

# 《计算机网络原理》

课程编号：40240513

讲课教师：吴建平 徐明伟 尹霞

本科生必修课

计算机科学与技术系

# 主要教学内容和学时分配

第一章	引言	3
第二章	计算机网络的体系结构	6
第三章	数据通信的基本原理	3
第四章	物理层	3
第五章	数据链路层	9
第六章	局域网和介质访问控制	6
第七章	网络层	6
第八章	传送层	3
第九章	计算机网络应用	6
第十章	计算机网络新技术/复习	3
共计		48

# 第七章 网络层

## 第二部分

# 主要内容 (1)

## 7.1 网络层概述

## 7.2 路由算法

- 7.2.1 最优化原则
- 7.2.2 最短路径路由算法
- 7.2.3 洪泛算法
- 7.2.4 基于流量的路由算法
- 7.2.5 距离向量路由算法
- 7.2.6 链路状态路由算法
- 7.2.7 分层路由
- 7.2.8 移动主机的路由

# 主要内容 (2)

## 7.3 拥塞控制算法

7.3.1 拥塞控制的基本原理

7.3.2 拥塞控制算法

## 7.4 网络互连

7.4.1 级联虚电路

7.4.2 无连接网络互连

7.4.3 隧道技术

7.4.4 互联网路由

7.4.5 分段

7.4.6 防火墙

# 主要内容 (3)

## 7.5 互联网网络层协议

7.5.1 IPv4协议

7.5.2 IP地址

7.5.3 互联网控制协议

7.5.4 路由信息协议 RIP

7.5.5 内部网关路由协议: OSPF

7.5.6 外部网关路由协议: BGP

7.5.7 无类域间路由CIDR

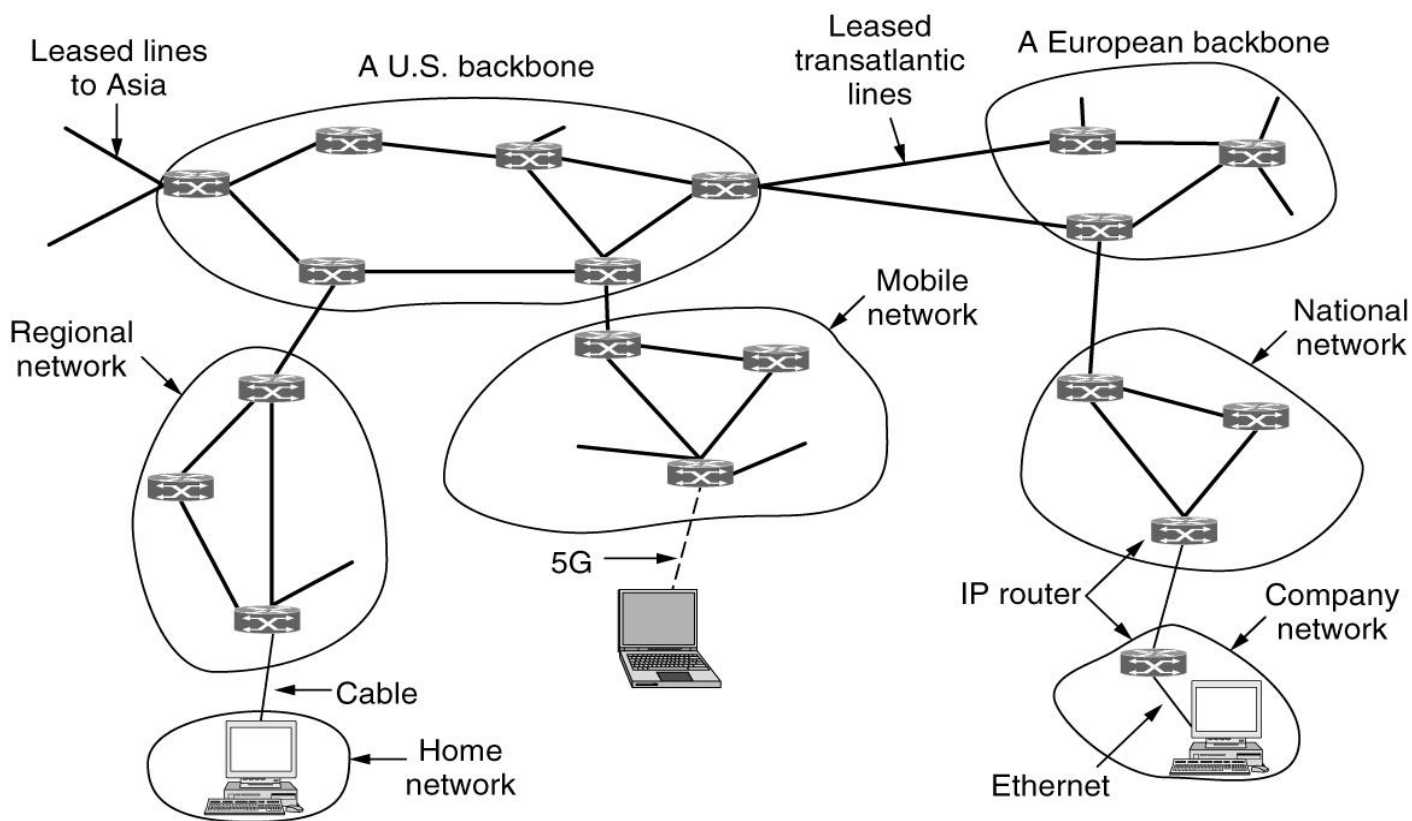
7.5.8 IPv6

## 7.6 路由器体系结构和关键技术

# 互联网网络层的10大设计原则 (RFC 1958)

- 保证工作
- 保持简单
- 明确选择
- 采用模块化思想
- 预期异构性
- 避免静态选择和参数
- 寻找好的设计，不需要最优
- 严格发送，宽松接收
- 考虑可扩展性
- 考虑性能和成本

## 7.5 互联网的网络层 (1)



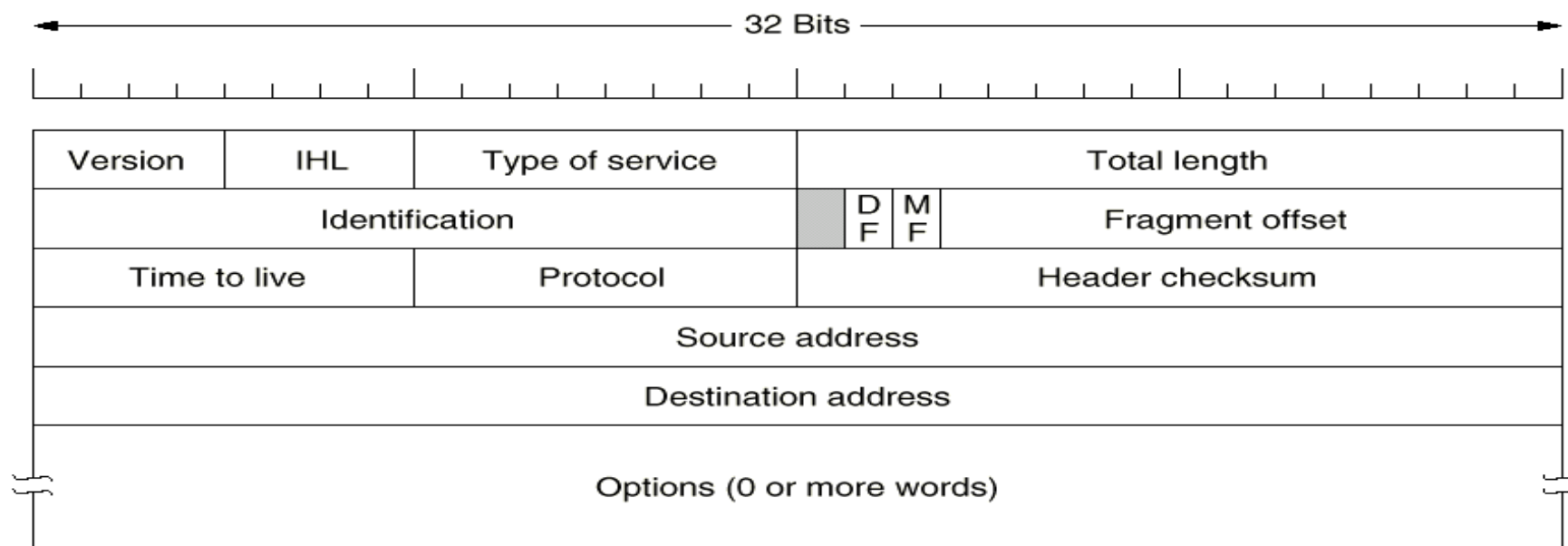
- 网络之间互连的纽带是IP (Internet Protocol) 协议



# 7.5 互联网的网络层协议 (2)

## 7.5.1 IPv4协议

### – IPv4头格式



IPv4头包括20个字节的固定部分和变长（最长40字节）的可选部分，从左到右传输；

## 7.5 互联网的网络层协议 (3)

- 版本域 (Version) ;
- IHL: IP 包头长度, 最小为5, 最大为15, 单位为32-bit word;
- 服务类型域 (Differentiated Service)
  - 3个优先级位;
  - 3个标志位: D (Delay) 、 T (Throughput) 、 R (Reliability) ;
  - 2个保留位;
  - 目前, 很多路由器都忽略服务类型域。
- 总长度域 (Total length)
- 标识域 (Identification)

## 7.5 互联网的网络层协议 (4)

- DF: Don't Fragment;
  - 所有机器必须能够接收小于等于576字节的段
- MF: More Fragments
  - 除最后一个段外的所有段都要置MF位。
- 段偏移量 (Fragment offset)
  - 除最后一个段外的所有段的长度必须是8字节 (基本段长) 的倍数。
- 生存期 (Time to live)
  - 实际实现中, IP包每经过一个路由器TTL减1, 为0则丢弃, 并给源主机发送一个告警包。

## 7.5 互联网的网络层协议 (5)

- 协议域 (Protocol) : 上层为哪种传输协议, TCP、UDP...
- 头校验和 (Header checksum)
  - 只对IP包头做校验;
  - 算法: 每16位求反, 循环相加 (进位加在末尾), 和再求反。
- 源地址 (Source address)
- 目的地址 (Destination address)
- 选择项 (Options)
  - 变长, 长度为4字节的倍数, 不够则填充, 最长为40字节;

## 7.5 互联网的网络层协议 (6)

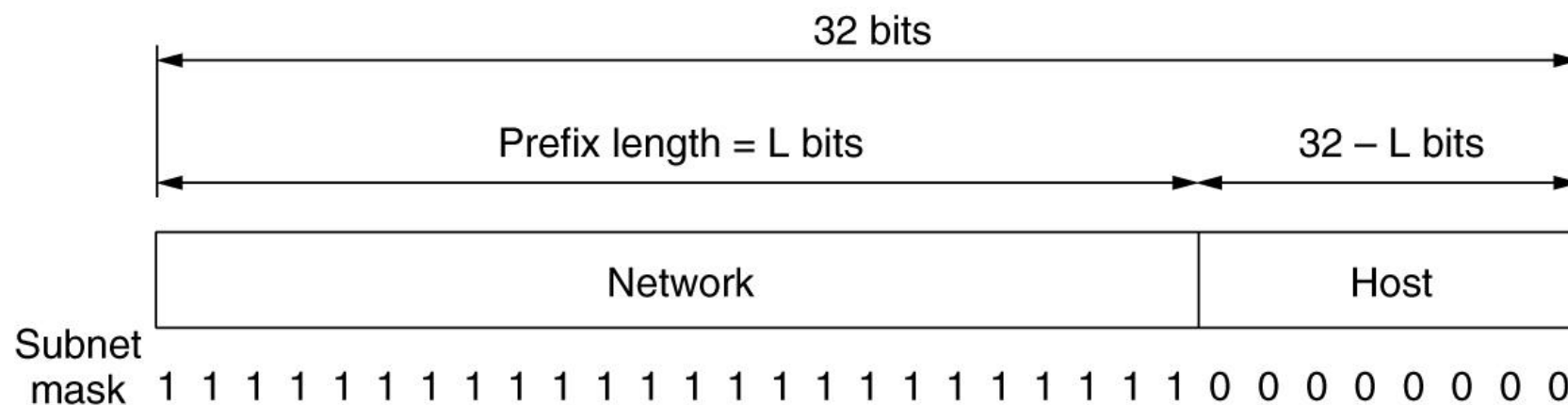
Option	Description
Security	Specifies how secret the datagram is
Strict source routing	Gives the complete path to be followed
Loose source routing	Gives a list of routers not to be missed
Record route	Makes each router append its IP address
Timestamp	Makes each router append its address and timestamp

**Fig. 5-46.** IP options.

# 7.5 互联网的网络层协议 (7)

- 7.5.2 IP地址

- 前缀：一个网络对应于IP地址空间的一个连续块

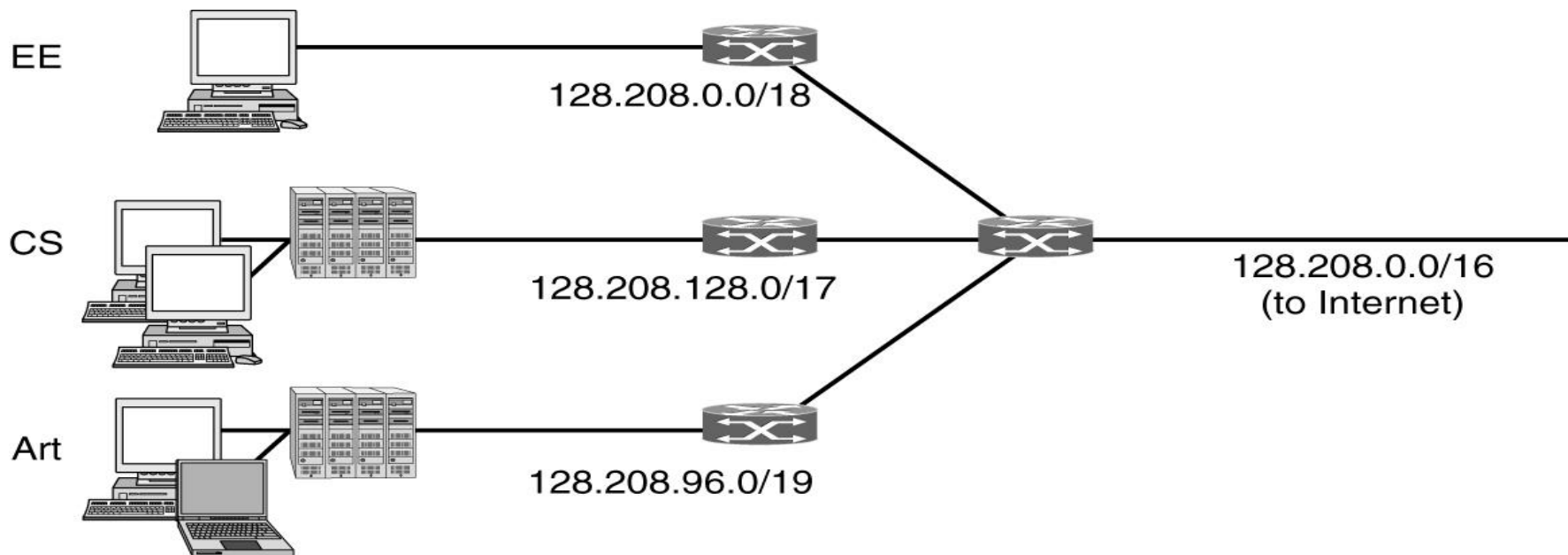


一个前缀和一个子网掩码

## 7.5 互联网的网络层协议 (8)

- 7.5.2 IPv4地址

- 子网

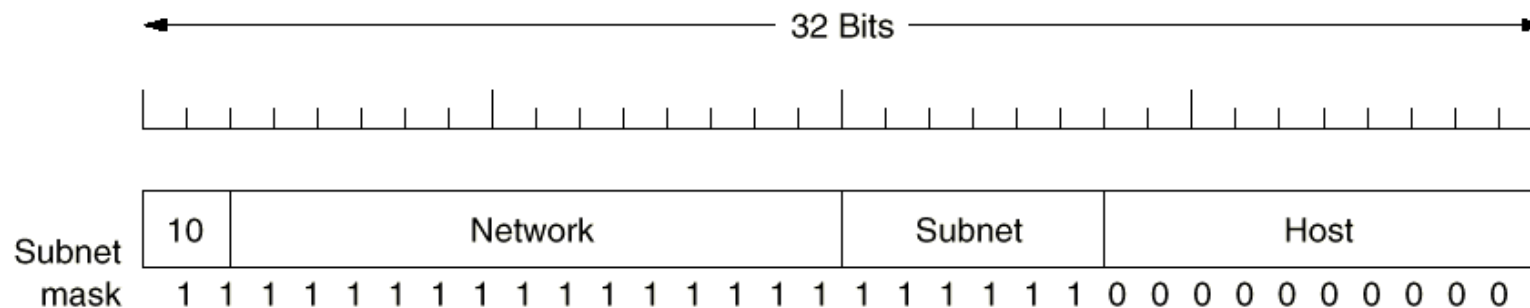


通过子网划分的做法将一个IP前缀切分为单独的网络

# 7.5 互联网的网络层协议 (9)

- 7.5.2 IP地址

- 子网：分而治之的思想：为了便于管理和使用，可以将网络分成若干供内部使用的部分，称为子网。对外界，该网络还是一个单独的网络。





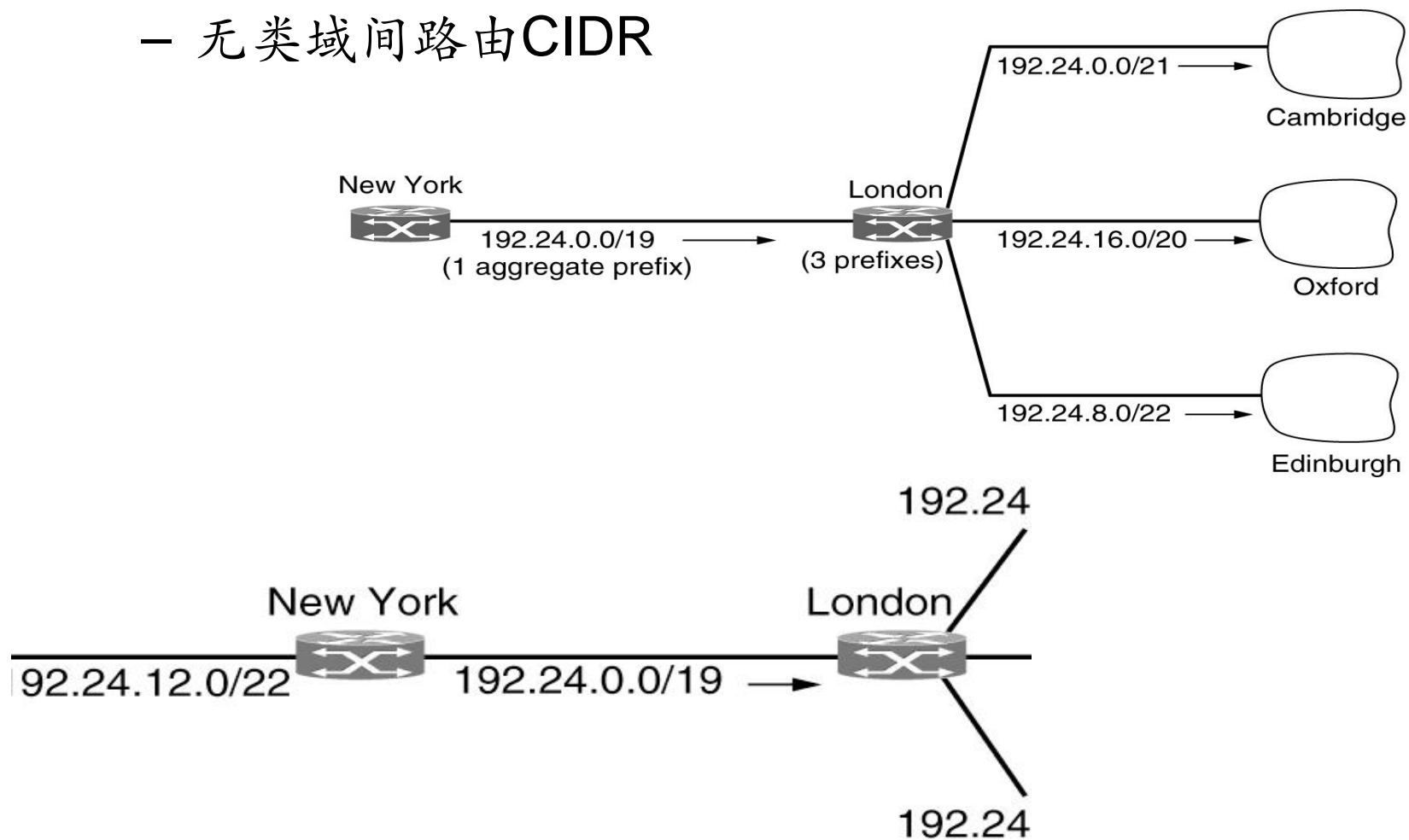
# 7.5 互联网的网络层协议 (10)

- 7.5.2 IP地址
  - 无类域间路由CIDR, RFC4632 (2006)

University	First address	Last address	How many	Prefix
Cambridge	194.24.0.0	194.24.7.255	2048	194.24.0.0/21
Edinburgh	194.24.8.0	194.24.11.255	1024	194.24.8.0/22
(Available)	194.24.12.0	194.24.15.255	1024	194.24.12.0/22
Oxford	194.24.16.0	194.24.31.255	4096	194.24.16.0/20

# 7.5 互联网的网络层协议 (11)

- 7.5.2 IP地址
  - 无类域间路由CIDR

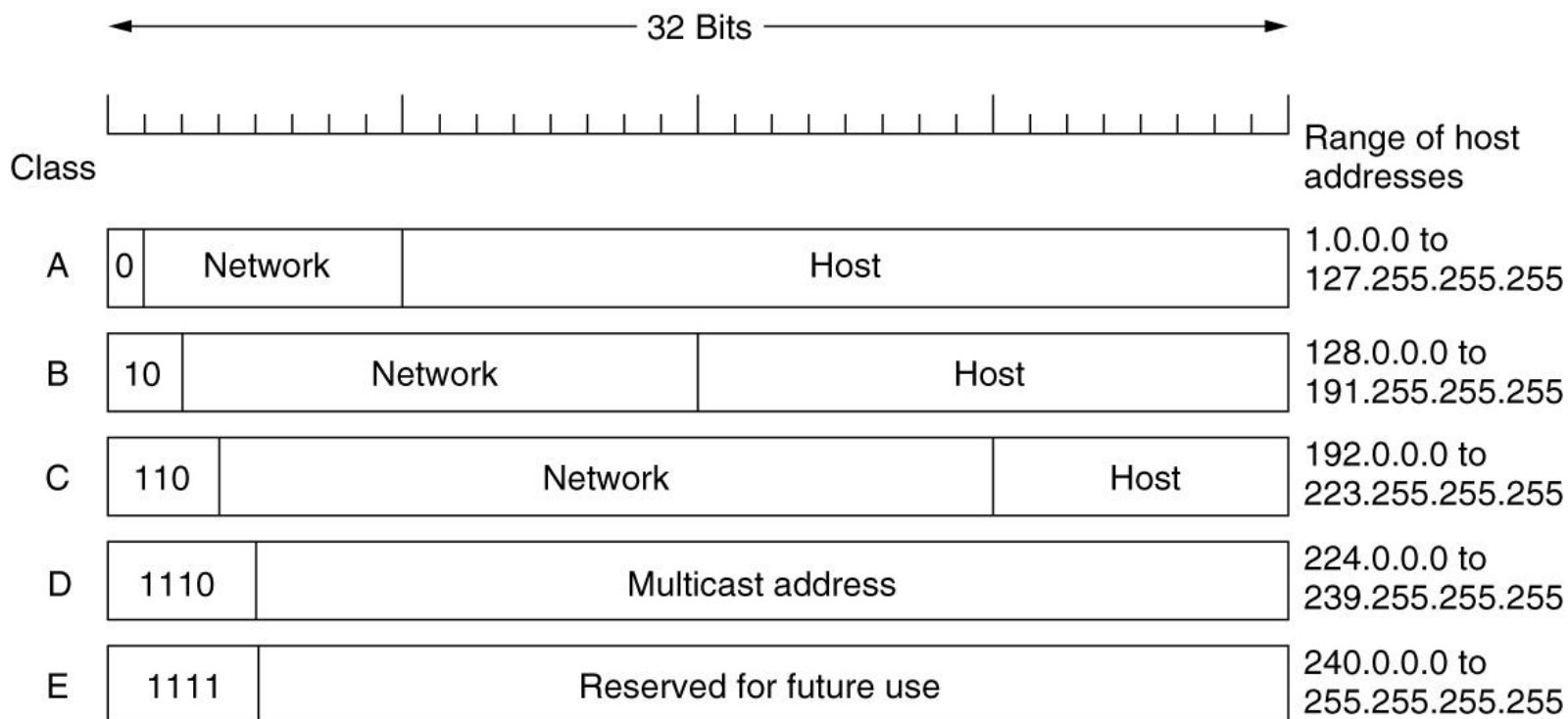


# 7.5 互联网的网络层协议 (12)

- 7.5.2 IP地址

- 分类和特殊寻址，1993

地址组成分类：网络号 + 主机号；



# 7.5 互联网的网络层协议 (13)

- 7.5.2 IP地址

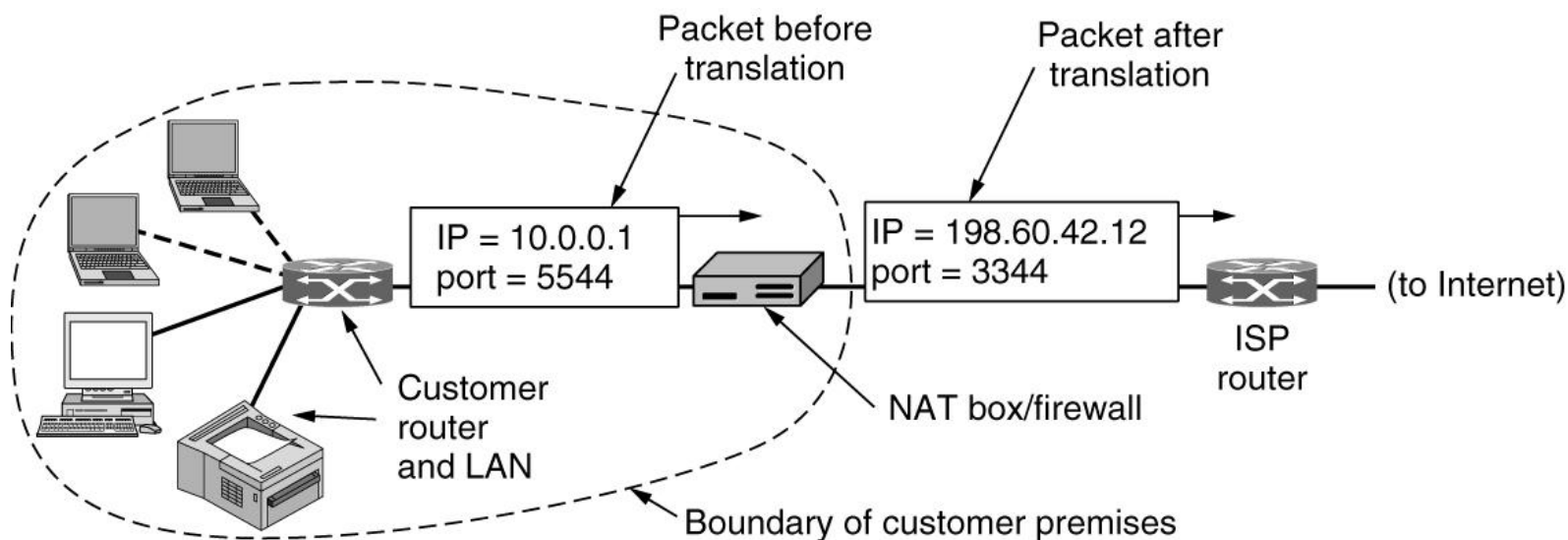
- 分类和特殊寻址

特殊IP地址：全0：表示本网络或本主机，全1：表示广播地址；

0 0																														This host										
0 0										...										0 0										Host										A host on this network
1 1																														Broadcast on the local network										
Network										1 1 1 1										...										1 1 1 1										Broadcast on a distant network
127										(Anything)																				Loopback										

# 7.5 互联网的网络层协议 (14)

- 7.5.2 IP地址
  - 网络地址翻译NAT



# 7.5 互联网的网络层协议 (15)

## 7.5.3 互联网控制协议

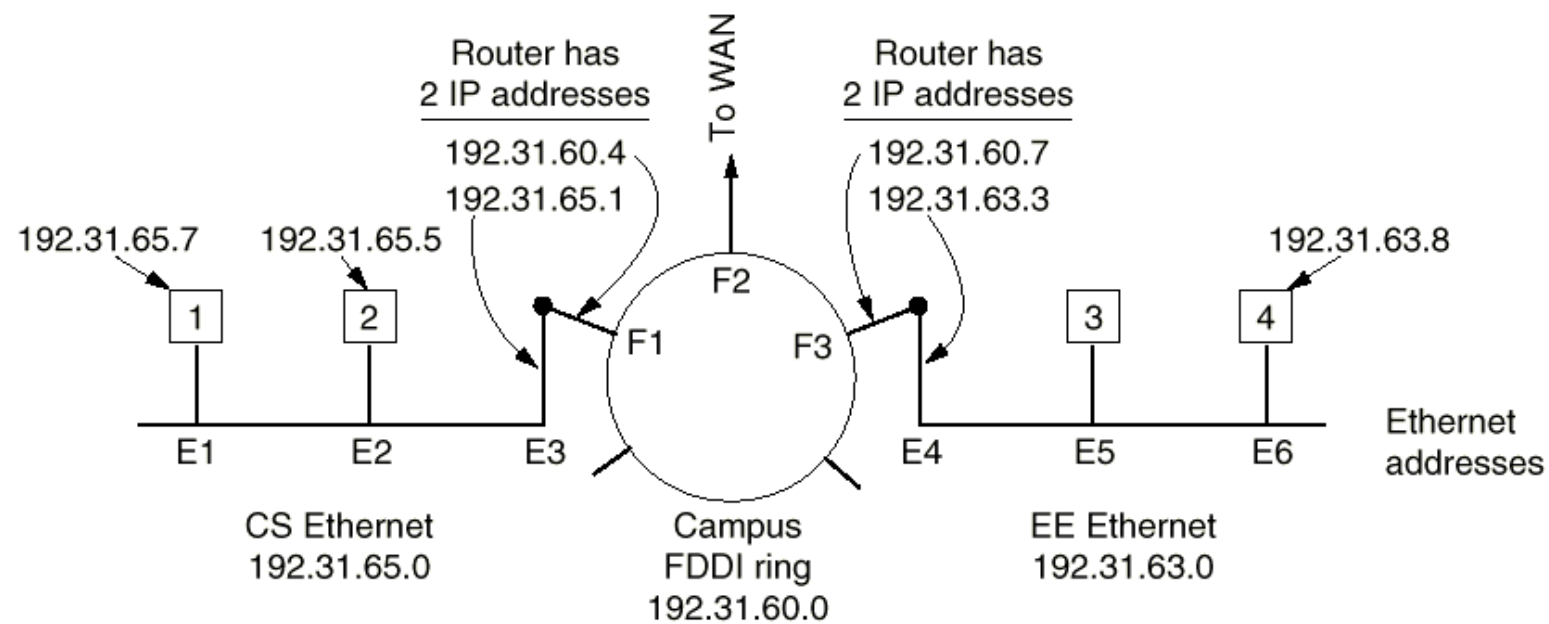
- 互联网控制报文协议ICMP(Internet Control Message Protocol)
  - 主要用来报告出错和测试;
  - 报文类型
  - ICMP报文封装在IP包中。

Message type	Description
Destination unreachable	Packet could not be delivered
Time exceeded	Time to live field hit 0
Parameter problem	Invalid header field
Source quench	Choke packet
Redirect	Teach a router about geography
Echo request	Ask a machine if it is alive
Echo reply	Yes, I am alive
Timestamp request	Same as Echo request, but with timestamp
Timestamp reply	Same as Echo reply, but with timestamp

**Fig. 5-50.** The principal ICMP message types.

## 7.5 互联网的网络层协议 (16)

- 地址解析协议ARP (Address Resolution Protocol)
  - 解决网络层地址 (IP地址) 与数据链路层地址 (MAC地址) 的映射问题;
  - 工作过程
    - 建立一个ARP表,表中存放(IP地址, MAC地址)对;
    - 若目的主机在同一子网内, 用目的IP地址在ARP表中查找, 否则用缺省网关的IP地址在ARP表中查找。若未找到, 则发送广播包, 目的主机收到后给出应答, ARP表增加一项;
    - 每个主机启动时, 广播它的(IP地址, MAC地址)映射;
    - ARP表中的表项有生存期, 超时则删除。



**Fig. 5-51.** Three interconnected class C networks: two Ethernets and an FDDI ring.



## 7.5 互联网的网络层协议 (17)

- 反向地址解析协议RARP (Reverse Address Resolution Protocol)
  - 解决数据链路层地址 (MAC地址) 与网络层地址 (IP地址) 的映射问题;
  - 主要用于无盘工作站启动;
  - 缺点: 由于路由器不转发广播帧, RARP服务器必须与无盘工作站在同一子网内。一种替代协议BOOTP (使用UDP) 。

## 7.5 互联网的网络层协议 (18)

### 7.5.4 路由信息协议RIP ( Routing Information Protocol)

- 内部网关协议IGP (interior gateway protocol) : 自治系统AS内使用的路由算法, RIP、OSPF
- 外部网关协议EGP (exterior gateway protocol) : 自治系统AS之间使用的路由算法, BGP
- RIP
  - 距离向量协议
  - 1982年被包含进BSD-UNIX
  - RIPv1, RFC1058; RIPv2, RFC1723
  - 距离的衡量采用跳数 (max = 15 hops)
  - 距离向量: 每30秒交换一次

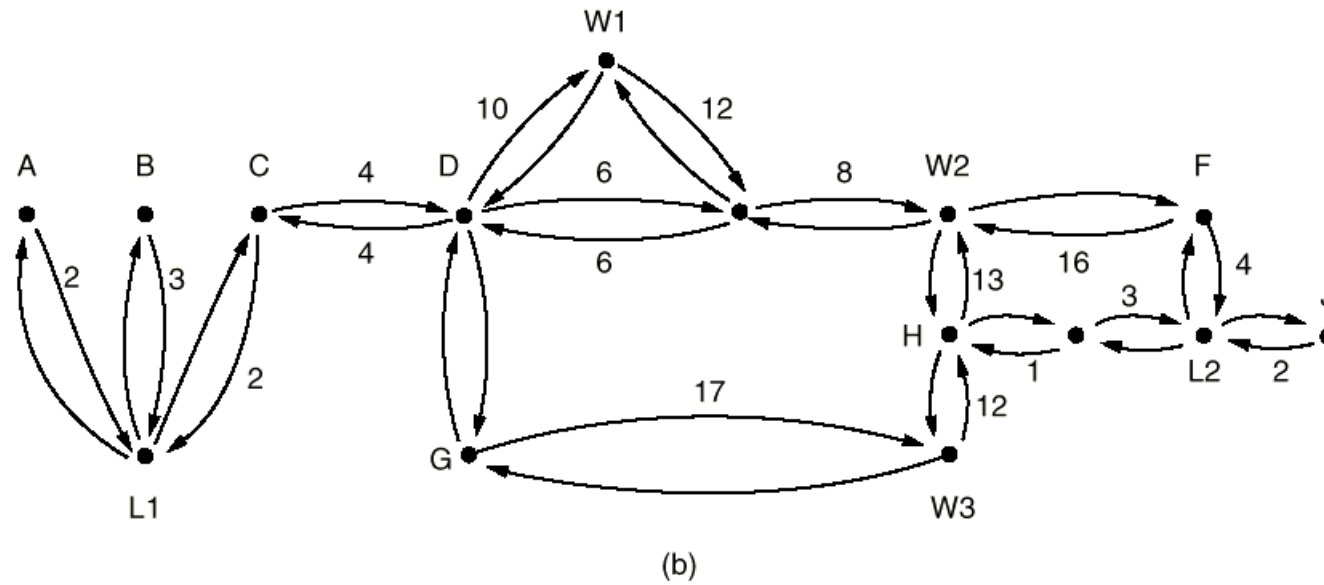
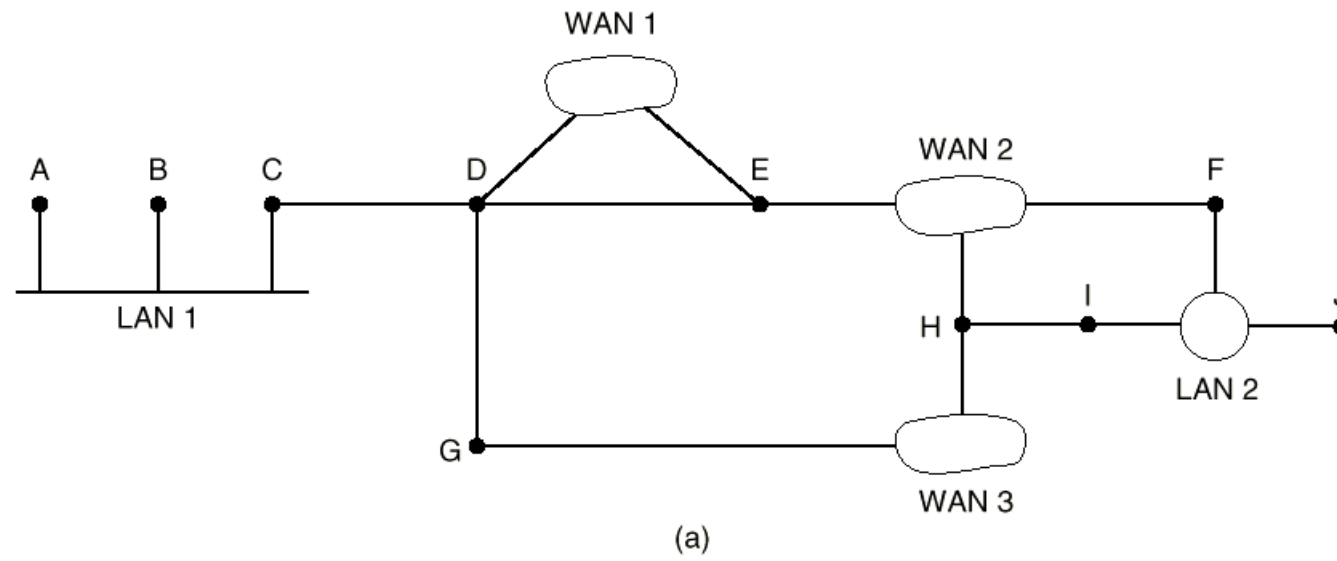
# 7.5 互联网的网络层协议 (19)

## 7.5.5 内部网关路由协议：OSPF

- 开放最短路径优先OSPF (Open Shortest Path First)
  - 开放，公开发表；
  - 支持多种距离衡量尺度，例如，物理距离、延迟等
  - 动态算法；
  - 支持基于服务类型的路由；
  - 负载平衡；
  - 支持分层系统；
  - 适量的安全措施；
  - 支持隧道技术。

## 7.5 互联网的网络层协议 (20)

- 构造有向拓扑图
  - 根据实际的网络、路由器和线路构造有向图；
  - 每个弧赋一个开销值；
  - 两个路由器之间的线路用一对弧来表示，弧权可以不同；
  - 多路访问 (multiaccess) 网络，网络用一个结点表示，每个路由器用一个结点表示，网络结点与路由器结点的弧权为0。



**Fig. 5-52.** (a) An autonomous system. (b) A graph representation of (a).

## 7.5 互联网的网络层协议 (21)

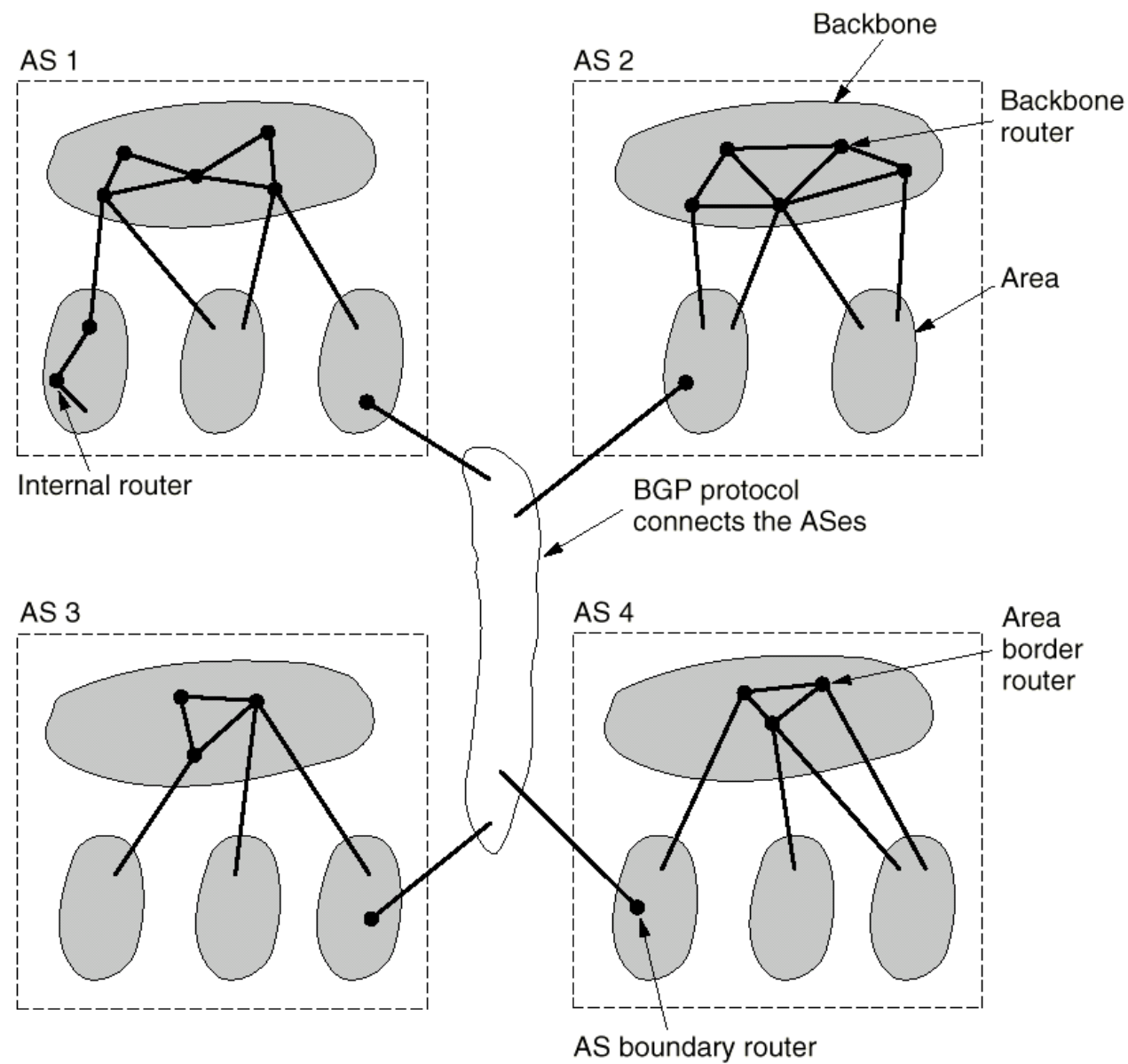
- 分层路由
  - 自治系统AS可以划分区域 (areas) ；
  - 每个AS有一个主干 (backbone) 区域，称为区域0，所有区域与主干区域相连；
  - 一般情况下，有三种路由
    - 区域内
      - 从源路由器到主干区域；
      - 穿越主干区域到达目的区域；
      - 到达目的路由器。
    - 区域间
- 自治系统间

## 7.5 互联网的网络层协议 (22)

- 四类路由器，允许重叠
  - 完全在一个区域内的内部路由器；
  - 连接多个区域的区域边界路由器；
  - 主干路由器；
  - 自治系统边界路由器。
- 信息类型

Message type	Description
Hello	Used to discover who the neighbors are
Link state update	Provides the sender's costs to its neighbors
Link state ack	Acknowledges link state update
Database description	Announces which updates the sender has
Link state request	Requests information from the partner

**Fig. 5-54.** The five types of OSPF messages.



**Fig. 5-53.** The relation between ASes, backbones, and areas in OSPF.



# 7.5 互联网的网络层协议 (23)

## 7.5.6 外部网关路由协议 BGP

— 为什么域间和域内的路由有所不同？

- 策略

- 域间路由跨越不同管理域，要控制流量如何路由

- 域内路由属于同一管理域，不需要定义策略

- 规模

- 分层路由降低了路由表的大小，减小了路由更新的流量

- 性能

- 域内路由：着重于性能

- 域间路由：策略更为重要

## 7.5 互联网的网络层协议 (24)

- 边界网关协议BGP (Border Gateway Protocol)
  - 通过TCP连接传送路由信息;
  - 采用路径向量 (path vector) 算法, 路由信息中记录路径的轨迹
    - 与距离向量协议相似
    - 每个BGP网关向邻居广播所有通往目的地的路径
    - 比如, Gateway X 发送它通往Z的路径:
$$\text{Path}(X,Z) = X,Y1,Y2,Y3,\dots,Z$$

## 7.5 互联网的网络层协议 (25)

- **假设:** 网关X把它的路径发给peer网关W
  - W可以选择是否使用X提供的路径，依据是开销，策略（不使用通过竞争对手网络的路径），防止出现路由循环等
  - 如果W使用X提供的路径
$$\text{Path}(W,Z) = w, \text{Path}(X,Z)$$
  - 注意：X能通过其发往peers的路径广播来控制进入的流量。不想转发到Z的流量，可以通过不通告通往Z的路径来实现

## 7.5 互联网的网络层协议 (26)

### – BGP 消息:

- **OPEN**: 建立与对等方的TCP连接并认证身份
- **UPDATE**: 声明新的路径或撤销旧的路径
- **KEEPALIVE**: 在没有UPDATE消息的时候保持连接有效。也用来回应OPEN request
- **NOTIFICATION**: 报告上一个消息的错误; 也用来关闭连接。

# 7.5 互联网的网络层协议 (27)

## 7.5.7 无类域间路由CIDR (Classless Inter Domain Routing)

- CIDR的提出

- 互联网指数增长, IP地址即将用完
  - IP is rapidly becoming a victim of its own popularity.
- 基于分类的IP地址空间的组织浪费了大量的地址
  - The real villain is the class B network.

- CIDR

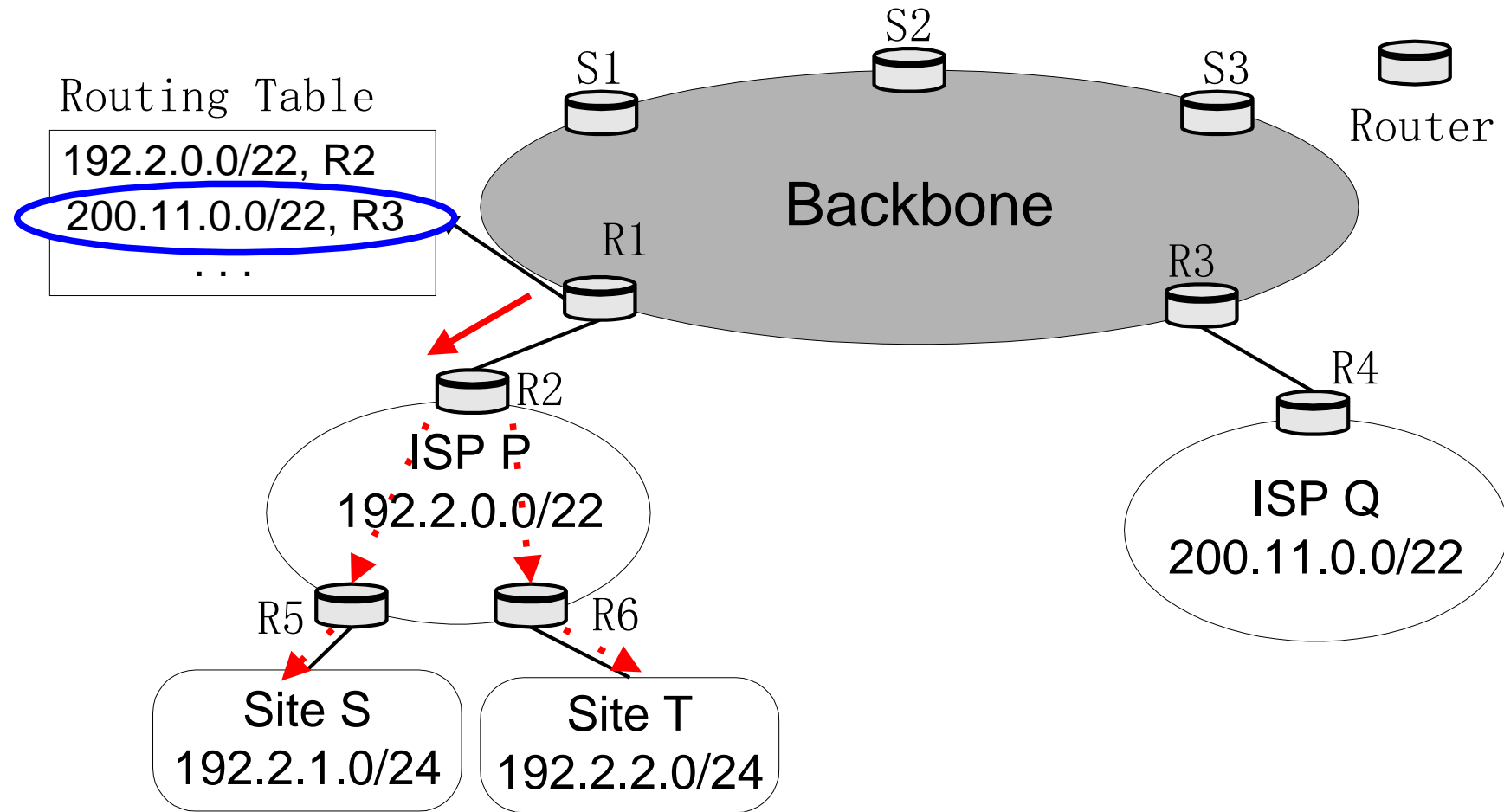
- RFC 1519
- 基本思想: 将剩余的C类地址分成大小可变的地址空间;

## 7.5 互联网的网络层协议 (28)

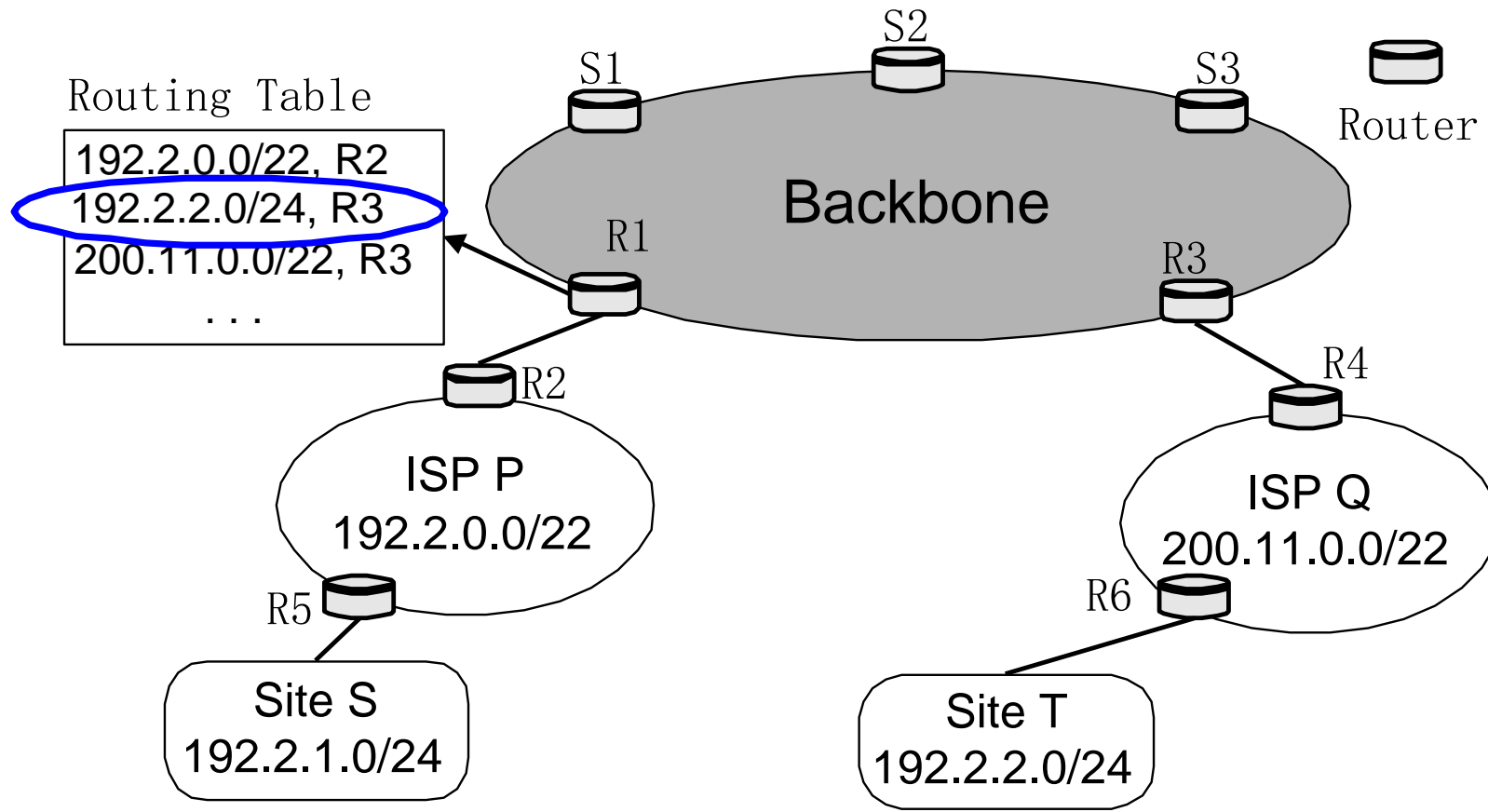
- 例如，需要2000个地址，则分配一个2048个地址（8个C类地址）的连续地址块，而不是一个B类地址；
- RFC 1519 改变了过去C类地址的分配规则，将世界分成4个区，每个区分配一块连续的C类地址空间
  - 欧洲：194.0.0.0 ~ 195.255.255.255
  - 北美：198.0.0.0 ~ 199.255.255.255
  - 中、南美：200.0.0.0 ~ 201.255.255.255
  - 亚太：202.0.0.0 ~ 203.255.255.255
- 路由表中增加一个32位的掩码（mask）域

## 7.5 互联网的网络层协议 (29)

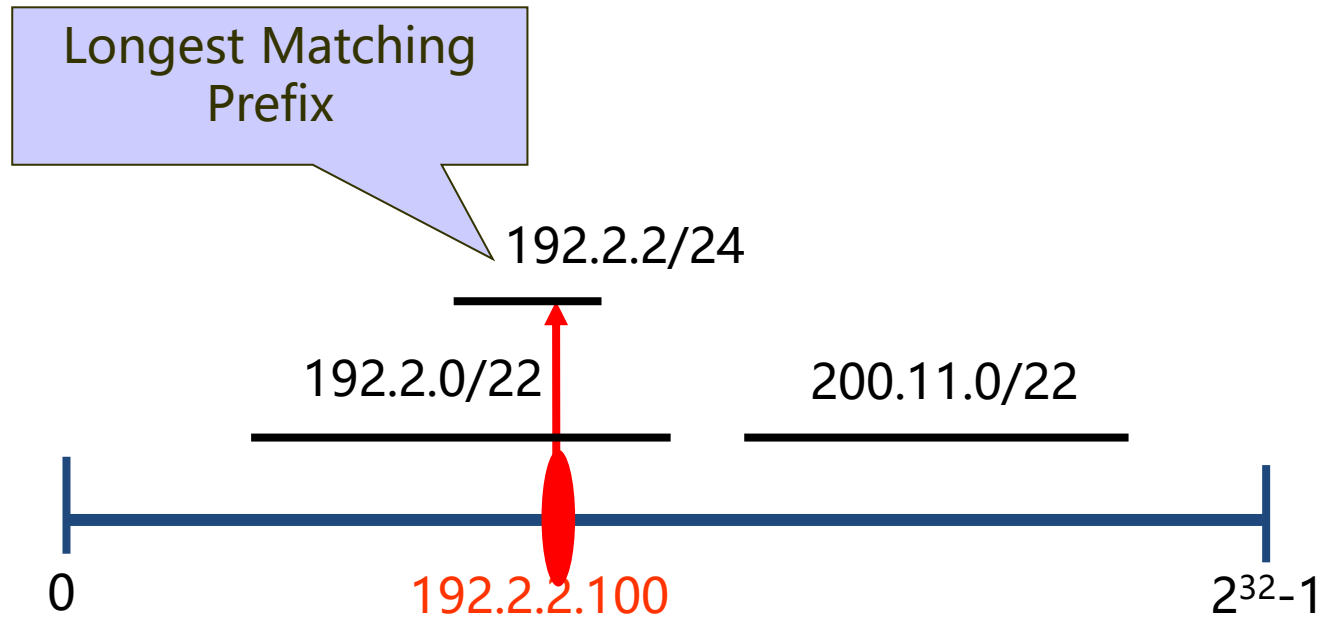
- 最长匹配原则：路由查找时，若多个路由表项匹配成功，选择掩码长（1比特数多）的路由表项；
- CIDR思想可用于所有IP地址，没有A、B、C类之分。
- address format: **a.b.c.d/x**, where x is # bits in network portion of address







# CIDR地址结构下的查找问题



- 路由查找: 在所有前缀中寻找最长匹配

## 7.5 互联网的网络层协议 (30)

- 例

- Cambridge University 需要2048个地址, 194.24.0.0 ~ 194.24.7.255, 掩码255.255.248.0
- Oxford University 需要4096个地址, 194.24.16.0 ~ 194.24.31.255, 掩码255.255.240.0;
- Edinburgh University 需要1024个地址, 194.24.8.0 ~ 194.24.11.255, 掩码255.255.252.0
- 路由表内容

地址	掩码
11000010 00011000 00000000 00000000	11111111 11111111 11111000 00000000
11000010 00011000 00010000 00000000	11111111 11111111 11110000 00000000
11000010 00011000 00001000 00000000	11111111 11111111 11111100 00000000

- 或者表示成

- 194.24.0.0/21, 194.24.16.0/20, 194.24.8.0/22

## 7.5 互联网的网络层协议 (31)

- 一个目的地址为194.24.17.4的包到达路由器，路由表查找过程如下：
  - 194.24.17.4的二进制表示为  
11000010 00011000 00010001 00000100
  - 该地址与第一项Cambridge的掩码做 AND 操作，得到  
11000010 00011000 00010000 00000000  
与Cambridge的地址不同，匹配不成功；
  - 该地址与下一项Oxford的掩码做AND操作，得到  
11000010 00011000 00010000 00000000  
与Oxford的地址相同，匹配成功。
  - 继续匹配，最后选择前缀最长的路由表项

# 7.5 互联网的网络层协议 (32)

## 7.5.8 IPv6协议

### – IPv6的提出

- CIDR仅仅是一种临时补救措施，不能从根本上解决IP地址空间匮乏的问题；
- 移动电话、家电上网需要大量的IP地址；
- 1990年，IETF开始IPv6的工作，收到21个建议；
- 1992年，7个建议被讨论；
- 1993年，3个比较好的建议发表在IEEE Network上，进一步讨论、修改、结合后，形成IPv6；
- IPv6与IPv4不兼容，但与其它Internet协议兼容，如TCP、UDP、OSPF、BGP、DNS等，但实际上还是需要开发另外一套协议栈。

# 7.5 互联网的网络层协议 (33)

## — IPv6的目标

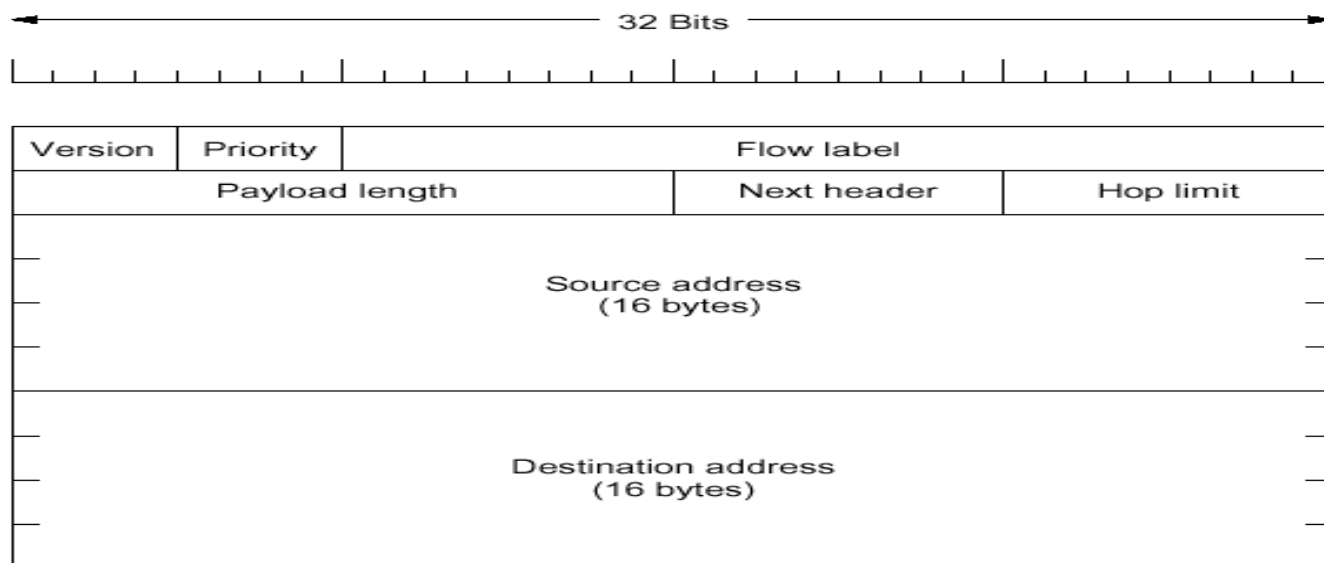
- 即使在不能有效分配地址空间的情况下，也能支持数十亿的主机；
- 减少路由表的大小；
- 简化协议，使得路由器能够更快的处理包；
- 提供比IPv4更好的安全性；
- 更多的关注服务类型，特别是实时数据；
- 支持Multicast；
- 支持移动功能；
- 协议具有很好的可扩展性；
- 增强安全性；
- 在一段时间内，允许IPv4与IPv6共存。

## 7.5 互联网的网络层协议 (34)

- 与IPv4相比，IPv6的主要变化
  - 地址变长，由32位变成128位；
  - IP头简化，由13个域减少为7个域，提高路由器处理速度
    - 由于IPv6包头定长，取消IHL域；
    - Protocol域取消，用Next header域表示；
    - 取消与分段有关的域，IPv6采用不同的分段方法：所有主机和路由器必须支持576字节的包，当主机发送一个大包时，路由器不做分段，而是给主机发一个错误信息，由主机做分段；
    - 取消Checksum域。
  - 更好的支持选项功能；
  - 安全性提高；
  - 更注重服务类型。

## 7.5 互联网的网络层协议 (35)

- IPv6包头



**Fig. 5-56.** The IPv6 fixed header (required).

- Version, 值为6;
- Priority, 用来区分源端可以流控或不能流控的包, 值0 ~ 7表示发生拥塞时源端可以降速, 值8 ~ 15表示发送速率固定的实时负载, 值越小优先级越低;



## 7.5 互联网的网络层协议 (36)

- Flow label, 用来允许源和目的建立一条具有特殊属性和需求的伪连接;
- Payload length, 用来指示IP包中40字节包头后面部分的长度, 与IPv4的total length域不同;
- Next header, 指示扩展包头, 若是最后一个包头, 则指示传输协议类型 (TCP/UDP) ;
- Hop limit, IP包的生存时间;
- Source address, destination address, 16字节定长地址

# 7.5 互联网的网络层协议 (37)

- IPv6地址表示

- 16字节地址表示成用冒号 (:) 隔开的8组，每组4个16进制位，例如，

8000:0000:0000:0000:0123:4567:89AB:CDEF

- 由于有很多“0”，有三种优化表示
    - 打头的“0”可以省略，0123可以写成123；
    - 一组或多组16个“0”可以被一对冒号替代，但是一对冒号只能出现一次。上面的地址可以表示成  
8000::123: 4567:89AB:CDEF
    - IPv4地址可以写成一对冒号和用“.”分隔的十进制数，  
例如::192.31.20.46

## 7.5 互联网的网络层协议 (38)

- 扩展包头
  - 目前定义了六种类型的扩展包头
  - hop-by-hop header，用来指示路径上所有路由器必须检查的信息；
  - routing header，列出路径上必须要经过的路由器
  - fragmentation header，与IPv4相似，扩展头中包括IP包标识号、分段号和判断是否还有分段的位，只有源主机可以分段；

<b>Extension header</b>	<b>Description</b>
Hop-by-hop options	Miscellaneous information for routers
Routing	Full or partial route to follow
Fragmentation	Management of datagram fragments
Authentication	Verification of the sender's identity
Encrypted security payload	Information about the encrypted contents
Destination options	Additional information for the destination

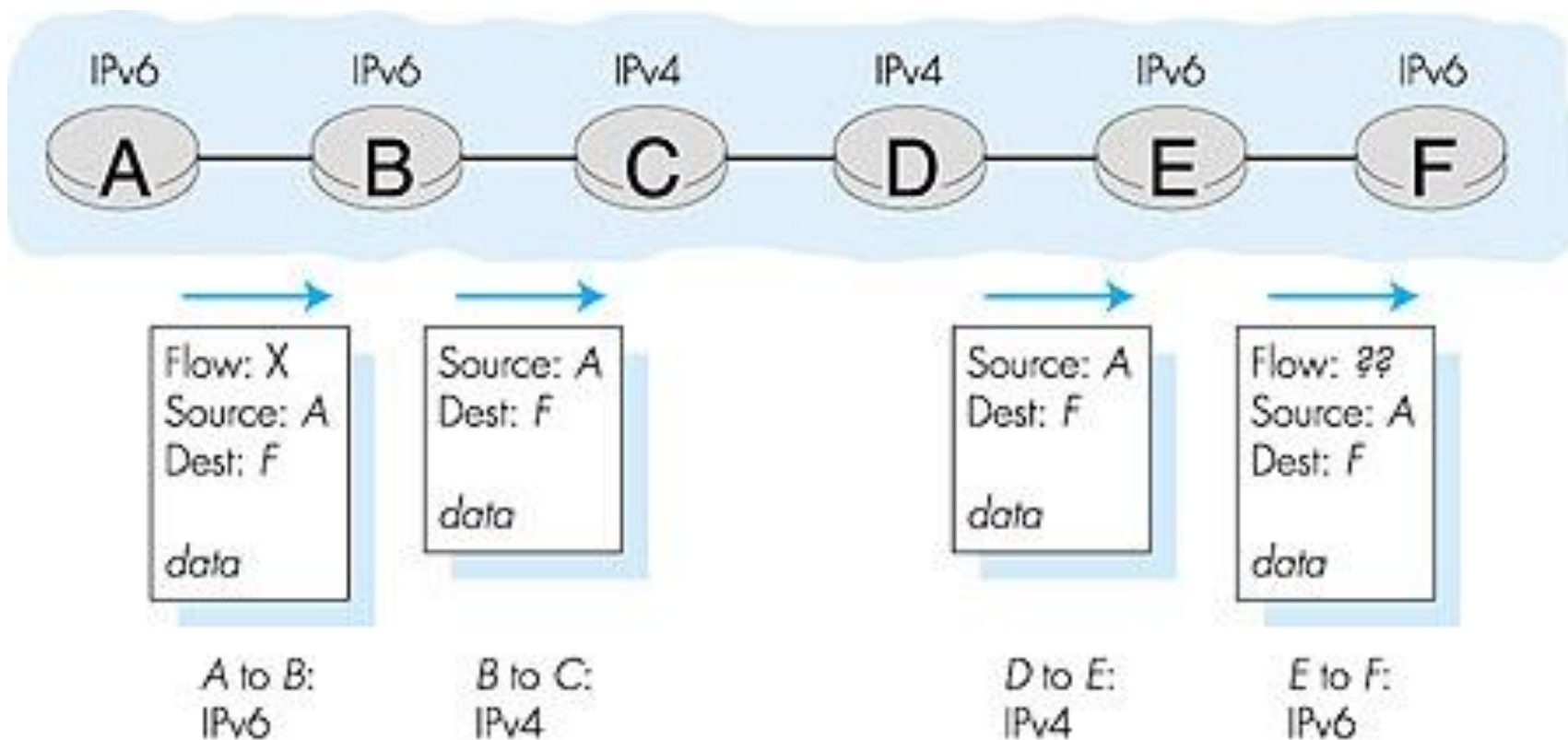
**Fig. 5-58.** IPv6 extension headers.

## 7.5 互联网的网络层协议 (39)

- 从IPv4往IPv6的过渡
  - 不可能所有的路由器同时升级
    - 没有一个确定的期限 (not like Y2K)
    - IPv4和IPv6路由器混合的网络如何操作?
  - 有两个方案
    - 双栈: 有些路由器同时运行两种协议栈, 并可以在这两种之间进行格式转换。
    - 隧道: IPv6的报文被作为IPv4报文的净荷由IPv4路由器处理。

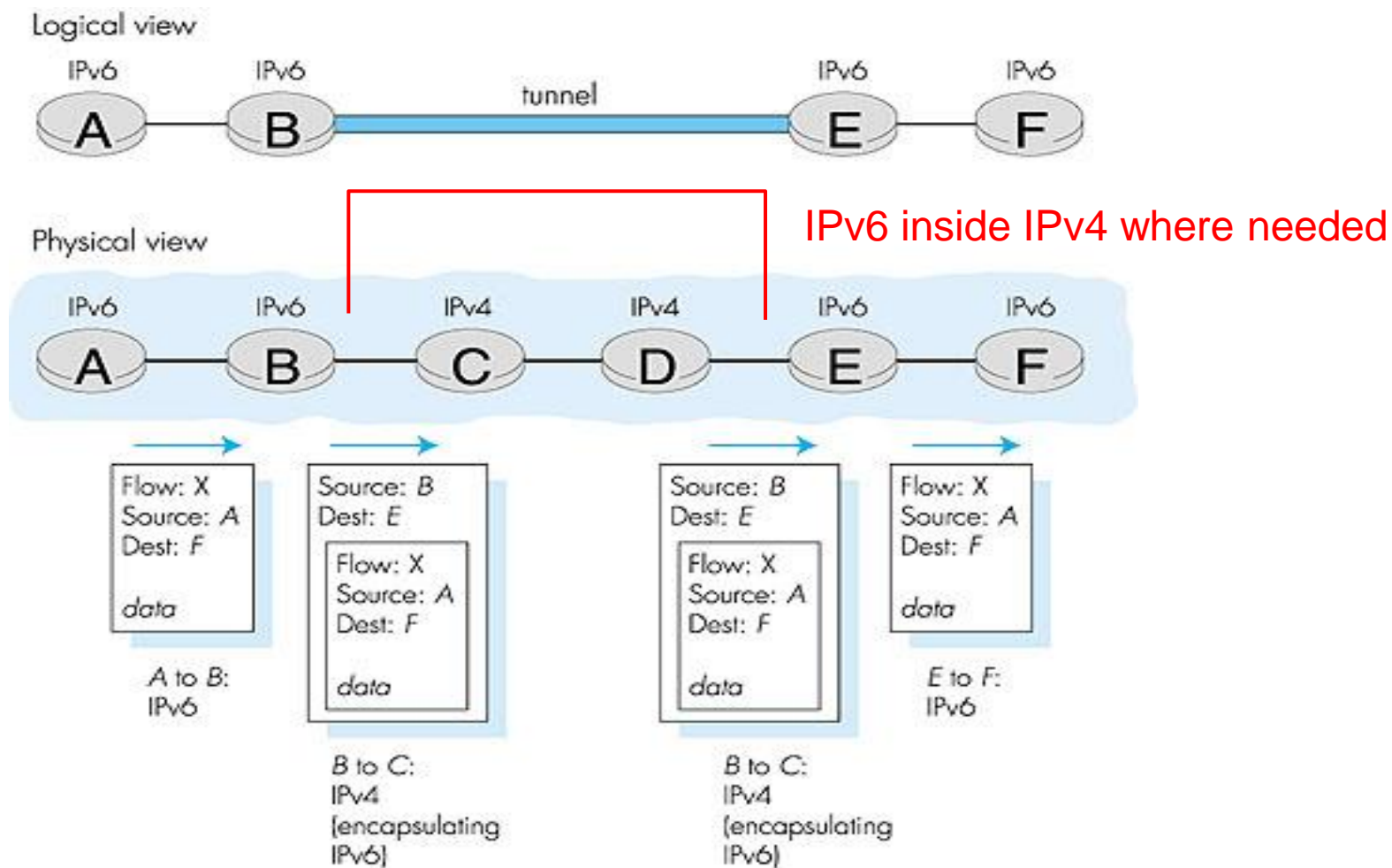
## 7.5 互联网的网络层协议 (40)

- 双栈



## 7.5 互联网的网络层协议 (39)

- 隧道



# 小结 (1)

- IPv4协议
- IPv4地址
  - 前缀，子网，CIDR，分类和特殊寻址
  - 网络地址翻译NAT
- ICMP协议
  - ICMP报文封装在IP包中
  - 地址解析协议ARP
    - ARP协议工作过程
  - 反向地址解析协议RARP



# 小结 (2)

- RIP
  - 内部网关协议
  - 距离向量算法
  - 基于UDP
- OSPF
  - 内部网关协议
  - 链路状态算法
  - 分层思想
  - 四种类型路由器
- BGP
  - 外部网关协议
  - 路径向量算法
  - 基于TCP

# 小结 (3)

- CIDR
  - CIDR的思想
  - IP地址分配和掩码配置
  - 最长匹配原则
- IPv6 协议
  - IPv6与IPv4的主要区别
  - IPv6地址
  - 双栈模式与隧道模式

**第七章 网络层**

**第二部分结束**