the TCP/IP stack. Per *network interface*. One machine can have multiple of them
  - *NIC*: network interface card
- Are 4 bytes (32 bits) long for IPv4, and16 bytes (128 bits) long for IPv6. All modern OS and devices support both
  - IPv6 has built-in security/authentication, it addresses IPv4 address space shortage
  - Will focus on IPv4 here

IP Address Examples

- IPv4: 130.245.65.129
  - 4 8-bit values, each separated by a dot
  - If any number is >=(2^8=256), it is wrong
    - Here's looking at you, CSI
- IPv6: fe80::7ed1:c3ff:fedb:e7b3
  - One hexadecimal digit encodes 4 bits
  - 8 x 4 hex digits = 128 bits
  - A string of consecutive 0's in the middle replaced with double colons (::)

How many IPv4 Addresses are there?

- 2^{32}= ~4.3 billion
- How many computers in the world?
  - ~2 billion on the internet in 2010
- How many people in the world?
  - ~7 billion
- So IPv4 will eventually run out
  - And management issues have caused problems already

Stony Brook University CSE/ISE 311: Systems Administration

The IP Protocol

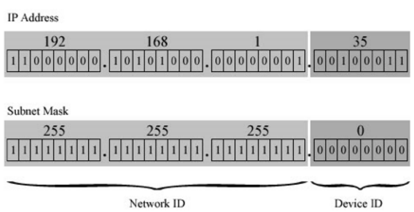
- Major job: routing packets from source machine to destination machines. Actually two tasks
- Forwarding vs. routing
  - Forwarding: move a packet from a router's input to an output
  - Routing: determine a route taken by a packet from source to destination
- Routing is done in the background. It produces IP forwarding tables
  - Will focus on forwarding here

Stony Brook University CSE/ISE 311: Systems Administration

Another way to look at IP Addresses

Consists of a network portion and a host portion

- Network portion: high order bits, identifies a logical network
- Host portion: the rest bits, identifies a node on the network



Stony Brook University CSE/ISE 311: Systems Administration

Packet Delivery (within network)

- You can figure out if an IP is in the same network by looking at the Network ID portion of the address
- Use a Data Link Layer protocol called Address Resolution Protocol (ARP) to ask:
  - "Does anyone know the MAC address of IP x.x.x.x?"
  - Cache results in a local table
  - Usually only ask once

Stony Brook University CSE/ISE 311: Systems Administration

Local Packet Delivery

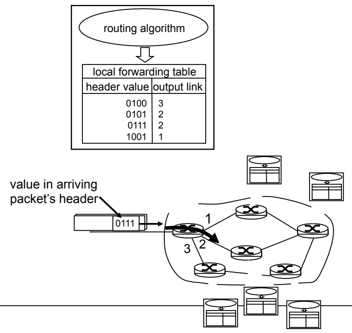
- Common case: look up ARP cache, send directly from one computer to another

Stony Brook University CSE/ISE 311: Systems Administration

Remote Packet Delivery

- Computers also include **routing tables** that map network names onto remote IP addresses
- Within a simple network (like campus), your routing table may simply send all remote packets to the edge router

Interplay between routing, forwarding



Subnet

- A subnet is made of those network interfaces that can reach each other without passing through a router. All of them have the same subnet (network) portion
- **Subnet mask:** specifies the length of the network portion. The 1's must be leftmost and contiguous
  - E.g., /24 or 255.255.255.0
  - What's equivalent of /26? Answer: 255.255.255.192
- How many hosts can be in a /26 network/subnet? Answer: 62
  - Host portion of all 0's denotes this subnet, all 1's is used as a multicast address. They cannot be assigned to hosts

Historical network types

- Class A: only first byte used for network address
  - Huge, hard to get
- Class B: second two bytes for network
  - Still pretty big, easier to get
- Class C: First three bytes identify network
  - Easy to get – Even I have one for my lab
- Netmasks give you finer-grained subdivision

How many computers in a Class C?

- Hint: 1 byte for host portion
- 254
- Why not 256?
  - .0 reserved for router
  - .255 reserved for broadcast

CIDR

- Classless Inter-Domain Routing. A method to allocate IP addresses and routing IP packets. Allows arbitrary length of the network portion
  - The previous classful addressing uses fixed length
    - Class A: 8 bits in network portion
    - Class B: 16 bits in network portion
    - Class C: 24 bits in network portion
- CIDR notation: 192.144.0.0/21
- Assume a site is given the block 192.144.0.0/21. The site could use the block in various ways. For example:

Supernetting

- 1 network /21 with 2,046 hosts
- 8 networks /24 with 254 hosts each
- 16 networks /25 with 126 hosts each
- 32 networks /26 with 62 hosts each

- Q: how many routing table entries for each case?
- A:
  - From the perspective of Internet, no need to have 8, 16, 32 entries
  - All refer to the same organization, go to the same ISP
  - A single entry 192.144.0.0/21 suffices.
- Supernetting aggregates several networks for purposes of routing

IP Forwarding Table

Destination IP address range	Link Interface
11001000 00010111 00010000 00000000 through *11001000 00010111 00010111 11111111*	*0*
11001000 00010111 00011000 00000000 through *11001000 00010111 00011000 11111111*	*1*
Otherwise	*2*

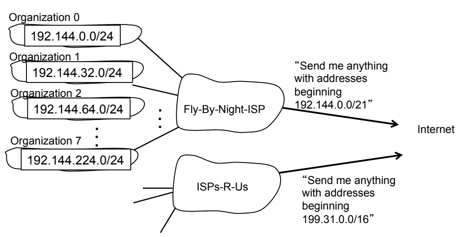
200.23.16.0/21	0
200.23.24.0/24	1
Otherwise	2

Longest Prefix Match

- In the previous example, assume the site has 8 /24 networks with 254 hosts each
  - We only need one entry to route to these 8 networks
- If one of the 8 networks moves to a new ISP, can add a more specific entry 192.144.32.0/24
  - In addition to the single 192.144.0.0/21 entry
- Longest prefix match. When both entries apply, use the one with longest prefix
  - Thus /24 is used. This way the packet is routed correctly to the new ISP

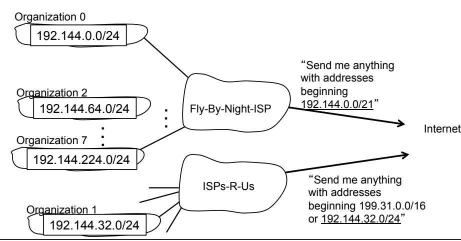
Hierarchical addressing: route aggregation

hierarchical addressing allows efficient advertisement of routing information:

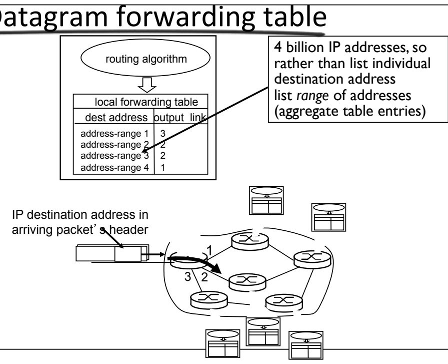


Hierarchical addressing: more specific routes

ISPs-R-Us has a more specific route to Organization 1



Datagram forwarding table



Longest prefix matching

longest prefix matching

when looking for forwarding table entry for given destination address, use *longest* address prefix that matches destination address.

Destination Address Range	Link interface
11001000 00010111 00010*** *********	0
11001000 00010111 00011000 *********	1
11001000 00010111 00011*** *********	2
otherwise	3

examples:

DA: 11001000 00010111 00010110 10100001 which interface?

DA: 11001000 00010111 00011000 10101010 which interface?

Human-Understandable Addressing

Hostname and Ports

- Hostname: IP addresses are hard to remember. Thus we name our machines
  - Hostnames generally managed by a transport layer protocol called Domain Name Service (DNS)
- DNS maps human-readable names to IP addresses

$ host www.cs.stonybrook.edu
www.cs.stonybrook.edu is an alias for www.cs.sunysb.edu.
www.cs.sunysb.edu has address 130.245.27.2

Human-Understandable Addressing

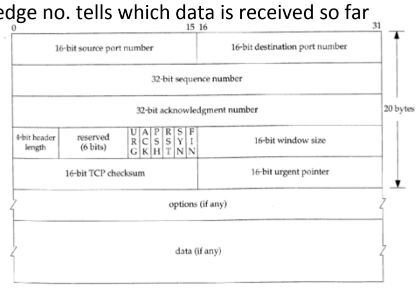
Hostname and Ports

- Ports: an IP address leads packets to a machine. A port number leads packets to a process or service. 80 for Web, 25 for email, 20/21 for ftp …
  - Ports are a TCP/UDP abstraction for a specific application/protocol
  - Where are ports in packets? They are in the transport segment header.
  - 16 bits to encode ports. How many ports overall possible? 2^16. Some are well-known. Some not. See http://iana.org/assignments/port-numbers

TCP Segment Header

- E.g., sequence no. tells which data is in payload; acknowledge no. tells which data is received so far

Minimum
(20 bytes)



What does each protocol do?

- HTTP: requests and serves web pages
- FTP: requests and serves files
- TCP: reliable, full-duplex, flow-controlled, error-corrected conversations
- UDP: unverified, one-way data delivery
- IP: routing packets from source machine to destination machines
- Ethernet: communication between adjacent nodes

To do these jobs, headers are added

TCP/IP Protocols

A suite of protocols. Each defined by one or more RFCs. Some major ones:

- IP: Internet Protocol. RFC791
- ICMP: Internet Control Message Protocol. RFC792
- ARP: Address Resolution Protocol. RFC826
- UDP: User Datagram Protocol. RFC768
- TCP: Transmission Control Protocol. RFC793