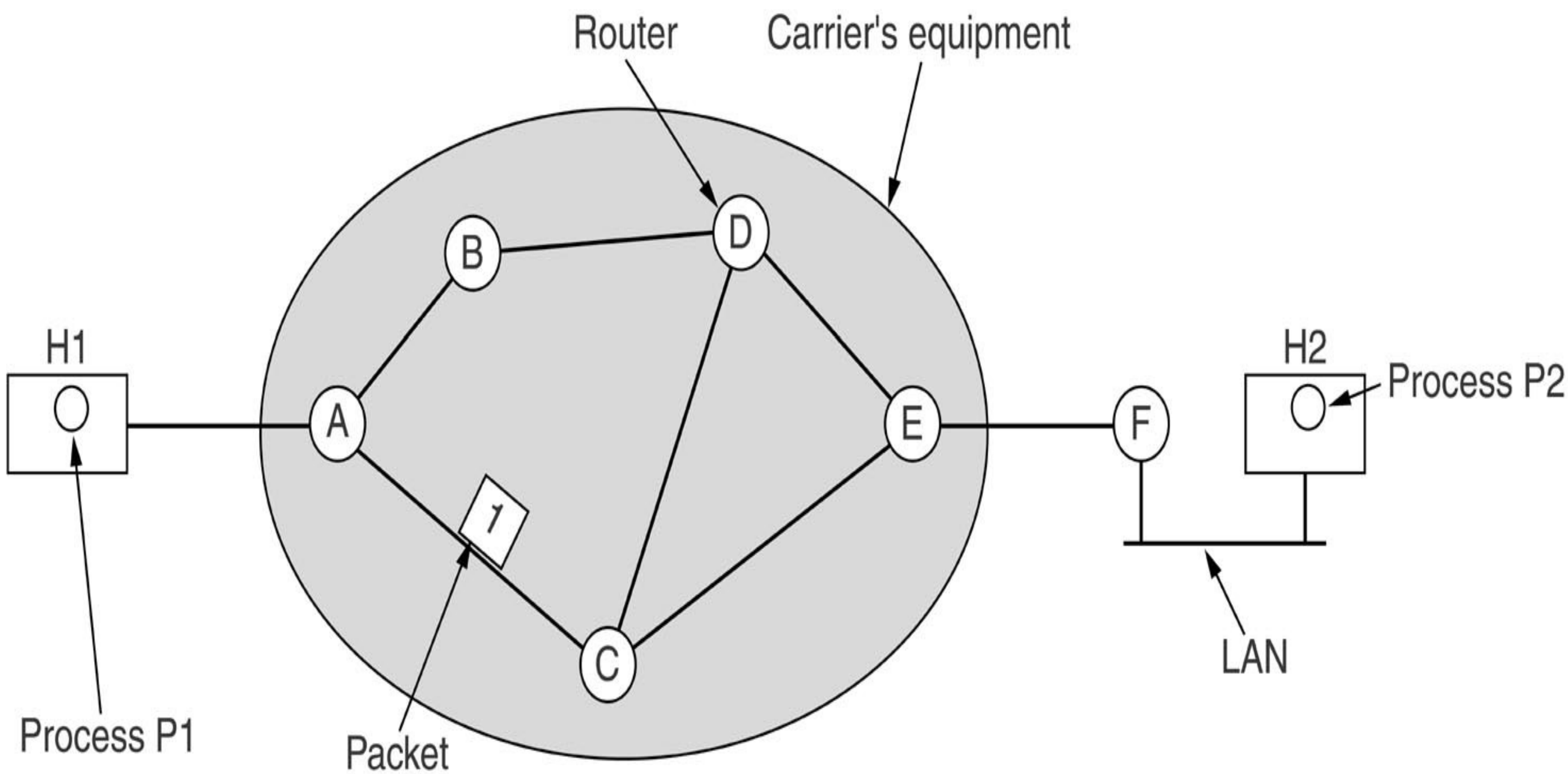
Chapter IIII
The IP Layer

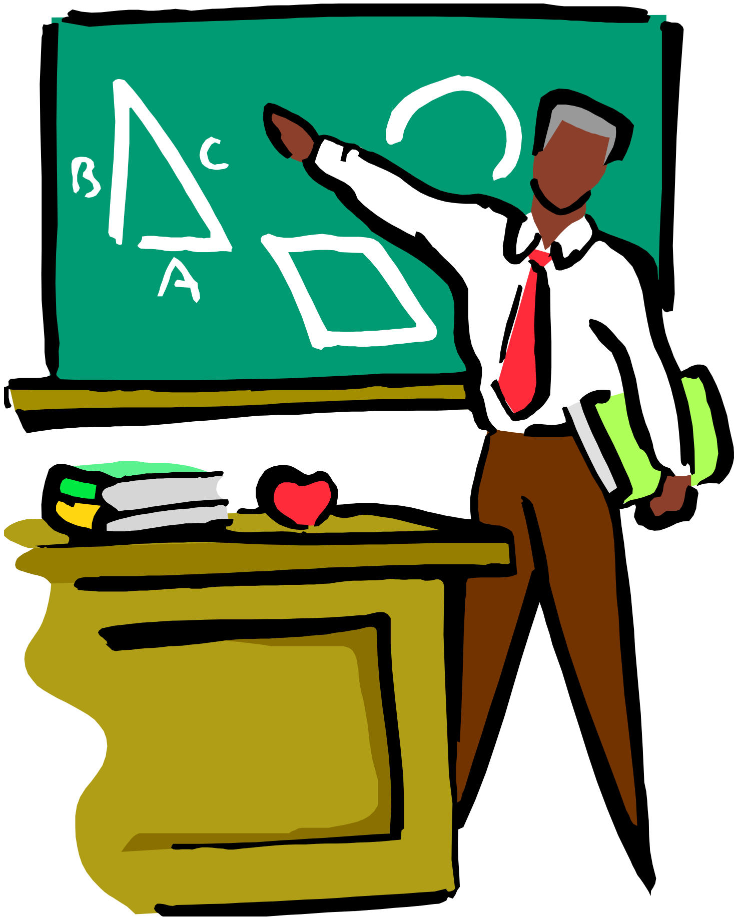
Network layer



Network layer

- Key design issues
  - Services provided to the transport layer (connection oriented vs. connectionless services)
  - Routing algorithms
  - Congestion control
  - Quality of services
  - Internetworking

The IP Layer (or network layer in Internet)



- 1 - Design choices
- 2 - IPv4 / IPv6
- 3 - Mobility management
- 4 - Routing in Internet

Design choices

The Internet

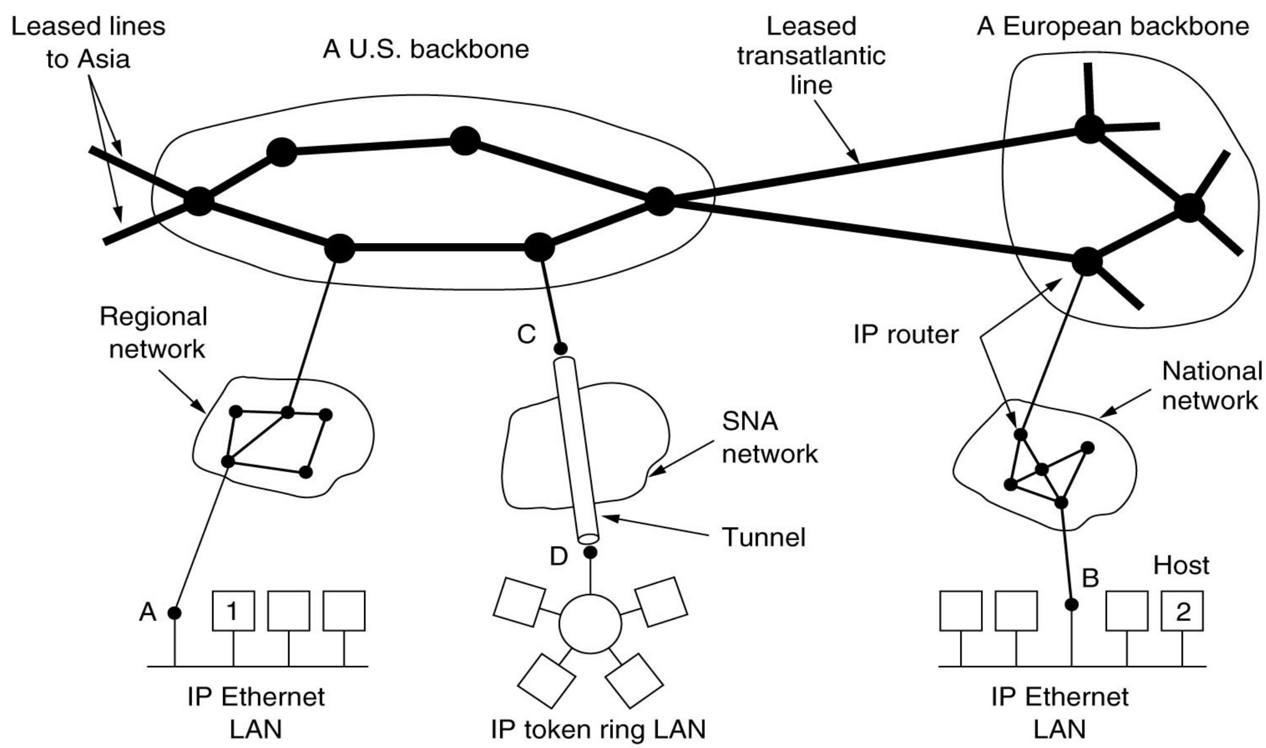


Figure 5.52 - Reference [1]

Design choices

- Key design objectives
  1. Make sure it works
  2. Keep it simple
  3. Make clear choices
  4. Exploit modularity
  5. Expect heterogeneity
  6. Avoid static options and parameters
  7. Look for a good design, it needs not be perfect
  8. Be strict when sending and tolerant when receiving
  9. Think about scalability
  10. Consider performance and cost

Design choices

- Choices
  - Services provided to the transport layer (connection oriented vs. connectionless services)
    - Connectionless only
  - Routing algorithms
    - Interior Gateway Routing Protocol
      - Open Shortest Path First (OSPF)
    - Exterior Gateway Routing Protocol
      - Border Gateway Protocol (BGP)

Design choices

- Choices
  - Congestion control
    - Left to upper layers
  - Quality of services
    - Best effort
      - More sophisticated/refined features left to upper layers
  - Internetworking
    - IP as the glue

IPv4

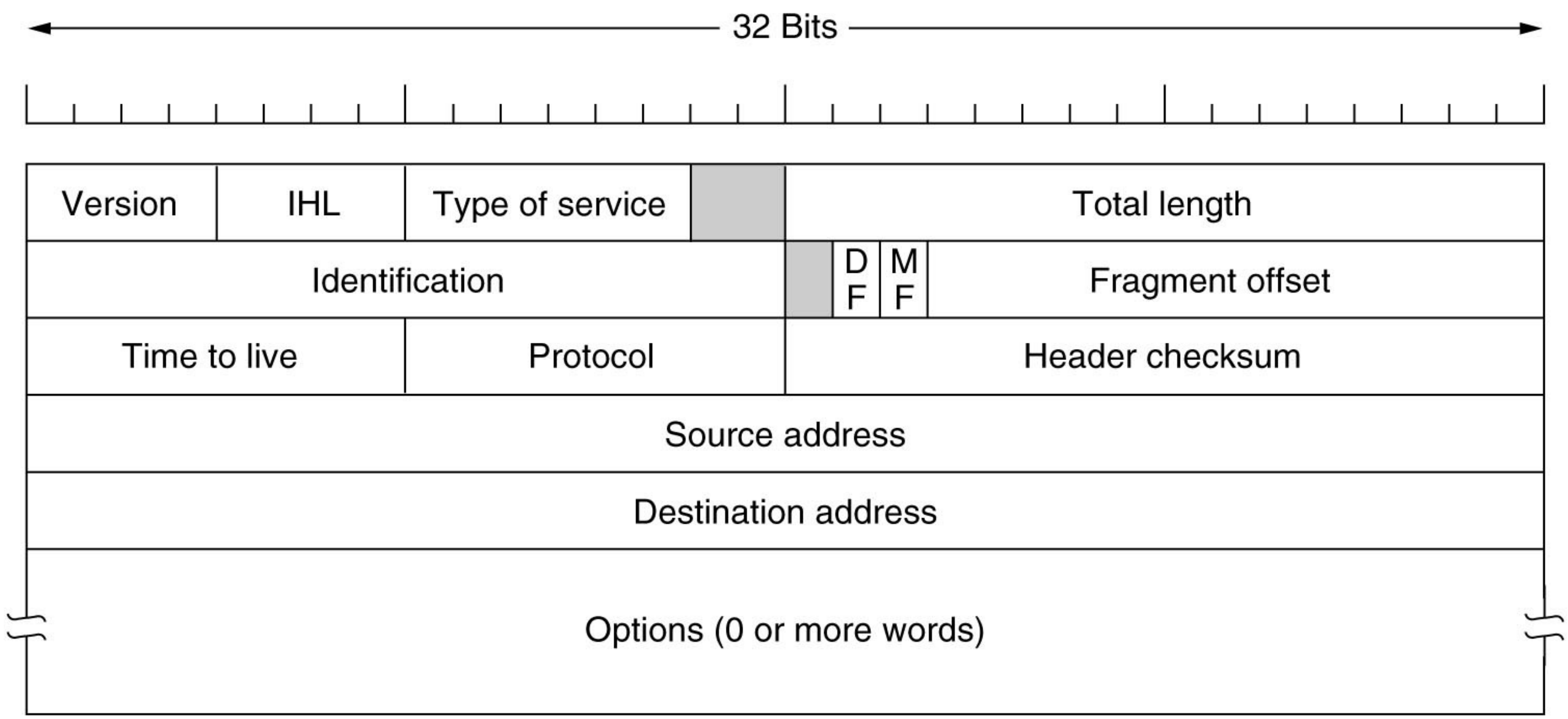


Figure 5.53 - Reference [1]

IPv4

- Header (20 byte fixed and variable length optional part)
  - Version
  - IHL: Length in 32 bit words
    - Minimum: 5 (No option is present)
    - Maximum: 15 (header 60 bytes and options 40 bytes)
  - Type of service (Early efforts for quality of services)
  - Total length: header + data (65,535 bytes)
  - Identification: Determine to which datagram a fragment belongs to
  - Fragment / do not fragment
  - More fragments

IPv4

- Header (20 byte fixed and variable length optional part)
  - Fragment offset:
    - Where in the current datagram the fragment belongs
  - Time to live
  - Protocol:
    - to which transport process the datagram should be given to (UDP or TCP)
  - Header checksum
  - Source address / destination address
  - Options (e.g. strict source routing, loose source routing, record route, timestamp)

IPv4

- IP addresses

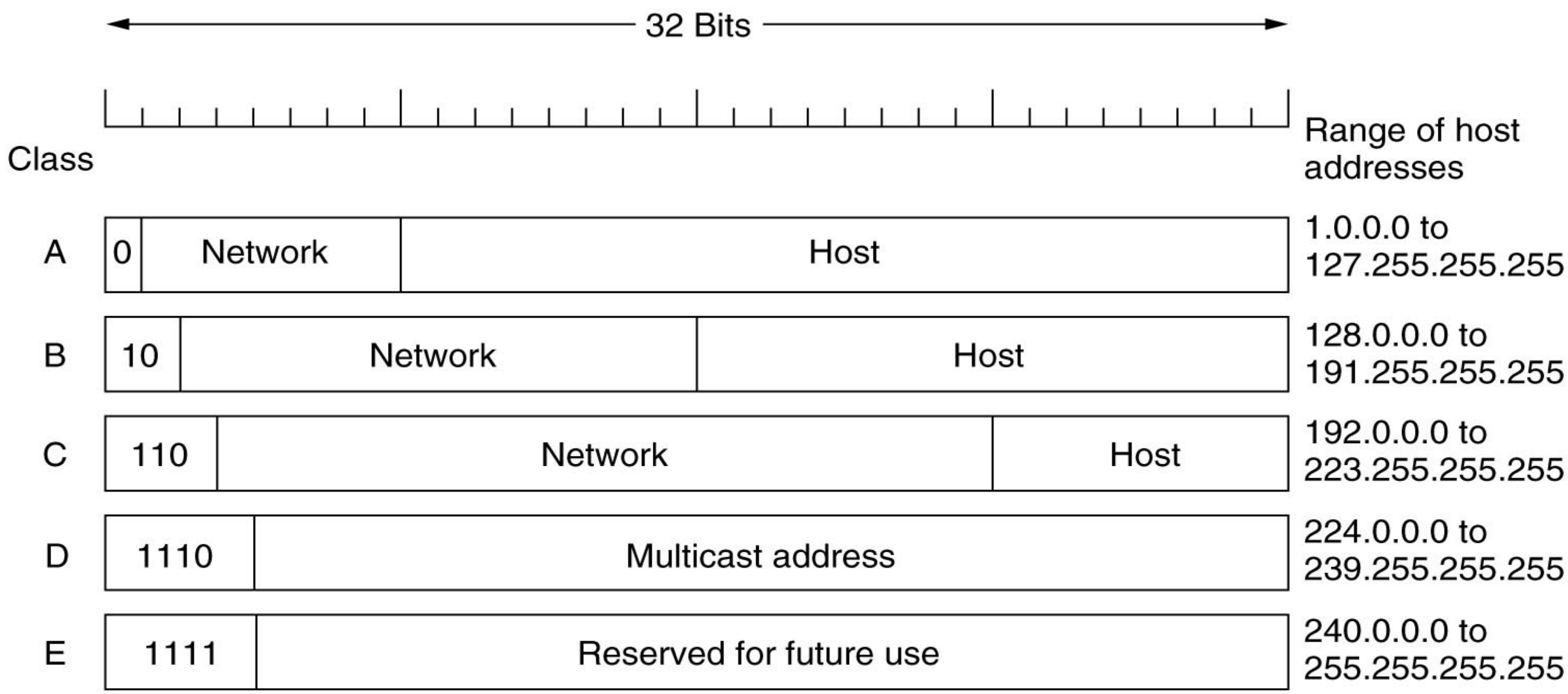


Figure 5.55 – Reference [1]

IPv4

- Some early quick fixes to the IP address shortage issue
  - Classless Inter Domain Routing (CIDR)
    - Allocate remaining addresses in variable size blocks, without regard to classes
    - Make routing much more complex
  - Network Address Translation (NAT)
    - Only 1 IP address seeing from outside
    - Several IP addresses inside (i.e. 1 per host)
    - Translation process
      - Same set of internal addresses could be used by different organizations

IPv4

- Some early quick fixes to the IP address shortage issue
  - Classless Inter Domain Routing (CIDR)
    - Allocate remaining addresses in variable size blocks, without regard to classes
    - Make routing much more complex
  - Network Address Translation (NAT)
    - Only 1 IP address seeing from outside
    - Several IP addresses inside (i.e. 1 per host)
    - Translation process
      - Same set of internal addresses could be used by different organizations

IPv4

- Network Address Translation (NAT)

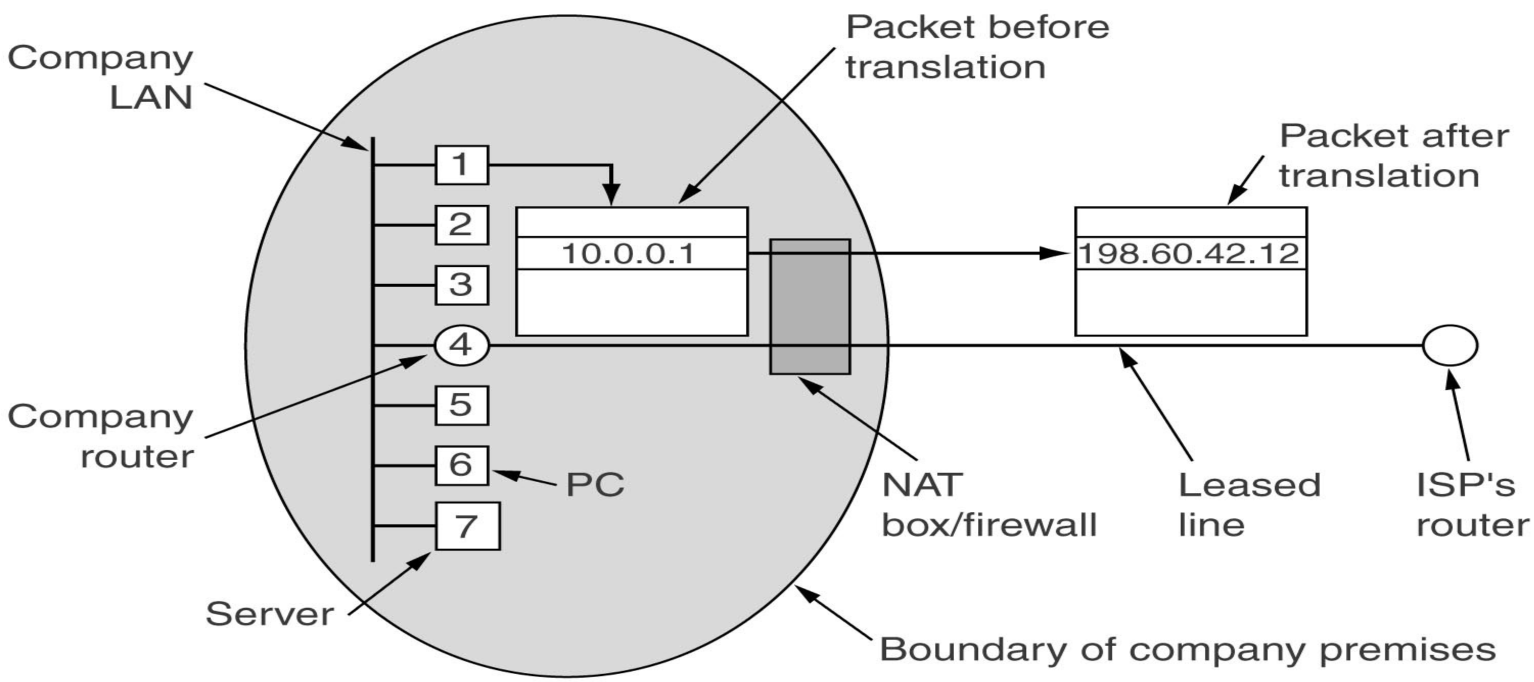


Figure 5.60 – Reference [1]

IPv6

- Some of the design goals
  - Support of billions of hosts
  - Reduce the size of routing tables
  - Simplify the protocol
  - Provide better security
  - Pay more attention to type of service
  - Aid multicasting
  - Enable roaming without address change
  - Enable evolution of the protocol
  - Enable co-existence IPv4 / IPv6

IPv6

- The main header

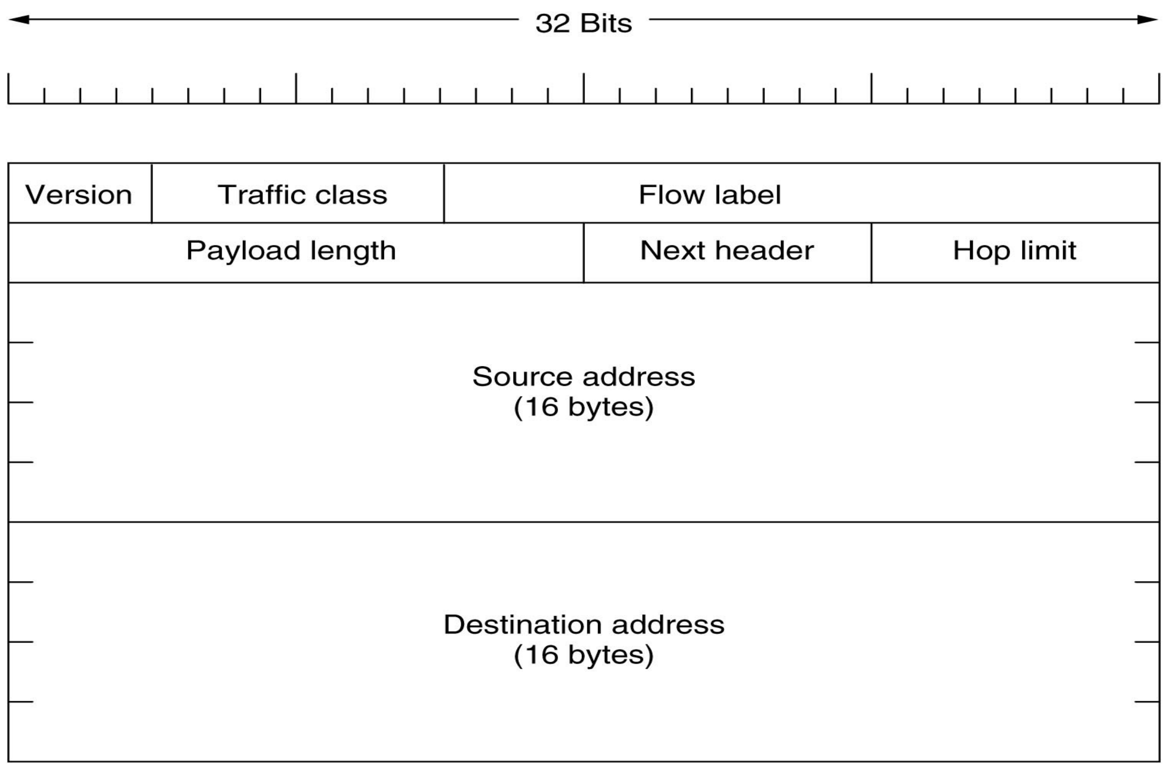


Figure 5.68 - Reference [1]

IPv6

- Header
  - Version (6 for IPv6 and 4 for IPv4)
  - Traffic class: Distinguish between packets with different delivery requirements
  - Flow label: still under experiment – Enable pseudo connection to mimic connection oriented services
  - Payload length: how many bytes follow the 40-byte header
  - Hop limit
  - Next header: which one of the currently 6 optional headers follows this one, if any
  - Source address, destination address: 16 bytes addresses instead of 4 in IPv4

IPv4/IPv6 Integration and Coexistence

- Key working assumptions
  - Deployment of IPv6 at the edge first
  - Full deployment (including core) last
    - Realistic and easy because most OS deployed on user sites are IPv6 capable.
  - Key related issues
    - Transportation of IPv6 packets from edge to edge through an IPv4 capable core
    - Conversion of IPv4 packets into IPv6 packets and vice versa on user sites.

IPv4/IPv6 Integration and Coexistence

- Key techniques
  - IPv6 / IPv4 dual stack
    - Use of a new API that supports both IPv6 and IPv4
    - Requirements
      - Upgrade of entire infrastructure
      - Dual addressing scheme
      - Dual management
      - Dual routing tables
    - Validity
      - Specific network infrastructure with a mix of IPv4 and IPv6
        - » Campus network
        - » Points of presence

IPv4/IPv6 Integration and Coexistence

- Key techniques
  - IPv6 over IPv4 tunnels
    - IPv6 packets are encapsulated in IPv4 packets
    - Requirements
      - Support of a dual stack by the two end points of the tunnel
    - Validity
      - Quite suitable when dual stacks are implemented at the edges and the core remains IPv4

IPv4/IPv6 Integration and Coexistence

- Key techniques
  - IPv4 – IPv6 Translation mechanisms
    - Two categories
      - No change to IPv4 and IPv6
        - » TCP-UDP relay mechanism
          - » Runs on a dedicated server
          - » separate transport level connection with IPv4 and IPv6
      - Change to IPv4 and/or IPv6
        - » Name resolver, address mapper and translator added to IPv4 between the network layer and the higher layer

IPv4/IPv6 Integration and Coexistence

- Key techniques
  - Requirements
    - Vary depending on the category and the specific mechanism
    - Examples
      - » Dedicated server
  - Validity
    - Will enable use of legacy IPv4 applications when IPv6 becomes widely deployed

IPv4/IPv6 Integration and Coexistence

- Key techniques
  - IPv6 over MPLS (Multiple Protocol Label Switching) backbone
    - MPLS
      - Switching using labels instead of IP addresses
        - » Inherent VPN features
    - No reconfiguration of core routers
    - Requirements
      - Depend on mechanisms used
    - Validity
      - No impact on MPLS infrastructure

Mobility Management

- Service continuity when moving from sub-networks to sub-networks
  - Should be transparent to higher layer protocols
    - Key challenge
      - IP address no more valid when hosts move to different networks
    - Different from the ability to detach from a network and attach to a new one
      - New IP address assigned in this case without service disruption because there is no requirement on service continuity

Mobility Management

- The way it is done in cellular networks
  - Location management
    - Registration / updating
    - Paging
  - Handoff management
    - Intra-cell (i.e. change of radio channel)
    - Inter-cell (i.e. change of base station)

Mobility Management

- Classification scheme for Internet
  - Macro mobility
    - Mobility across regional networks
    - Schemes: Mobile IP (MIP): MIPv4, MIPv6
  - Micro mobility
    - Mobility within regional networks
    - Examples of schemes: Cellular IP, HAWAI
  - Seamless mobility
    - “Right” mix of macro mobility and micro mobility

Mobility Management

- Macro mobility
  - Mobile IPv4
    - Key concepts
      - Mobile host (MH)
      - Two IP addresses
        - » Home address
        - » Care of (COA) address
      - Two new entities
        - » Home agent (HA)
        - » Foreign agent (FA)

Mobility Management

- Macro mobility
  - Mobile IPv4
    - Key phases
      - Agent discovery
      - Registration
      - Routing

Mobility Management

- Macro mobility
  - Mobile IPv4
    - Agent discovery (i.e. Need to detect MH has changed point of attachment)
      - Agent advertisements transmitted periodically by HA and FA
      - Extension of Internet Control Message Protocol (ICMP)
      - Detection may be based on lifetime field of the router advertisement
        - » ICMP
          - » Reports when something unexpected happens / Test Internet
            - » Ex: destination unreachable, time exceeded, echo/echo reply

Mobility Management

- Macro mobility
  - Mobile IPv4
    - Registration
      - Goal: Make HA aware of the whereabouts of MH
      - May (or may not) go through FA
      - Two messages (carried over UDP)
        - » Registration request
        - » Registration reply

Mobility Management

- Macro mobility
  - Mobile IPv4
    - Routing
      - HA
        1. Intercepts packets sent to MH home address
           » Gratuitous Address Resolution Protocol (ARP) packets
             » ARP address maps IP address on MAC address
             » Gratuitous ARP packets enables the re-directions to HA of all packets sent to MH home address

Mobility Management

- Macro mobility
  - Mobile IPv4
    - Routing
      - HA
        1. Tunnels packets to CoA
           - » End of tunnel
             - » MH
               - » Or
             - » FA

Mobility Management

- Macro mobility
  - Mobile IPv6
    - Same fundamental principles as Mobile IPv4
      - Some differences
        1. No foreign agent (FA)
           » IPv6 MH acquire their CoA without the assistance of FA
        2. HA discovery done using anycast
           » More efficient than the broadcast used in Mobile IPv4

Mobility Management

- Macro mobility
  - Shortcomings
    - High signalling load
      - Especially when MH is within a region
        - » Macro signalling actually not needed
    - Latency when restoring communications paths
      - Packets may be dropped

Mobility Management

- Micro mobility
  - Cellular IP (CIP)
    - Columbia University and Ericsson
      - CIP access network
        - » Base stations
          - » Wireless interface to MH
          - » Routing and location management
        - » CIP nodes
          - » Routing and location management only
        - » Gateway
          - » CIP nodes that bridge CIP access network and a MIP networks

Mobility Management

- Micro mobility
  - Cellular IP (CIP)
    - Columbia University and Ericsson
      - Routing
        - » Hierarchical (MH – BS – CIP node – Gateway)
          - » Example of advantage compared to MIP use in the same access network
            - » MH to MH packets within a CIP network do not leave the CIP network
              - » No need to travel back and forth between HA and FA

Routing in Internet

- Open Short Path First (OSPF)
  - Routing within autonomous systems
    - Key design goals/requirements
      - 1. Openness although used with autonomous systems
      - 2. Plurality of metrics (e.g. physical distance, delay)
      - 3. Dynamicity (i.e. adaptation to changes in the network)
      - 4. Load balancing
      - 5. Support of hierarchical systems
      - 6. Security

Routing in Internet

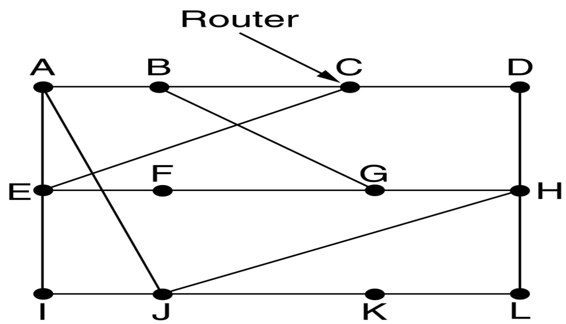
- Open Short Path First (OSPF)
  - Key features
    - Link state algorithm
    - Flooding algorithm
    - Shortest path algorithm
    - Authenticated exchanges

Routing in Internet

- Open Short Path First (OSPF)
  - Link state algorithm
    - Replacement of the distance vector algorithm
      - Distance vector algorithm in a nutshell
        - » Also known as Bellman Ford routing
          - » Use of a plurality of metrics
          - » Each router maintains a table with the best known distance to each destination and the next hop to reach the destination
          - » Table updated with the information received from the neighbours

Routing in Internet

- Distance vector algorithm



(a)

To	A	I	H	K	New estimated delay from J	Line
A	0	24	20	21	8	A
B	12	36	31	28	20	A
C	25	18	19	36	28	I
D	40	27	8	24	20	H
E	14	7	30	22	17	I
F	23	20	19	40	30	I
G	18	31	6	31	18	H
H	17	20	0	19	12	H
I	21	0	14	22	10	I
J	9	11	7	10	0	–
K	24	22	22	0	6	K
L	29	33	9	9	15	K
	JA delay is 8	JI delay is 10	JH delay is 12	JK delay is 6	New routing table for J	

$\underbrace{\text{A, I, H, K}}_{\text{Vectors received from J's four neighbors}}$

(b)

Figure 5.9 – Reference [1]

Routing in Internet

- Distance vector algorithm

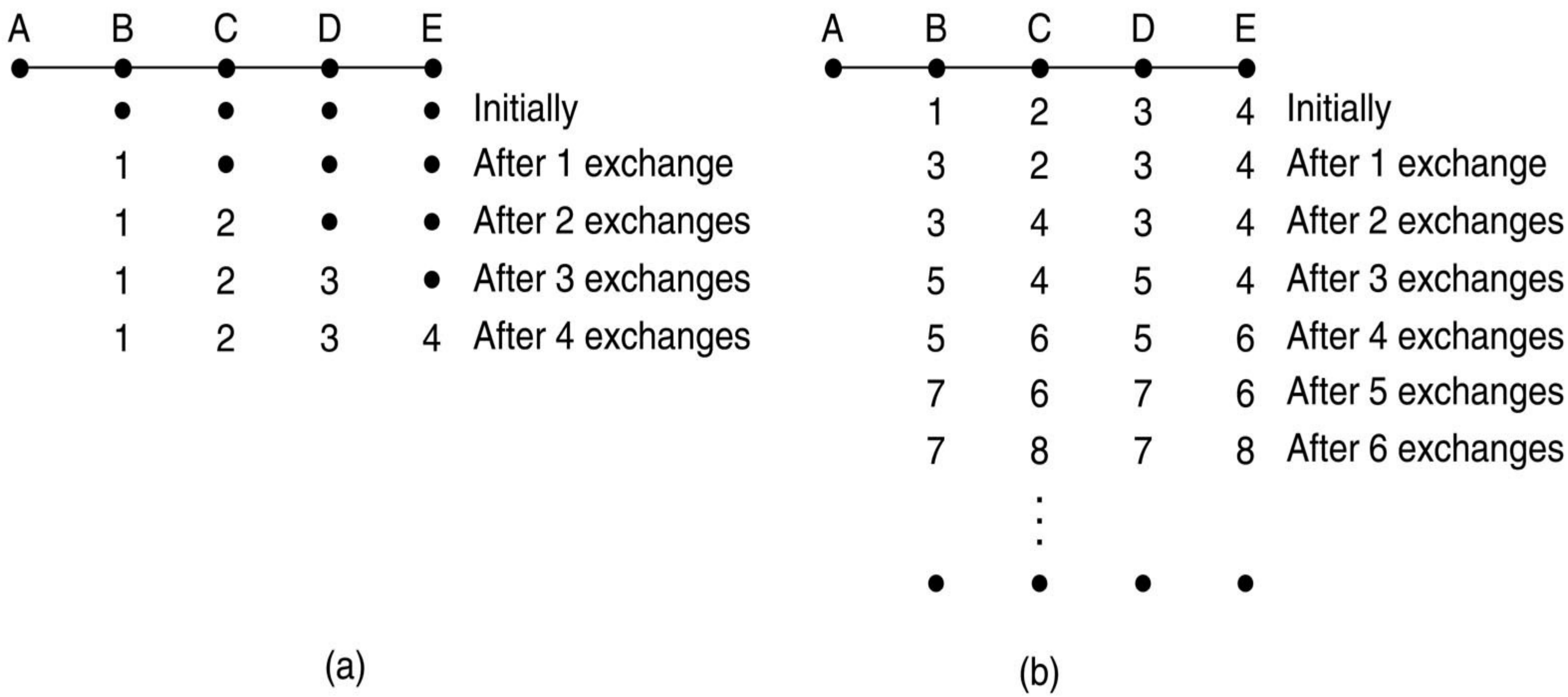


Figure 5.10 – Reference [1]

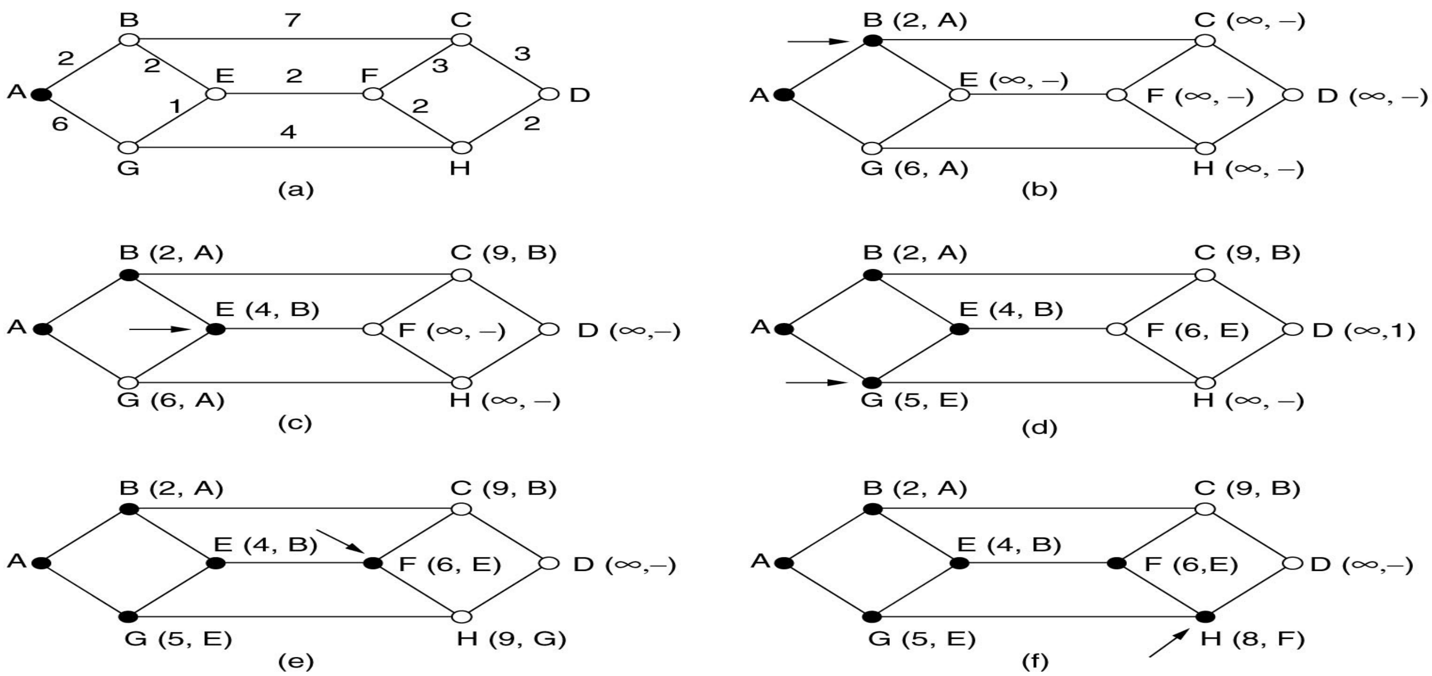
Routing in Internet

- Link state algorithm
  - Solves the count to infinity problem
    - Information received from all the other routers instead of just the neighbouring routers.
- Flooding algorithm
  - Used by each router to send information to all the other routers
    - Every packet received by a router is sent to all the neighbouring routers
    - Maybe selective (e.g. sent to all the other routers except the one from which it was received).

Routing in Internet

Shortest path algorithm (i.e. Djikstra algorithm)

- Rooted in graph theory



Fiugre 5.7 - Reference [1]

Routing in Internet

- Border Gateway Protocol (BGP)
  - Routing across autonomous systems
    - Additional requirement
      - Business model / political considerations
        - » Traffic from a given source AS should not transit by in a given Ass to reach a given destination AS
          - » Business model
          - » Political issues
    - Design choice to address the new requirement
      - Configurable policies on each router

Routing in Internet

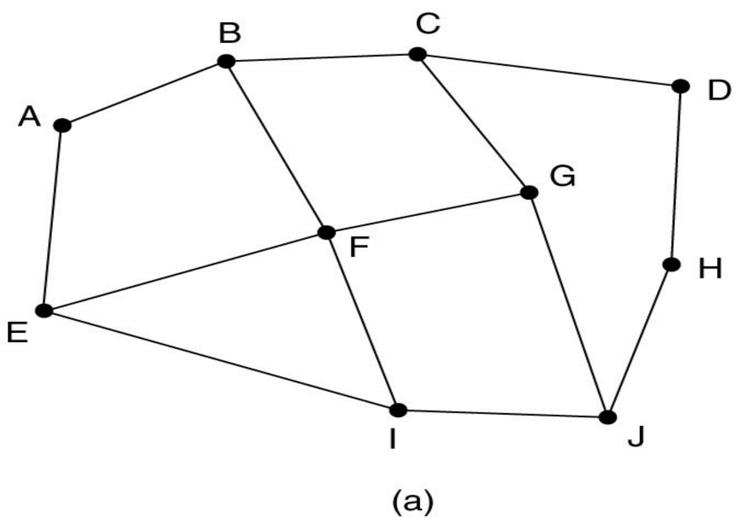
- Border Gateway Protocol (BGP)
  - Three types of autonomous systems
    - 1. Stub networks
      - Only 1 connection in the graph
      - No possibility to carry transit traffic
    - 2. Multi-connected networks
      - May be used to carry transit traffic (if they wish)
    - 3. Transit networks (e.g. backbones)
      - Willing to carry traffic
        - » For pay
        - » Eventual restrictions

Routing in Internet

- Border Gateway Protocol (BGP)
  - Key features
    - Distance vector algorithm
      - Restrictions in graph topology
        - » No possibility of count to infinity
    - Shortest path algorithm
      - Paths that do not respect configured policies are excluded even if they are the shortest
    - Authenticated exchanges

Routing in Internet

- Border Gateway Protocol (BGP)



Information F receives
from its neighbors about D

From B: "I use BCD"
From G: "I use GCD"
From I: "I use IFGCD"
From E: "I use EFGCD"

(b)

Figure 5.67 - Reference [1]

References

1. A. Tanenbaum, Computer Networks, Fourth Edition, Prentice Hall, 2003 (Introduction)
2. M. Tatipamula and P. Grossetete, IPv6 Integration and Co-existence Strategies for Next Generation Networks, IEEE Communications Magazine, January 2004
3. A. Salkintzis, editor, Mobile Internet – Enabling Technologies and Services, Chapter 5, CRC Press, 2004