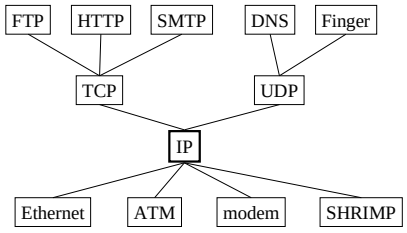
The Network Layer



Internetworking

- challenge: how to tie together diverse networks into a single meta-network
- goes beyond bridges
  - spans heterogeneous networks
  - must scale to millions of hosts
- terminology
  - internetwork: any combination of networks
  - Internet: a particular internetwork we all use

Internetworking

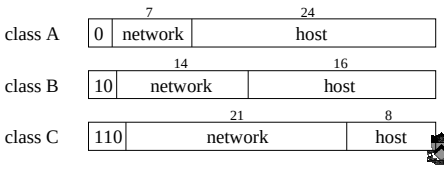
- three facets:
  - addressing: how hosts are named
  - service model: how hosts interact with the network
  - routing: how a route from sender to destination is chosen
- several versions of IP exist
  - everybody uses version 4 now
  - version 6 is coming

Who Runs the Internet?

- nobody!
- the Internet Engineering Task Force (IETF) declares technical standards, and then everybody follows them
- the IETF negotiates a standard with the major players, and people do what they want, influenced by the IETF standard
- who elects the IETF?

Addressing

- unique 32-bit address for each host
- two-level address: (network, host)
- dotted-decimal notation (128.112.153.27)
- three address formats



Two-Level Addressing

- network is a single (extended) LAN
  - provides physical address for each host
  - can deliver packets between all hosts in network
- outside hosts treat a network as a black box
- advantages of two-level approach
  - can remember single route for a whole network
  - host numbers can be allocated locally

Three-Level Addressing

- (network, subnet, host)
  - subnet and host numbers managed locally
  - outside can't tell two-level from three-level
- variable division of address bits between subnet and host parts
  - determined by *subnet mask*
- subnet is a single LAN
- subnets treat each other as black boxes

Translating Addresses

- IP uses IP addresses, but a network uses physical addresses
  - must translate to communicate within network
- each host keeps a list of IP-to-physical translations
  - forgotten every 15 minutes or so
- Address Resolution Protocol (ARP) translates IP address to physical address

ARP

- querying host broadcasts a packet asking for a translation for some IP address
- host with that IP address answers with ARP response containing the physical address
  - all hosts know their own addresses
- query contains translation for the querying host
  - other hosts learn from the query

RARP

- reverse Address Resolution Protocol (RARP) translates physical to IP
- works just like ARP
- allows a host to learn its own IP address after it boots
  - table of address mappings stored on a server
    - » maintained by administrator

DHCP

- Dynamic Host Configuration Protocol (DHCP) allows host to learn its own network parameters
- on startup, host broadcasts DHCP query
- server responds with all the information host needs to function on the Internet
- used on many PC networks

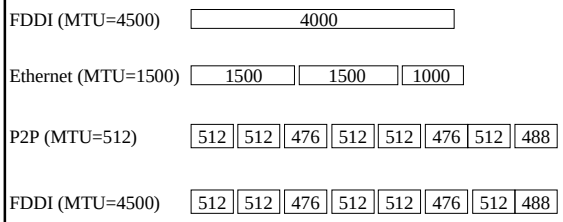
Service Model

- unreliable datagram service
  - transmits data in variable size chunks (datagrams)
  - may drop, reorder, or duplicate datagrams
  - best effort approach

Fragmentation and Reassembly

- each network has a Maximum Transmission Unit (MTU): the largest IP packet it can carry
- packet may travel through diverse networks with different MTUs
- solution: if packet is bigger than MTU, break it into fragments
  - may be broken more finely in a later network
- reassemble at ultimate destination

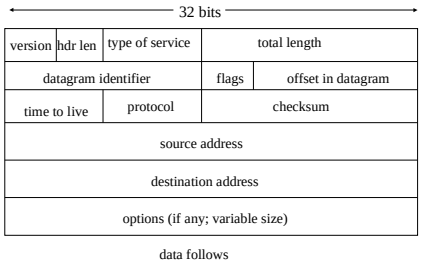
Fragmentation Example



Reassembly

- performed by destination host
- store fragments in memory until they all show up
- fragments might be dropped, so discard stored fragments after a timeout period
- for best performance, avoid fragmentation
  - higher-level protocols can avoid sending large datagrams (use common MTU, or discover the right size dynamically)

IP Packet Format



IP Packet Format: Details

- version: currently 4
- header length: in 32-bit words, 5 if no options
- type of service: almost never used
- total length: in bytes, header+data
- datagram identifier: allows destination to match up fragments of the same datagram

IP Packet Format: Details

- flags
  - more-fragments: says this isn't the last fragment of the datagram
  - don't-fragment: prohibits fragmentation; packet will be dropped rather than fragmented
- offset: offset within datagram at which this fragment begins

IP Packet Format: Details

- time to live: initially set to 64; decremented on each hop; packet dropped if TTL==0
- protocol: identifies which higher-level protocol this datagram belongs to
- checksum: 16-bit ones-complement sum
- source address, destination address obvious
- options: rarely used

ICMP

- Internet Control Message Protocol
- companion protocol to IP
- used to report errors in IP
- types of packets
  - destination host unreachable
  - reassembly failed
  - checksum didn't match
  - don't send so fast
  - there's a better route

Host Names and DNS

- Domain Naming System (DNS) provides a mapping from human-readable names (idea.cs.princeton.edu) to IP addresses (128.112.153.27)
- benefits
  - easier to remember
  - level of indirection allows reconfiguration without telling clients

Domains

- hierarchy of domains
  - top-level: .edu .gov .uk .su and so on
  - deeper levels: princeton.edu cs.princeton.edu
- name-servers provide mappings for a domain
  - at least two name servers per domain
  - servers for a domain must know IP addresses of name-servers for all sub-domains

DNS Naming

- top-level servers know IP addresses of name-servers for top-level domains
- lookup process
  - ask top-level server for IP address for desired name
  - result is either
    - » the answer
    - » referral to name-server for a more specific domain
  - iterate until you have the answer

DNS Details

- most machines just ask a local DNS server, which carries out the lookup process
- cache results
  - forget after timeout period
  - also cache lookup failures
- locations of top-level servers hard-wired into local DNS servers
- location of local DNS server hard-wired into each machine

Administrative Details

- to allocate a DNS domain name
  - contact InterNIC (Network Information Center)
  - pay them a fee
  - they check that you have the necessary servers running
  - they tell top-level servers that you're there
- similar for IP network number allocation