Introduction To TCP/IP
and
The Internet
Part II

DECUS Europe, Fall 1994
Friday, September 16 1994

Vladimir Sukonnik
Process Software Corporation
sukonnik@process.com

PART II

Internet Model and The Basics of Internetworking

Our Goal

To maximize interoperability of distributed computations across vendors and products

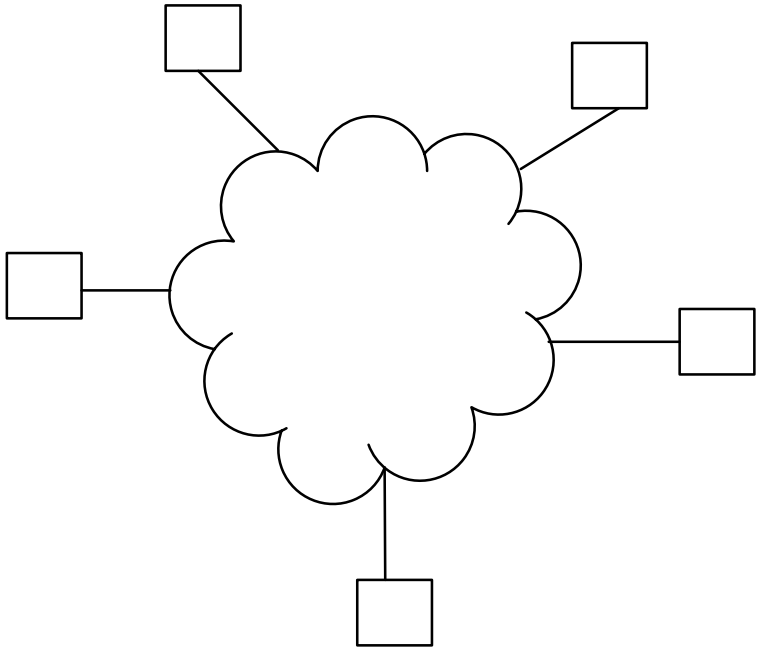
To Achieve Our Goal We Need

- Universal connectivity
- Agreement on procedures and formats for exchanging information
- Agreement on names and addresses for services, machines, etc.

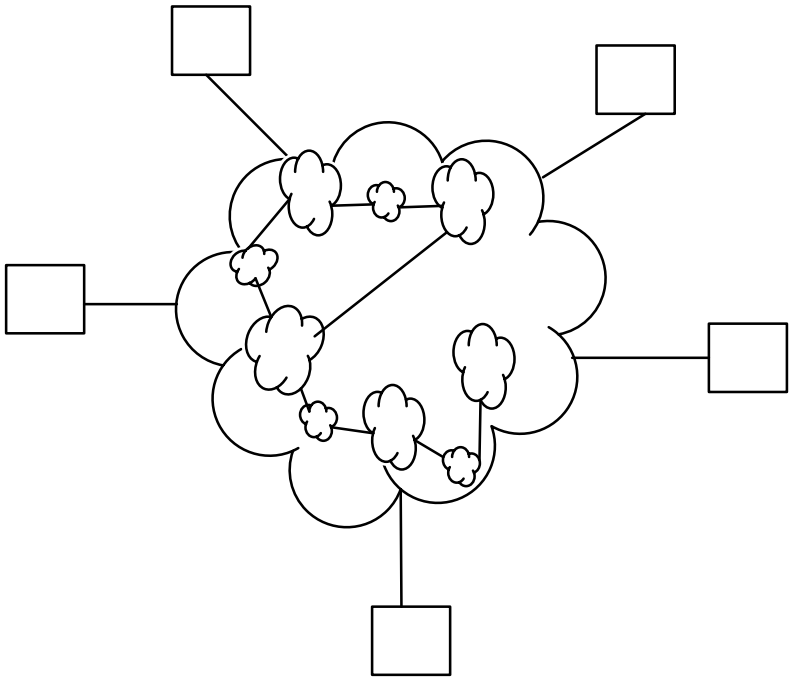
The Internet Idea

- Use standard network hardware
- Interconnect physical networks
- Devise abstraction to hide details of network hardware, physical addresses, and routing

Conceptual View Of The Internet



user's view



actual connections

Architectural Model

- Hosts connect to physical networks
- Dedicated computers called gateways or routers interconnect networks

Addressing Model

- Abstract Internet addresses (IP addresses) assigned to every host or gateway that connects to the Internet
- Application software on hosts and gateways use Internet addresses when sending and delivering packets
- Software in the operating system maps Internet address into physical hardware address automatically

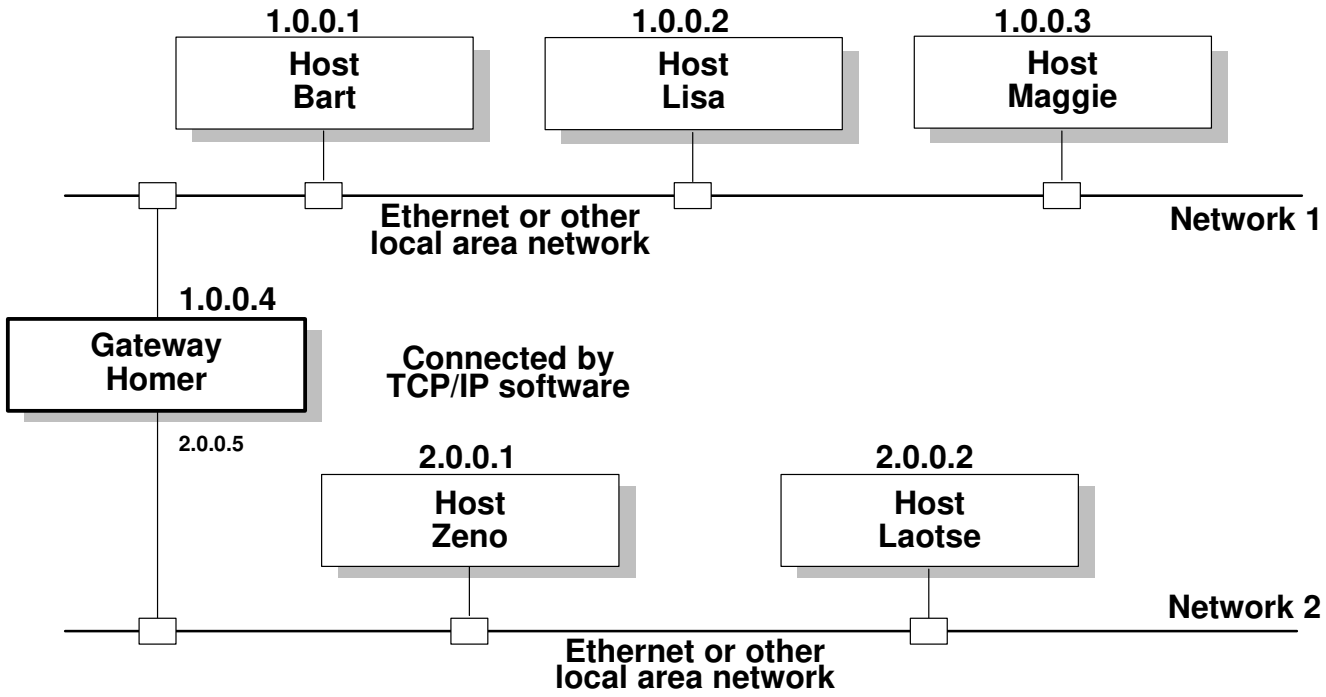
How Hosts Transmit Packets On The Internet

- Host forms a packet to be sent including the destination Internet address
- If host can reach the destination directly, it sends the packet over the physical network to its destination
- If host cannot reach the destination directly, it sends the packet to the nearest gateway

How Gateways Handle Packets

- If gateway can reach the destination directly, it send the packet over the physical network to its destination
- If gateway cannot reach the destination directly, it chooses another gateway that it can reach directly and sends the packet to that gateway

Illustration of Gateway Routing



Desirable Properties Of An Internet Addressing Scheme

- Compact
- Universal
- Work with all network hardware
- Support efficient decision making
  - Test whether a destination can be reached directly
  - Decide which gateway to use for indirect delivery
  - Choose next gateway along a path to the destination

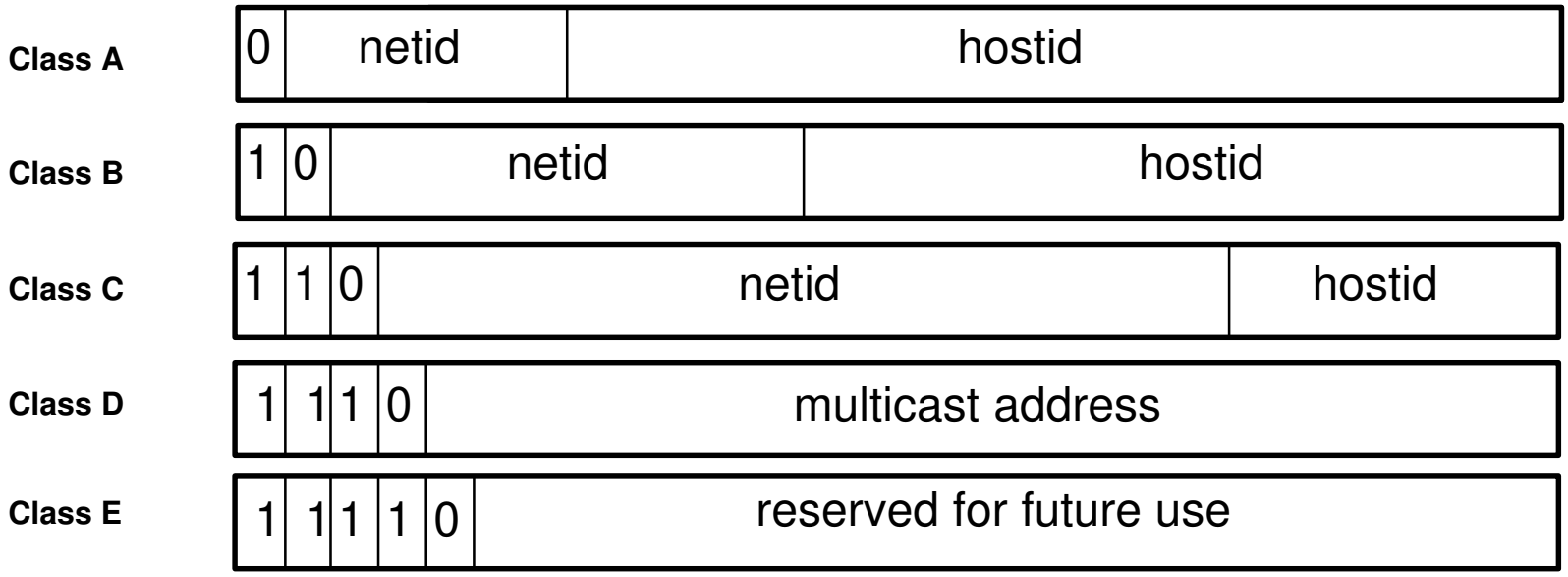
The TCP/IP Internet Address Scheme

- Each host assigned a 32–bit, hardware independent address (IP address)
- Prefix of the address uniquely identifies the network to which the host attaches (all machines on the same net share the same prefix)
- Prefixes assigned by central authority (SRI–NIC)
- Suffix of the address gives a unique id for the host on the network (no two machines on a given network share the same suffix)
- Suffixes assigned by local network administrator

Division of Internet Address Into Prefix and Suffix

- How should division be made ?
  - — Large prefix, small suffix means many possible networks, but each is limited in size
  - — Large suffix, small prefix means each network can be large, but there can only be a few networks
- Internet address scheme accommodates both possibilities

The Three Classes of Internet Addresses



	Total	Allocated	Allocated (%)
Class A	126	49	38%
Class B	1638	7354	45%
Class C	2097151	44014	2%

Why Partition Internet Addresses Into Network And Host Portions ?

- Each address must be unique
- Want to make routing efficient
- Partition allows us to assign unique id to each network and do routing using network portion and delivery using unique host portion

Dotted Decimal Notation

- Syntactic form for expressing 32–bit address
- Used throughout the Internet and associated literature
- Represents each byte in decimal separated by periods (dots)

Example of Dotted Decimal Notation

- Here is a 32–bit number in binary

  10000000 00001010 00000010 00000011

- Here is the same 32–bit number expressed in dotted decimal notation

  128.10.2.3

Subnet Addressing

- Not part of original Internet scheme
- Allows a site to use a single network address for multiple physical networks
- Subdivides the *hostid* portion of an address into a pair of fields for physical network and host
- Interpreted only by gateways and hosts at the site; treated like normal address elsewhere
- Like Internet address scheme in miniature

Example Of Subnet Addressing

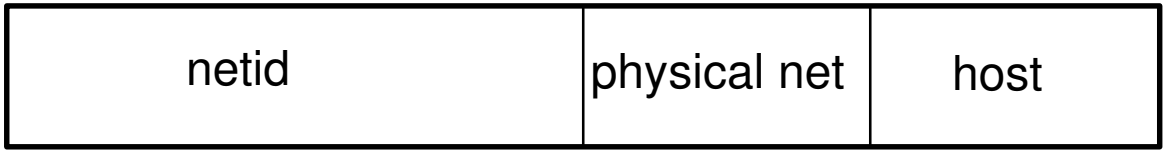
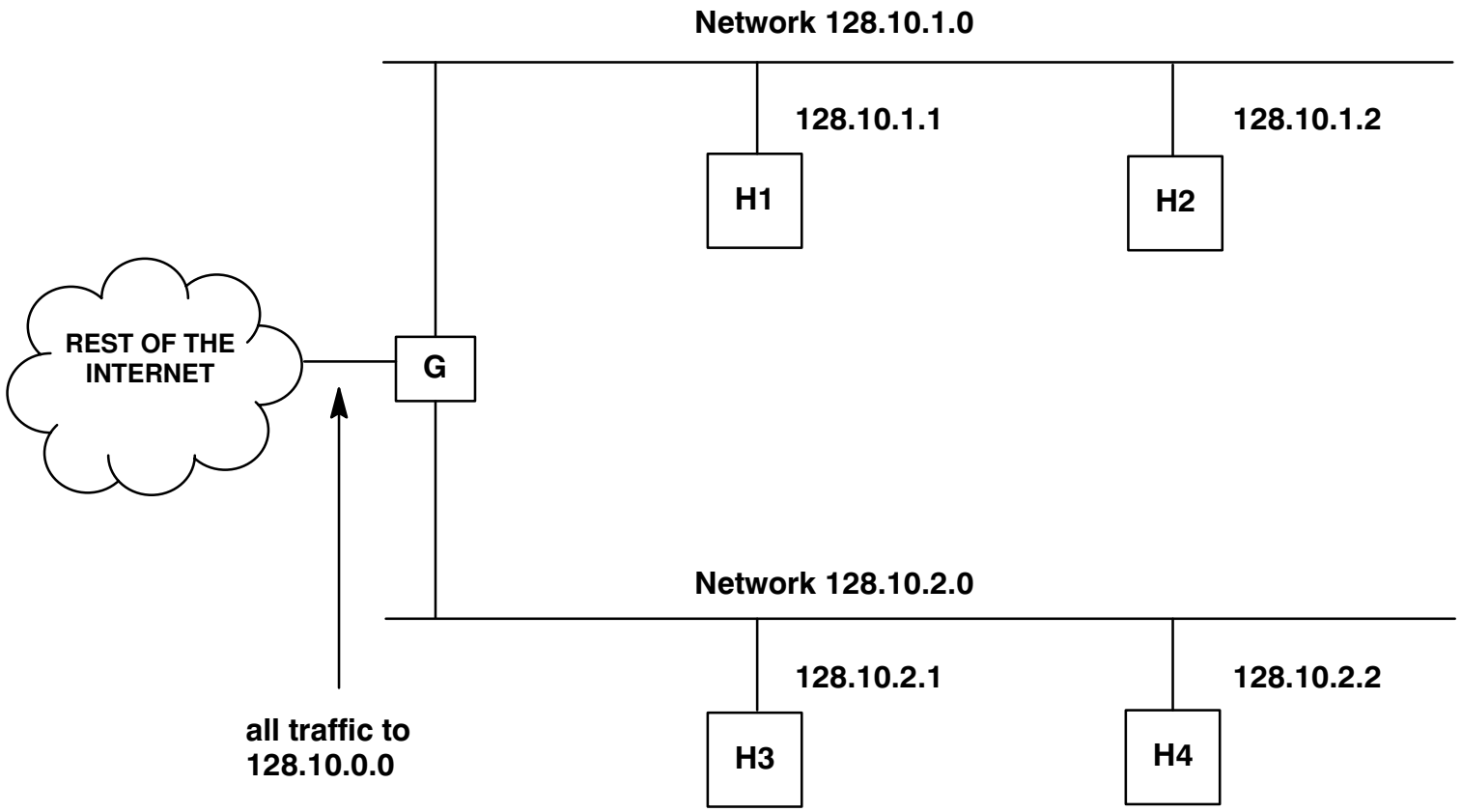


Illustration of Subnet Addressing



Subnet Address Details

- Each network is assigned subnet mask
- Mask covers *netid* portion plus zero or more bits of *hostid* portion
- Site may choose different masks for each of its networks (not advisable)
- Mask–based addressing can generalize to handle host–specific routes, network specific routes, and subnet routes

Binding Internet Addresses To Physical Hardware Addresses

- Must use physical addresses to send and receive packets
- Therefore, must map Internet addresses to physical addresses
- Method depends on underlying hardware technology
- Software to perform mapping built into network interface software in the operating system

Static Address Binding

- Useful when hardware addresses chosen from small, dense set
- Key ideas:
  - Local site administrator is free to choose host portion of an Internet address
  - So, encode hardware address in host portion of Internet address
  - Mapping becomes trivial

Example of Static Binding

- Hardware is proNET–10
- Network portion of Internet address fixed when Internet address assigned
- Hardware address, assigned by customer, are typically: 1, 2, 3, etc.
- Choose Internet address with low–order byte equal to hardware address
- To map destination Internet address to hardware address, extract the low–order byte

Dynamic Address Binding

- Useful when hardware addresses are large and fixed
- Requires hardware broadcast
- Best for local area network
- Key ideas:
  - Maintain a table of pairs (Ia,Ha) where Ia is Internet address and Ha is hardware address for machine A
  - Use the network to obtain new bindings

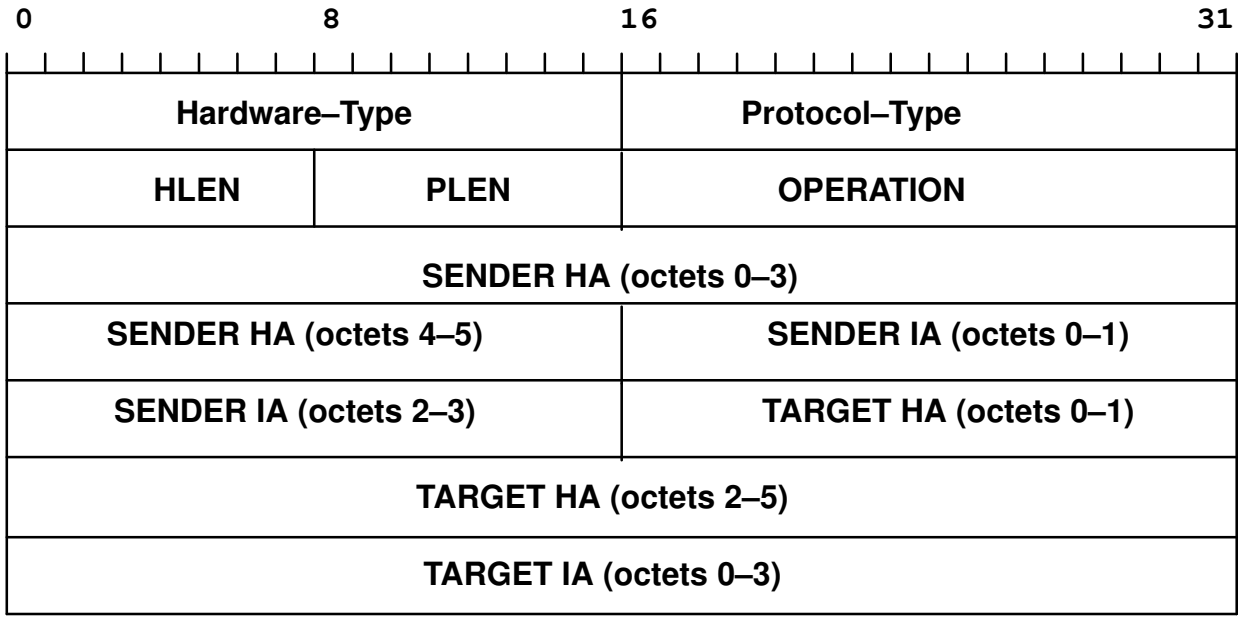
Internet Address Resolution Protocol (ARP)

- Internet standard for dynamic address binding
- Allows machine A to find machine B's physical address knowing only B's Internet address
- Uses hardware broadcast
- Note: ARP only used to map addresses within a single physical network, never across multiple networks

Address Resolution Protocol (continued)

- Machine A broadcasts ARP request with B's Internet address
- All machines on local net receive broadcast
- Machine B replies, with its physical address
- Machine A delivers packet directly to B

ARP Packet Format When Used With Ethernet



Algorithm For Processing ARP Requests

- Extract sender's pair (Ia,Ha) and update local ARP table
- If this is a request and the target is "me"
  - Exchange sender and target entries
  - Fill in target hardware address
  - Set option to *reply*
  - Send reply back to requester

ARP Details

- ARP table is merely a cache
- Entries should time out and be invalidated
- Machine can broadcast new binding when it boots

Comments on Address Binding

- Conceptual Purpose
  - — Isolates hardware addresses at low level
  - — Allow application programs (and other protocol software) to work with Internet addresses only
- Consequences
  - — Occurs at lowest level of system
  - — Should not be used for routing
  - — Provides uniform, virtual network interface

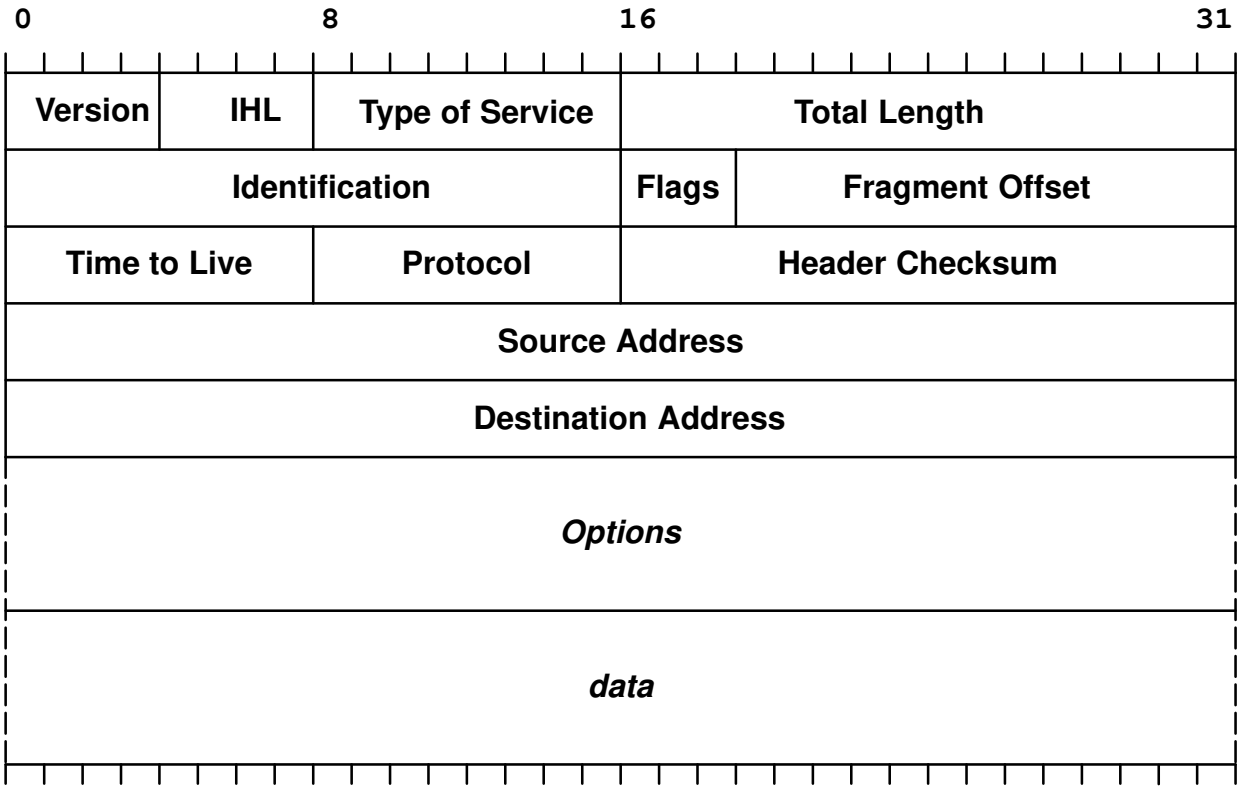
Internet Protocol (IP)

- Embodies the Internet connectionless packet delivery service
- Defines the format of packets for the (virtual) Internet
- Specifies Internet routing

IP Datagram

- Basic unit of Internet transfer
- Analogous to physical network packet
- Composed of
  - Header that contains source and destination Internet addresses, datagram type field, etc.
  - Data area that contains data being carried

IP Datagram Format



Encapsulation

- IP datagram travels in physical network packet or frame
- Complete datagram is treated as data by the hardware
- Internet defines standards for encapsulation on most network hardware technologies

Datagram Encapsulated In Physical Network Frame

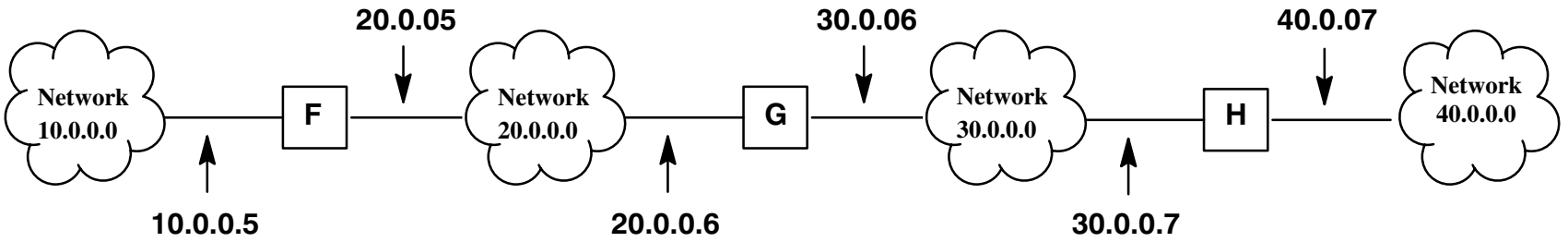
Routing In The Internet

- Host delivers datagrams to directly connected machines
- Host sends datagrams that cannot be delivered directly to gateways
- Gateways forward datagrams to other gateways
- Final gateway delivers datagram directly

Efficient Routing

- Routing decisions based on table lookup
- Routing tables keep only network portion of addresses (size proportional to number of networks, not number of hosts)
- Algorithm is efficient and easy to understand
- Easy to automate routing table update

Example IP Routing Table (Gateway)



An example Internet with IP addresses

TO REACH NETWORK	ROUTE TO THIS ADDRESS
20.0.0.0	DELIVER DIRECT
30.0.0.0	DELIVER DIRECT
10.0.0.0	20.0.0.5
40.0.0.0	30.0.0.7

The routing table for gateway G

Network MTU

- Each network hardware technology imposes a fixed limit on packet size
- Size limit call *Maximum Transmission Unit* (MTU)
- Encapsulated datagram must be less than network MTU

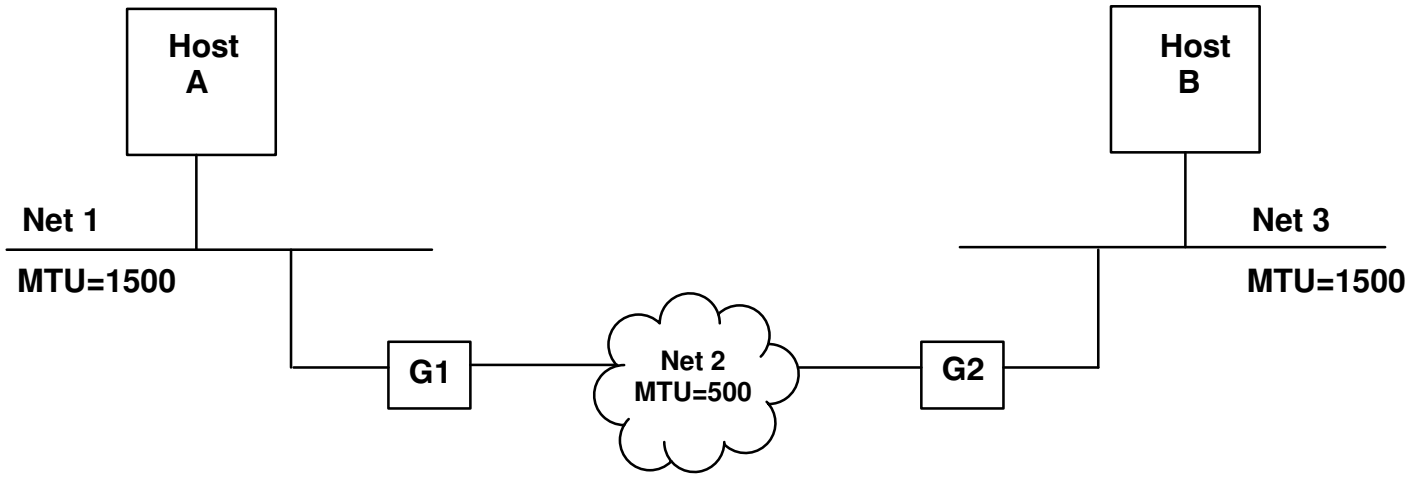
Possible Ways To Accommodate Networks With Different MTUs

- Force datagram to be less than smallest possible MTU
  - — inefficient
  - — cannot know minimum MTU
- Choose initial datagram size that seems appropriate and handle problems later

Datagram Fragmentation

- Needed when datagram larger than network MTU over which it must travel
- Performed by gateways
- Divides datagram into several, smaller datagrams called fragments
- Each fragment routed as independent datagram
- Final destination reassembles fragments

Illustration Of Fragmentation

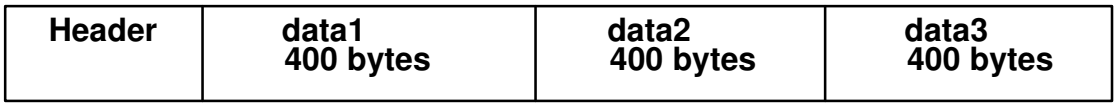


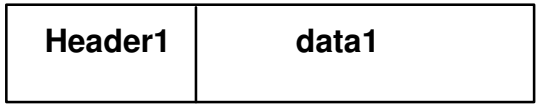
- Hosts A and B send datagrams of up to 1500 octets
- Gateway G1 fragments large datagrams from Host A before sending over Net 2
- Gateway G2 fragments large datagrams from Host B before sending over Net 2

Datagram Fragmentation Details

- Each fragment is a datagram
- Gateway replicates initial datagram header for all fragments
- Offset field in header gives offset in original datagram for data in this fragment
- Fragmentation bit in header indicates this is a fragment
- Additional bit set in header to indicate last fragment

Example of Fragmentation





fragment #1 (offset of 0)

fragment #2 (offset of 400)

fragment #3 (offset of 800)

- Offset specifies where data belongs in original datagram
- Offset actually stored as multiples of 8 octets
- Last fragment bit turned on in header of fragment #3

Best Effort Delivery

- By definition, IP in unreliable
- Gateways can drop packets or change routes, meaning that:
  - datagrams can be lost
  - datagrams can be duplicated
  - datagrams can arrive out of order
  - datagrams can arrive scrambled
- Motivation: allow IP to operate over the widest possible variety of physical networks

In Practice

- IP works (reliably)
- Errors are the exception, not the rule

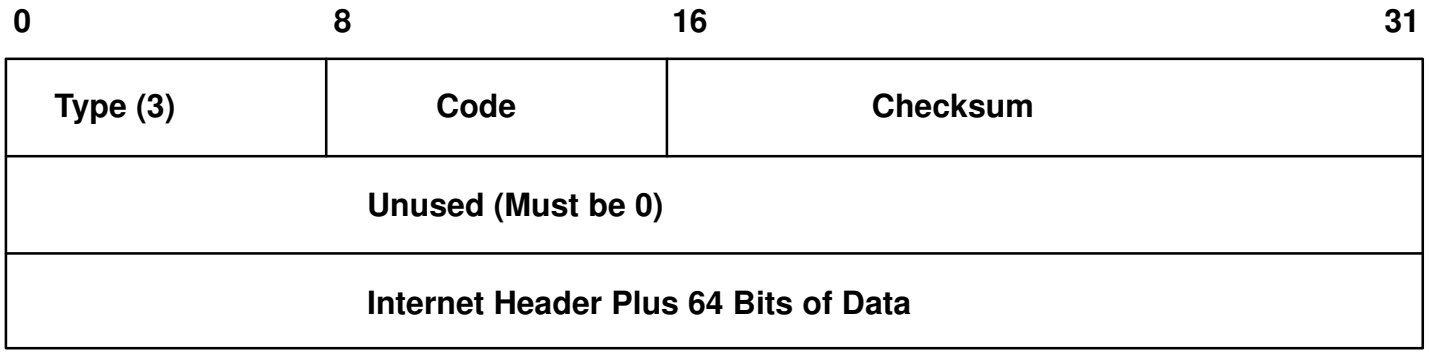
Error Detection And Reporting Mechanisms

- Checksum to detect transmission errors
- Error reporting mechanism to distinguish between lost datagrams and incorrect addresses
- Higher level protocols handle duplicate or out of order datagrams

Error Reporting Mechanism

- Integral part of IP
- Called Internet Control Message Protocol (ICMP)
- Used mostly by gateways to report delivery or routing problems to original source
- Uses IP to carry control messages
- No error messages about error messages

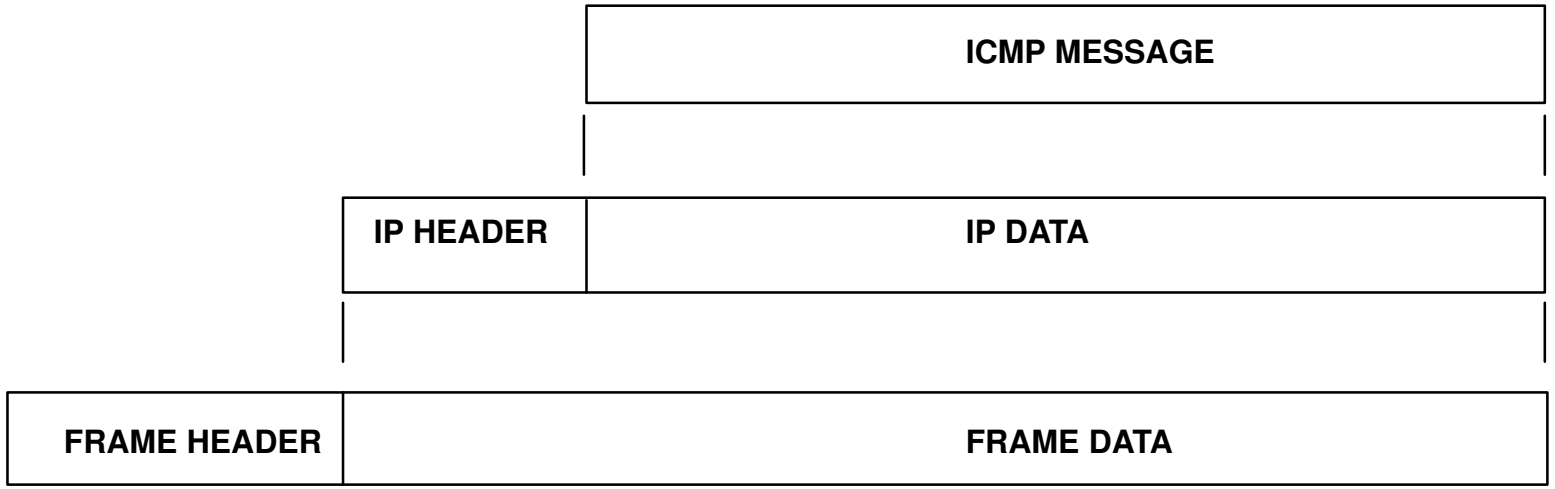
Example ICMP Message Format (Destination Unreachable)



Example ICMP Error Messages

- Destination Unreachable
- Source Quench
- Redirect (change route)
- Time Exceeded
- Echo Request
- Echo Reply
- Address Mask Request
- Address Mask Reply

ICMP Message Encapsulation



- ICMP message has header and data areas
- Complete ICMP message is treated as data in IP datagram
- Complete IP datagram is treated as data in physical network frame

Summary

- Internet consists of physical networks connected by gateways
- Each machine on Internet assigned a 32–bit Internet address
- Address binding software maps Internet addresses to physical hardware addresses
- Datagram is the Internet unit of transfer
- Each datagram carries Internet address of original source and ultimate destination

Summary (continued)

- Datagrams encapsulated in physical hardware frame for delivery
- Internet protocol (IP) defines datagram format and routing
- Datagrams can be fragmented if needed; ultimate destination reassembles them
- Error reporting mechanism (ICMP) uses IP to carry error messages back to original source