

Network Layer

- ❖ Network Layer Services
- ❖ Routing principles: path selection
- ❖ Hierarchical routing
- ❖ IP
- ❖ Internet Routing Protocols: Reliable Transfer
 - a. Intra AS Hierarchy
 - b. Inter AS Hierarchy
- ❖ What's inside a router
- ❖ IPv6
- ❖ Mobility

IP Fragmentation and Reassembly

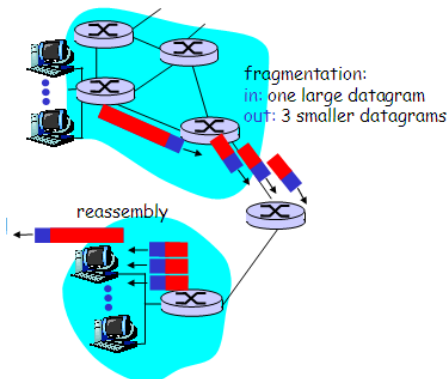
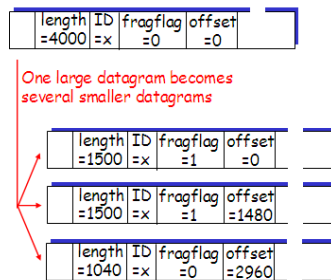
Each network has some MTU (Max Transmission Unit) Size

Strategy

- ❖ Fragment when necessary (MTU < Datagram)
- ❖ IP packet needs to fit in payload part of frame
- ❖ Try to avoid fragmentation at source host
- ❖ Re-fragmentation is possible at routers etc.
- ❖ Fragments are self-contained datagrams
- ❖ All fragments contain same value in **ident field**
- ❖ Each fragment is re-encapsulated in frame
- ❖ Delay reassembly until destination host
- ❖ Do not recover from lost fragments
- ❖ IP header bits used to identify, order related fragments

Example

- ❑ 4000 byte datagram
- ❑ MTU = 1500 bytes



Internet Control Message Protocol (ICMP)

	Type	Code	description
used by hosts, routers, gateways to communication network-level information	0	0	echo reply (ping)
error reporting:	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
echo request/reply (used by ping)	3	7	dest host unknown
	4	0	source quench (congestion control - not used)
network-layer "above" IP:	8	0	echo request (ping)
	9	0	route advertisement
ICMP message: type, code plus first 8 bytes of IP datagram causing error	10	0	router discovery
	11	0	TTL expired
	12	0	bad IP header

Dynamic Host Control Protocol (DHCP)

Goal: allow host to *dynamically* obtain its IP address from network server when it joins network

Can renew its lease on address in use

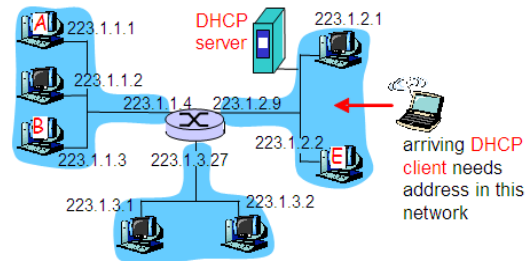
Allows reuse of addresses (only hold address while connected an "on")

Support for mobile users who want to join network (more shortly)

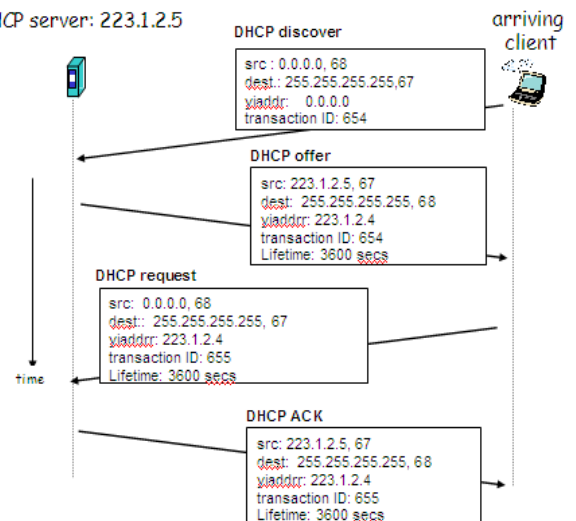
DHCP overview:

- ❖ host broadcasts "DHCP discover" msg
- ❖ DHCP server responds with "DHCP offer" msg
- ❖ host requests IP address: "DHCP request" msg
- ❖ DHCP server sends address: "DHCP ack" msg

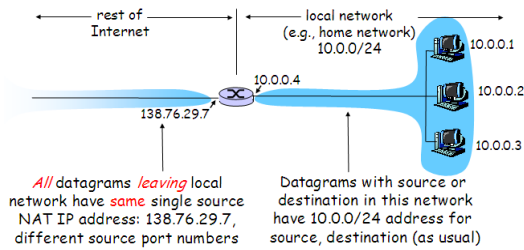
DHCP Client-Server Scenario



DHCP server: 223.1.2.5



Network Address Translation (NAT)

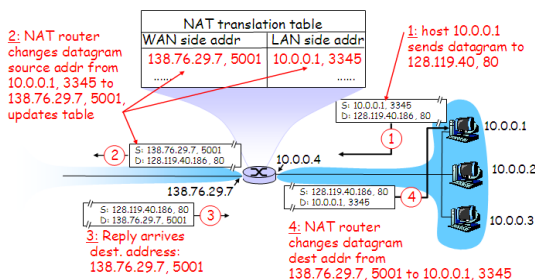


Motivation: local network uses just one IP address as far as outside world is concerned:

- ❖ no need to be allocated range of address 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 visible by outside world (a security plus).

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 address.
- ❖ *remember in translation table* every (source IP add, port#) to (NAT IP add, new port#) translation pair
- ❖ *incoming datagrams:* replace (NAT IP address, new port #) in destination fields of every incoming datagram with corresponding (source IP address, port #) stored in NAT table



Conclusion:

16-bit port-number field: 65,536 simultaneous connections with a single LAN-side address!
 NAT is controversial as routers should only process up to layer 3.
 NAT possibility must be envisaged by application developers IP address shortage instead solved by IPv6

Internet Routing Protocols: Reliable Transfer

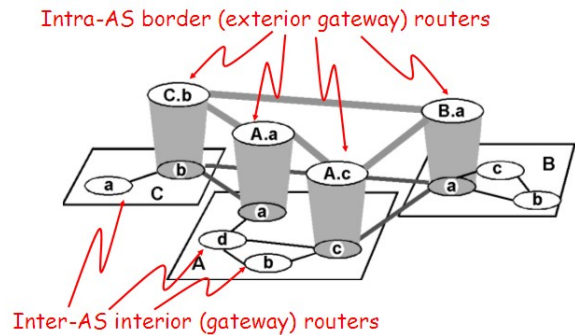
The Internet consists of **Autonomous Systems (AS)** interconnected with each other:

- Stub AS:** small corporation: 1 connection to other AS
- Multi homed AS:** large corporation (no transit): multiple connections to other AS's
- Transit AS:** provider hooking many AS's.

Two-level routing:

- Intra-AS:** administrator responsible for choice of routing algorithm within network
- Inter-AS:** one standard 4 inter-AS routing: BGP

Internet AS Hierarchy



Intra AS Hierarchy

- ❖ Also known as Interior Gateway Protocols (IGP)
- ❖ Most common Intra-AS routing protocols:
 - RIP:** Routing Information Protocol
 - OSPF:** Open Shortest Path First
 - IGRP:** Interior Gateway Routing Protocol (Cisco proprietary)

Inter-AS Routing in the Internet: BGP

- ❖ **BGP** (Border Gateway Protocol): *de facto* standard
- ❖ **Path Vector** protocol: similar to Distance Vector protocol each Border Gateway broadcast to neighbours *entire path* (i.e., sequence of AS's) to destination BGP routes to networks (ASs), not individual hosts

Why different Intra- and Inter-AS routing?

Policy:

- Inter-AS: administrator wants control over how its traffic routed, who routes through its network.
- Intra-AS: single administrator, so no policy decisions needed

Scale:

- hierarchical routing saves table size, reduced update traffic

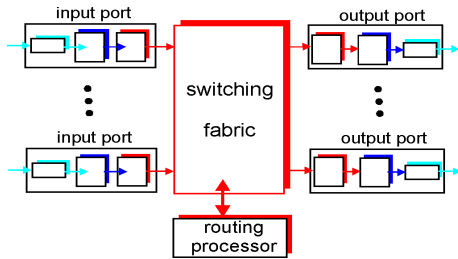
Performance:

- Intra-AS: can focus on performance
- Inter-AS: policy may dominate over performance

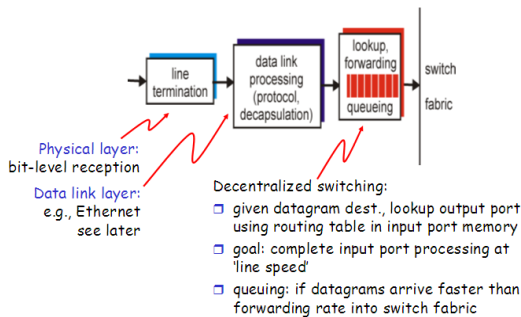
What's inside a router?

Two key router functions:

run routing algorithms/protocol (RIP, OSPF, BGP)
switching datagrams from incoming to outgoing link.

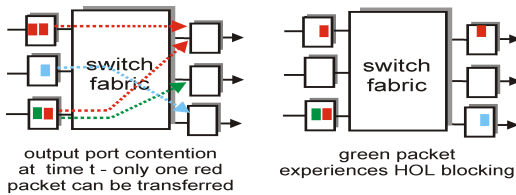


Input port functions



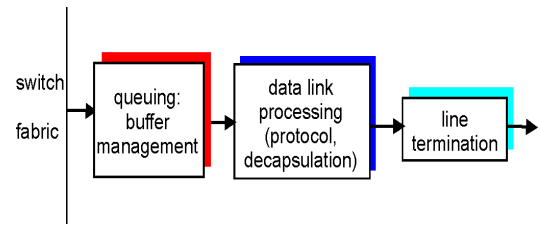
Input port queuing

- Fabric slower than input ports combined -> queuing may occur at input queues
- Head-of-the-Line (HOL) blocking: queued datagram at front of queue prevents others in queue from moving forward
- queuing delay and loss due to input buffer overflow!

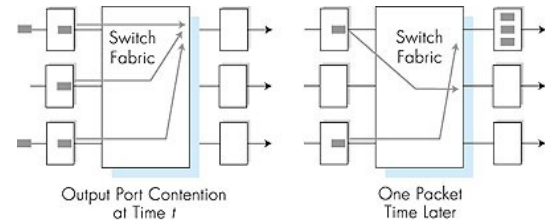


Output ports

- ❖ *Buffering* required when datagrams arrive from fabric faster than the transmission rate
- ❖ *Scheduling discipline* chooses among queued datagrams for transmission



- ❖ buffering when arrival rate via switch exceeds output line speed
- ❖ *queuing (delay) and loss due to output port buffer overflow!*



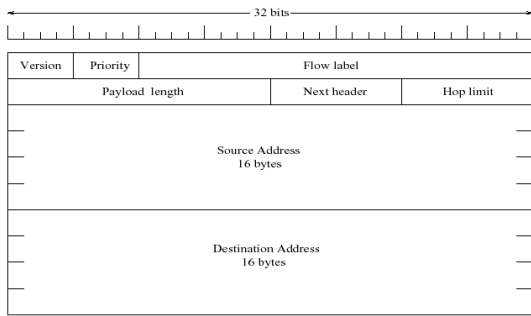
IPv6

- ❖ CIDR may last for a few years, but everyone realizes that the days of IP in its current form (IPv4) are numbered. (by 2008, all addresses will run out)
 - ❖ Besides growing number of mobile stations, it may not be long before every TV set is an Internet node, producing billion machines.
 - ❖ IETF has started working on a new version of IP, which would never run out of addresses, and solving a variety of other problems.
- Its major goals were to:
- ❖ Supports billions of hosts, even with inefficient address space allocation.
 - ❖ Reduce the size of the routing tables.
 - ❖ Simplify IP to allow routers to process packets faster.
 - ❖ Provide better security than current IP.
 - ❖ Pay more attention to type of service, particularly for real-time data.
 - ❖ Aid multicasting by allowing scopes to be specified.
 - Make it possible for a host to roam without changing its address.
 - Allow the protocol to evolve in the future.
 - Permit the old and new protocols to coexist for years.

Major Features

- ❖ 128-bit addresses (3.4 X 10³⁸ addresses)
- ❖ Auto configuration
- ❖ Real-time service
- ❖ Authentication and security
- ❖ End-to-end fragmentation
- ❖ Protocol extensions

IPv6 Header Format



- ❖ Fixed 40-byte “base” header
- ❖ **Extension headers** (fixed order, mostly fixed length)
 - fragmentation
 - source routing
 - authentication and security
 - other options
- ❖ The **Version field** is always 6 for IPv6 and 4 for Ipv4.
- ❖ The **Priority field** is used to distinguish between packets whose sources can be flow controlled and those that cannot.
 - 0-7 are capable of slowing down in the event of congestion. Lower numbers are less important (1 for news, 4 for FTP, etc).
 - 8-15 are for real-time traffic (audio, video).
- ❖ The **Flow label** is still experimental, but will be used to allow set up of particular properties between source and destination.
- ❖ The **Payload length** field tells how many bytes follow the 40-byte header
- ❖ The **Next header** field tells which of the (currently) six **extension headers** follows this one. If this header is last IP header the next header would be transport layer (TCP, UDP, etc.)
- ❖ The **Hop limit** field is used to keep packets from living forever. In IPv4 this was Time To Live.
- ❖ The **Source and Destination fields** are fixed-length 16-byte (128 bit) addresses.

Other Changes from Ipv4

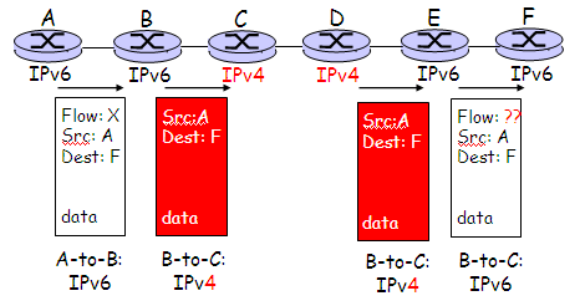
- ❖ **Checksum**: removed entirely to reduce processing time at each hop.

- ❖ **Options**: allowed, but outside of header, indicated by “Next Header” field.
- ❖ **ICMPv6**: new version of ICMP
 - extra message types, e.g. “Packet Too Big”

Transition from IPv4 to IPv6

- ❖ Not all routers can be upgraded simultaneously:
 - no “flag days”
 - How will the network operate with mixed IPv4 and IPv6 routers?
 -
- ❖ Two proposed approaches:
 - **Dual Stack**: some routers with dual stack (v6, v4) can “translate” between formats
 - **Tunneling**: IPv6 carried as payload in IPv4 datagram among IPv4 routers

Dual Stack Approach



Note: flow header info lost

Tunnelling Approach

