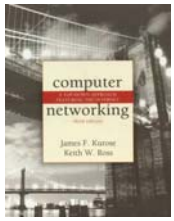
Chapter 4
Network Layer



Computer Networking: A Top Down Approach Featuring the Internet, 3rd edition.
Jim Kurose, Keith Ross
Addison-Wesley, July 2004.

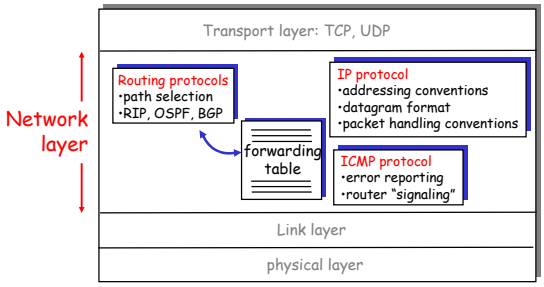
A note on the use of these ppt slides:
We're making these slides freely available to all (faculty, students, readers). They're in PowerPoint form so you can add, modify, and delete slides (including this one) and slide content to suit your needs. They obviously represent a *lot* of work on our part. In return for use, we only ask the following:
- If you use these slides (e.g., in a class) in substantially unaltered form, that you mention their source (after all, we'd like people to use our book!)
- If you post any slides in substantially unaltered form on a www site, that you note that they are adapted from (or perhaps identical to) our slides, and note our copyright of this material.

Thanks and enjoy! JFK/KWR

Chapter 4: Network Layer

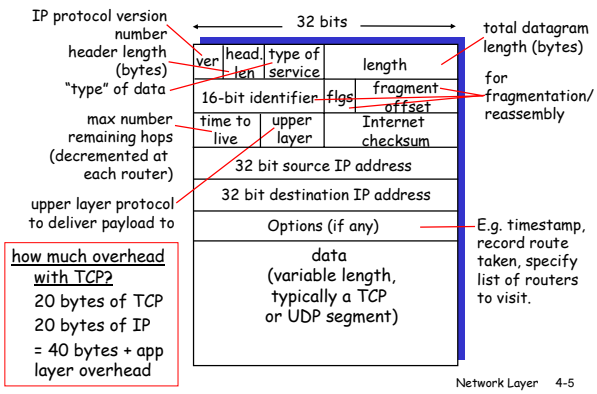
The Internet Network layer

Host, router network layer functions:



Chapter 4: Network Layer

IP datagram format

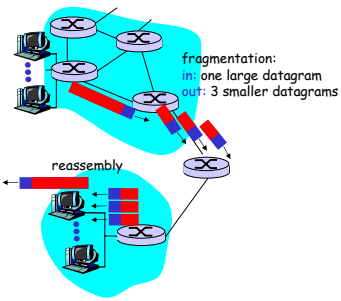


how much overhead with TCP?
- 20 bytes of TCP
- 20 bytes of IP
- = 40 bytes + app layer overhead

IP Fragmentation & Reassembly

network links have MTU (max.transfer size) - largest possible link-level frame.
- different link types, different MTUs

large IP datagram divided ("fragmented") within net
- one datagram becomes several datagrams
- "reassembled" only at final destination
- IP header bits used to identify, order related fragments

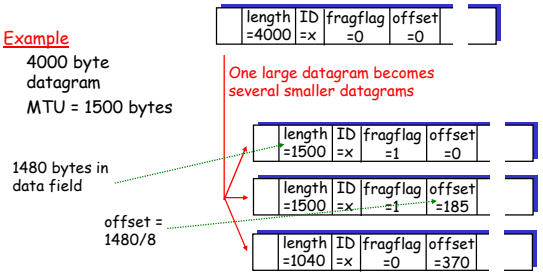


IP Fragmentation and Reassembly

Example

4000 byte datagram

MTU = 1500 bytes



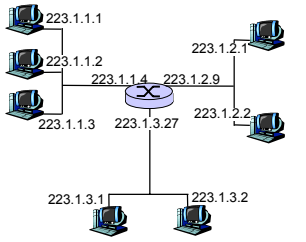
Chapter 4: Network Layer

IP Addressing: introduction

IP address: 32-bit identifier for host, router *interface*

interface: connection between host/router and physical link

- router's typically have multiple interfaces
- host may have multiple interfaces
- IP addresses associated with each interface



223.1.1.1 = 11011111 00000001 00000001 00000001

223 1 1 1

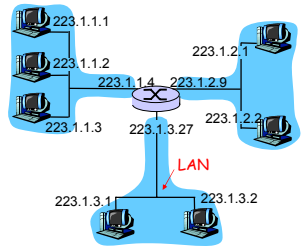
Subnets

IP address:

- subnet part (high order bits)
- host part (low order bits)

What's a subnet ?

- device interfaces with same subnet part of IP address
- can physically reach each other without intervening router

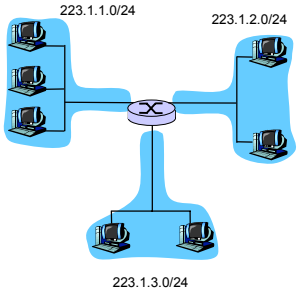


network consisting of 3 subnets

Subnets

Recipe

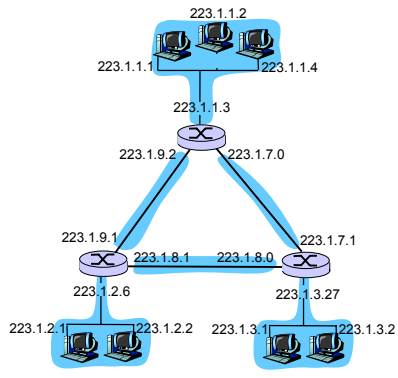
To determine the subnets, detach each interface from its host or router, creating islands of isolated networks. Each isolated network is called a subnet.



Subnet mask: /24

Subnets

How many?

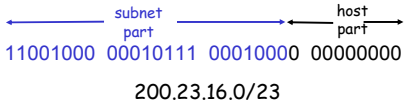


IP addressing: CIDR

CIDR: Classless InterDomain Routing

subnet portion of address of arbitrary length

address format: a.b.c.d/x, where x is # bits in subnet portion of address



200.23.16.0/23

IP addresses: how to get one?

Q: How does *host* get IP address?

hard-coded by system admin in a file

Wintel: control-panel->network->configuration->tcp/ip->properties

UNIX: /etc/rc.config

DHCP: Dynamic Host Configuration Protocol: dynamically get address from as server

"plug-and-play"

(more in next chapter)

IP addresses: how to get one?

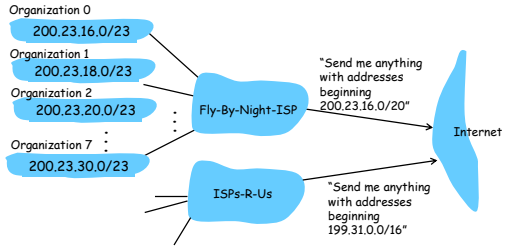
Q: How does *network* get subnet part of IP addr?

A: gets allocated portion of its provider ISP's address space

ISP's block	11001000 00010111 00010000 00000000	200.23.16.0/20
Organization 0	11001000 00010111 00010000 00000000	200.23.16.0/23
Organization 1	11001000 00010111 00010010 00000000	200.23.18.0/23
Organization 2	11001000 00010111 00010100 00000000	200.23.20.0/23
...	....	....
Organization 7	11001000 00010111 00011110 00000000	200.23.30.0/23

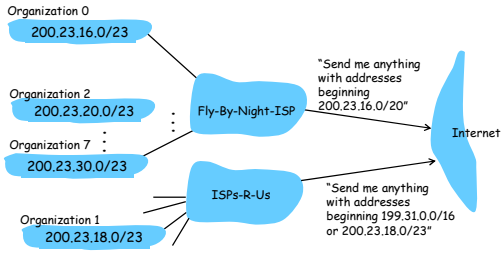
Hierarchical addressing: route aggregation

Hierarchical addressing allows efficient advertisement of routing information:



Hierarchical addressing: more specific routes

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



IP addressing: the last word...

Q: How does an ISP get block of addresses?

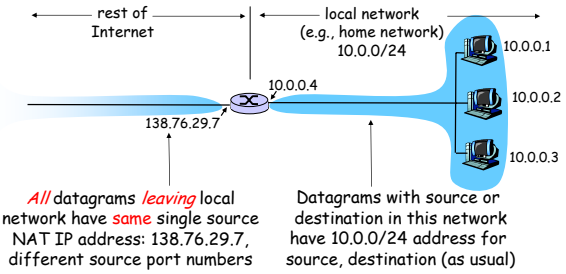
A: ICANN: Internet Corporation for Assigned Names and Numbers

allocates addresses

manages DNS

assigns domain names, resolves disputes

NAT: Network Address Translation



NAT: Network Address Translation

Motivation: local network uses just one IP address as far as outside word is concerned:

- no need to be allocated range of addresses from ISP: - just one IP address is used for all devices
- can change addresses of devices in local network without notifying outside world
- can change ISP without changing addresses of devices in local network
- devices inside local net not explicitly addressable, visible by outside world (a security plus).

NAT: Network Address Translation

Implementation: NAT router must:

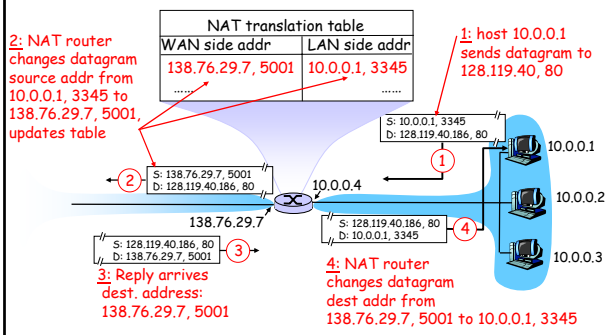
outgoing datagrams: replace (source IP address, port #) of every outgoing datagram to (NAT IP address, new port #)

. . . remote clients/servers will respond using (NAT IP address, new port #) as destination addr.

remember (in NAT translation table) every (source IP address, port #) to (NAT IP address, new port #) translation pair

incoming datagrams: replace (NAT IP address, new port #) in dest fields of every incoming datagram with corresponding (source IP address, port #) stored in NAT table

NAT: Network Address Translation



NAT: Network Address Translation

16-bit port-number field:

- 60,000 simultaneous connections with a single LAN-side address!

NAT is controversial:

- routers should only process up to layer 3
- violates end-to-end argument
  - NAT possibility must be taken into account by app designers, eg, P2P applications
- address shortage should instead be solved by IPv6

Chapter 4: Network Layer

ICMP: Internet Control Message Protocol

used by hosts & routers to communicate network-level information
- error reporting: unreachable host, network, port, protocol
- echo request/reply (used by ping)

network-layer "above" IP:
- ICMP msgs carried in IP datagrams

ICMP message: type, code plus first 8 bytes of IP datagram causing error

Type	Code	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control - not used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

Traceroute and ICMP

Source sends series of UDP segments to dest
- First has TTL =1
- Second has TTL=2, etc.
- Unlikely port number

When nth datagram arrives to nth router:
- Router discards datagram
- And sends to source an ICMP message (type 11, code 0)
- Message includes name of router& IP address

When ICMP message arrives, source calculates RTT

Traceroute does this 3 times

Stopping criterion
- UDP segment eventually arrives at destination host
- Destination returns ICMP "host unreachable" packet (type 3, code 3)
- When source gets this ICMP, stops.

Chapter 4: Network Layer

4. 1 Introduction

4.2 Virtual circuit and datagram networks

4.3 What's inside a router

4.4 IP: Internet Protocol
- Datagram format
- IPv4 addressing
- ICMP
- IPv6

4.5 Routing algorithms
- Link state
- Distance Vector
- Hierarchical routing

4.6 Routing in the Internet
- RIP
- OSPF
- BGP

4.7 Broadcast and multicast routing

IPv6

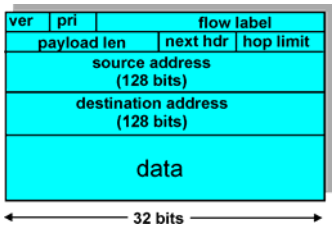
Initial motivation: 32-bit address space soon to be completely allocated.

Additional motivation:
- header format helps speed processing/forwarding
- header changes to facilitate QoS

IPv6 datagram format:
- fixed-length 40 byte header
- no fragmentation allowed

IPv6 Header (Cont)

Priority: identify priority among datagrams in flow
Flow Label: identify datagrams in same "flow." (concept of"flow" not well defined).
Next header: identify upper layer protocol for data



Other Changes from IPv4

Checksum: removed entirely to reduce processing time at each hop

Options: allowed, but outside of header, indicated by "Next Header" field

ICMPv6: new version of ICMP
- additional message types, e.g. "Packet Too Big"
- multicast group management functions

Transition From IPv4 To IPv6

Not all routers can be upgraded simultaneous
- no "flag days"
- How will the network operate with mixed IPv4 and IPv6 routers?

Tunneling: IPv6 carried as payload in IPv4 datagram among IPv4 routers

Tunneling