



## 第4章 网络层：数据平面

## 本章目标:

### ❖ 理解网络层服务的原理:

- 网络层服务模型
- 路由转发机制
- 路由器工作原理
- IP编址
- NAT
- 因特网中的实例和实现

1

4.1 概述

2

4.2 虚电路和数据报网络

3

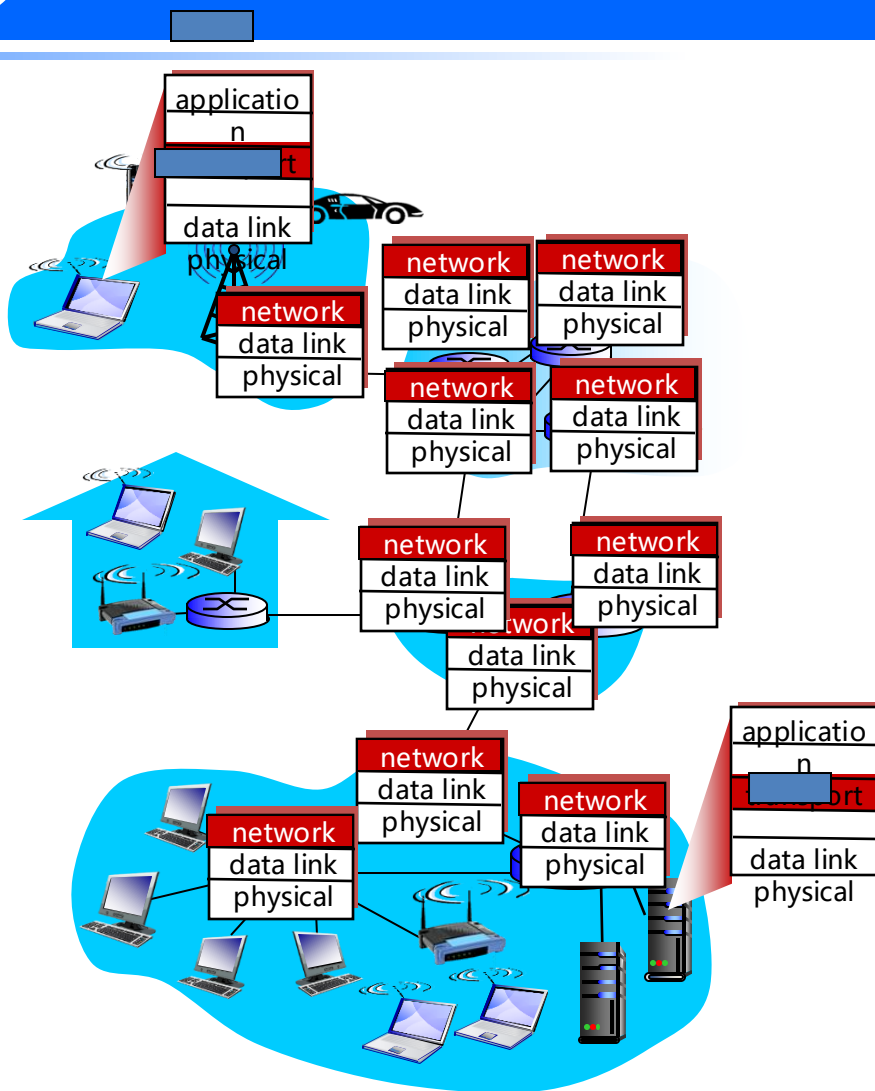
4.3 路由器工作原理

4

4.4 IP协议

# 网络层：概述

- 将**报文段**从发送方传输到接收方
- 在发送方将报文段封装为数据报
- 在接收方，提取报文段给传输层
- 网络层协议运行于每个主机、路由器
- 路由器会检测每个通过它的IP数据报的首部字段

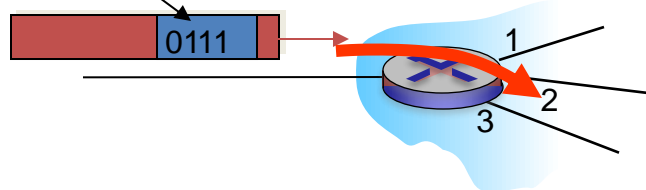


# 网络层： 数据平面 vs 控制平面

## 数据平面

- 主要关注每个路由器
- 转发功能；
- 转发机制：接口到接口

values in arriving  
packet header



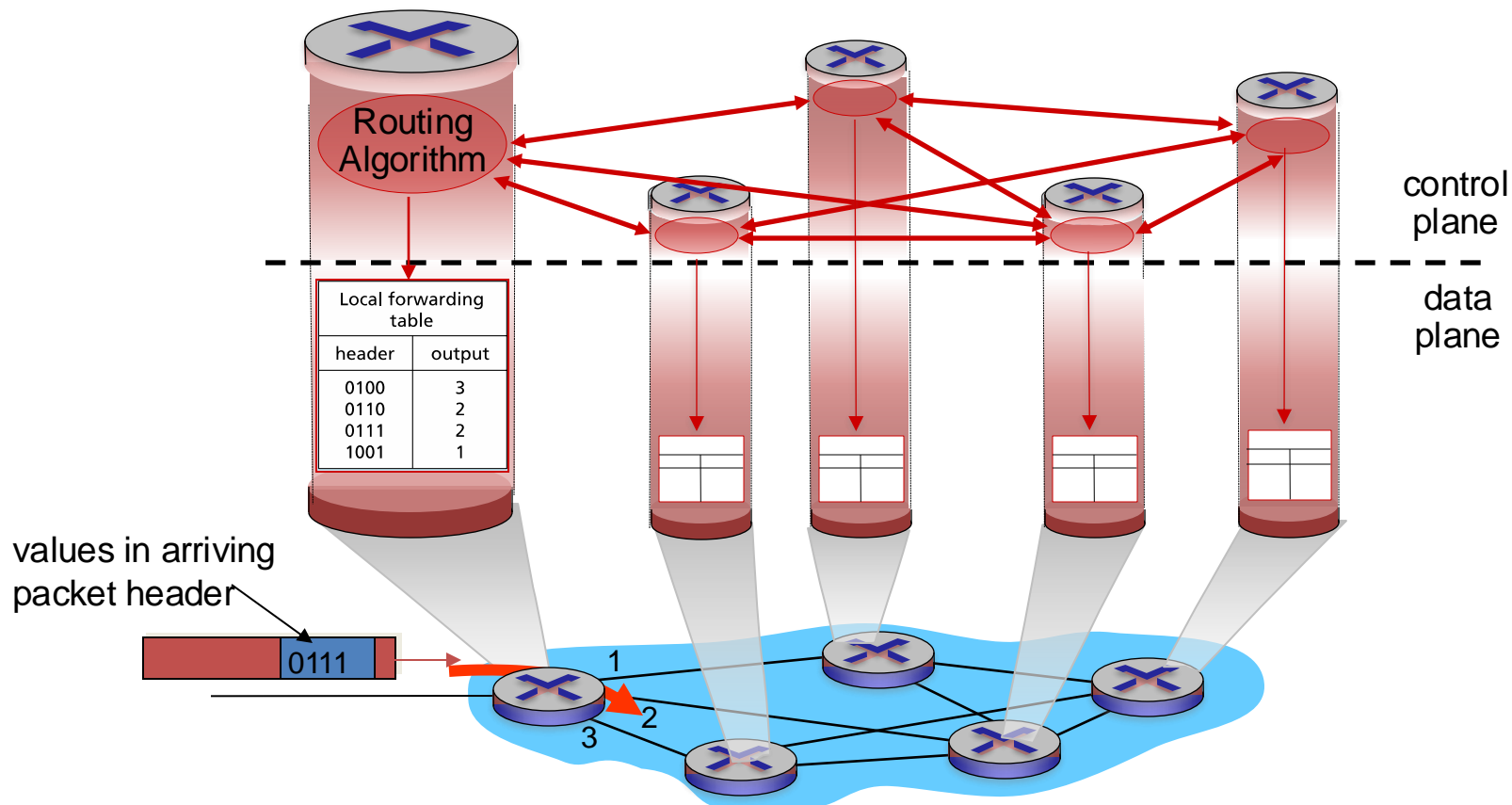
## 控制平面

- 整个网络范围
- 主机之间的逻辑连接
- 传统 vs 趋势:
  - *traditional routing algorithms:*  
implemented in routers
  - *software-defined networking (SDN):* implemented in  
(remote) servers

# 网络层： 数据平面 vs 控制平面

基于单独的路由器的路由算法

*in each and every router* interact in the control plane

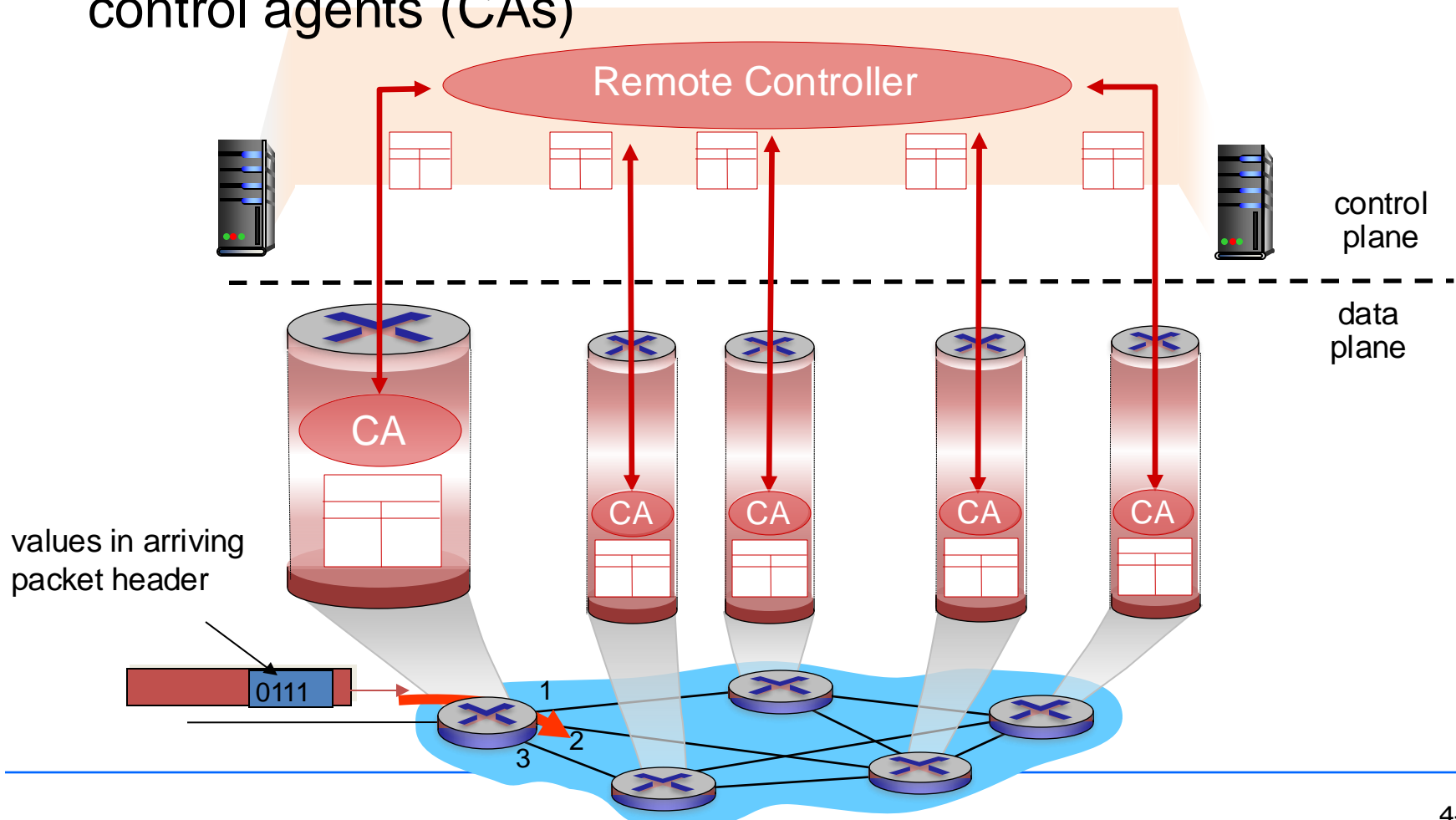


# 网络层： 数据平面 vs 控制平面

软件学院 · 计算机网络

远程，集中式管理：

A distinct (typically remote) controller interacts with local control agents (CAs)



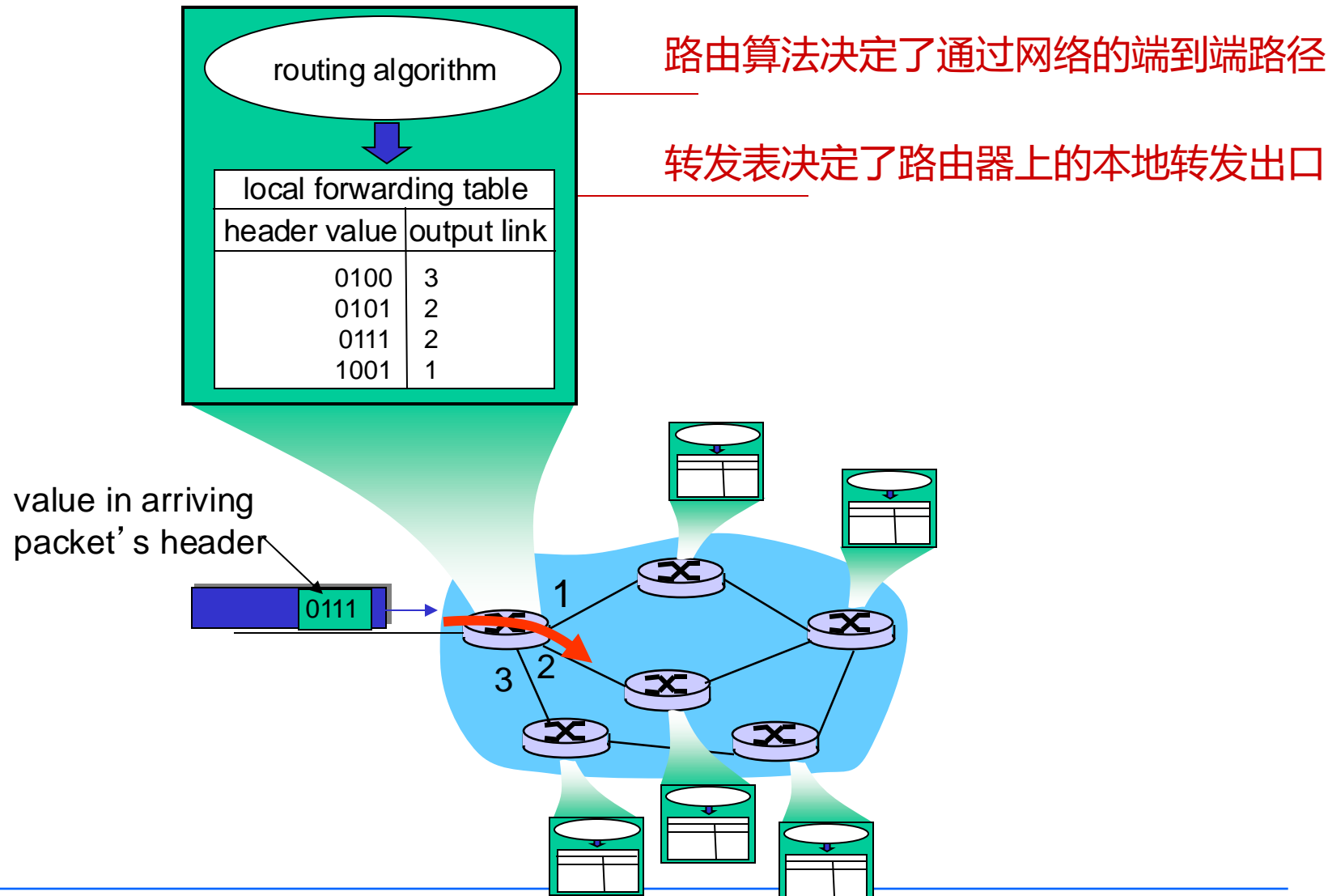
- **转发**: 将进入路由器的数据分组转发到另一个适合的出口
- **路由**: 确定分组从源到目的地所采用的路径
  - 路由算法
  - 路由协议

## 类比:

- ❖ **转发**: 通过单一立交桥的过程
- ❖ **路由**: 从源头到目的路由路径规划过程



# 路由与转发的相互作用



- 对于某些其他体系结构的网络，如ATM，帧中继，X.25，建立连接时是网络层提供的第三个重要功能
- 在数据报流动之前，源与目的端主机，以及所经由的路由器首先会建立一条虚拟连接
  - 路由器参与其中
- 网络与传输层连接服务的对比：
  - **网络层**：在两个端系统之间（在虚电路网络中，所经由的路由器也参与其中）
  - **传输层**：连接建立在两个进程之间

**Q:** 什么样的服务模型适合引导数据报从发送方传输到接收方?

*对于单个的数据报:*

- ❖ 确保交付
- ❖ 确保交付的时延小于40ms

*对于一组数据报流:*

- 数据报有序的交付
- 确保传输的最小带宽
- 确保最大时延抖动, 两个相继分组之间的时间间隔的变化不超过某些特定值

绝大多数网络层服务是  
“尽力而为”

# 网络层服务模型

Network Architecture	Service Model	Guarantees ?				Congestion feedback
		Bandwidth	Loss	Order	Timing	
Internet	best effort	none	no	no	no	no (inferred via loss)
ATM	CBR	恒定速率	yes	yes	yes	no congestion
ATM	VBR	变化速率	yes	yes	yes	no congestion
ATM	ABR	可用速率	no	yes	no	yes
ATM	UBR	不指定速率	no	yes	no	no

ATM: 英文全称为 “asynchronous transfer mode” ,  
中文名为 “异步传输模式” 。

1

4.1 概述

2

4.2 虚电路和数据报网络

3

4.3 路由器工作原理

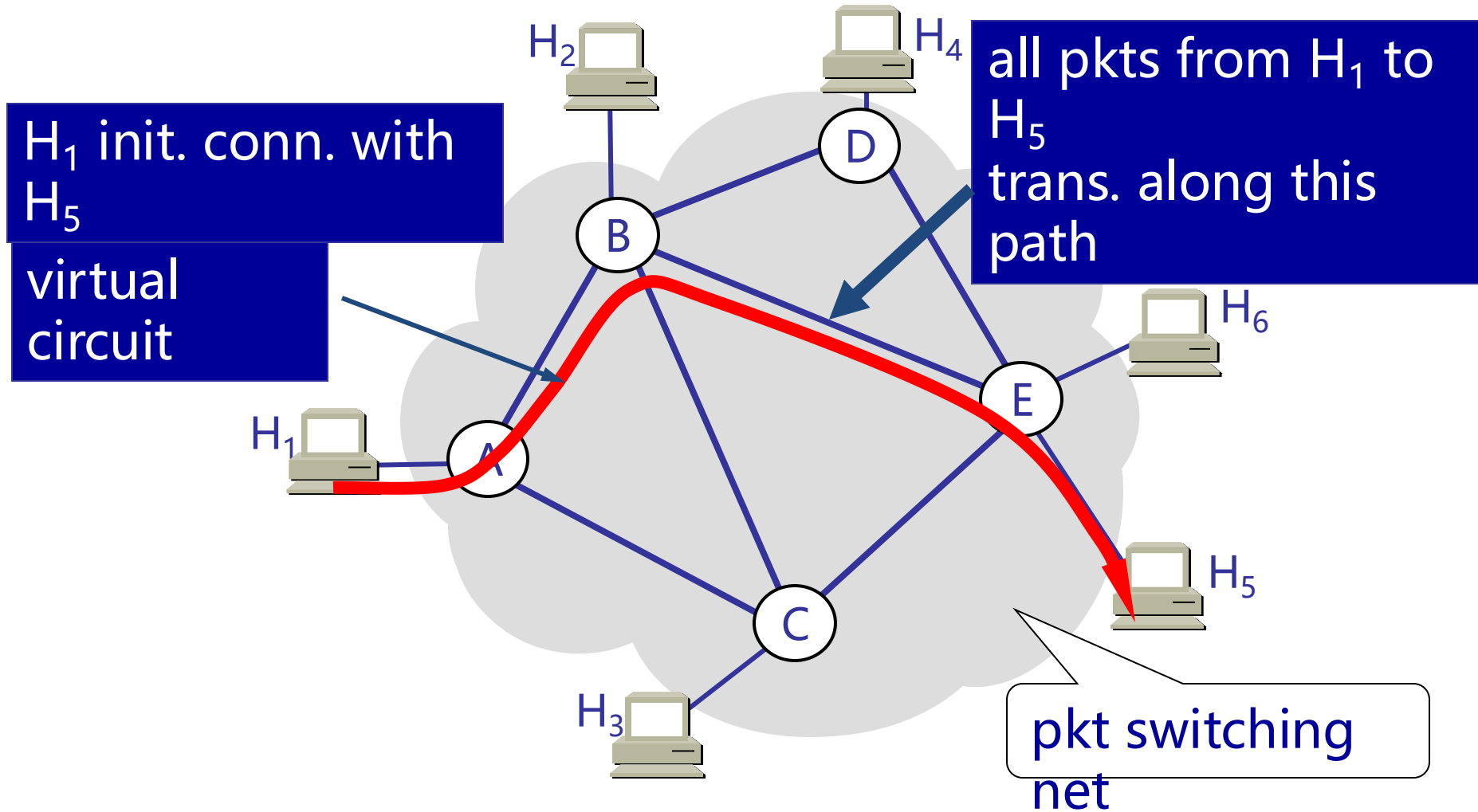
4

4.4 IP协议

- ❖ 数据报网络提供了网络层的无连接服务
- ❖ 虚电路网络提供了网络层的有连接服务
- ❖ 类似传输层TCP/UDP的面向连接/无连接服务，但是
  - 服务：主机到主机
  - 没有选择：每种网络体系结构只提供其中一种连接
  - 实现：端系统，以及位于网络核心的路由器

# 虚拟电路 (VC) : 直观视图

软件学院 · 计算机网络



“使收发双方之间的路径表现得如同电话线路一般”

- 网络内部有较多的功能和性能指标
- 沿收发路径上的网络结点的操作比较复杂

- 在数据发送前，需要建立连接，通信结束后要断开
- 每个分组携带 VC标识 (而不是信宿主机的ID)
- 每个在收发双方路径上的路由器需要为正在传输中的连接维持 “状态”  
传输层的连接仅涉及到两个端系统 (end system)
- 链路, 路由器资源 (带宽, 缓存等)可被分配 给 VC; 目的: 为了达到类似电路交换的性能

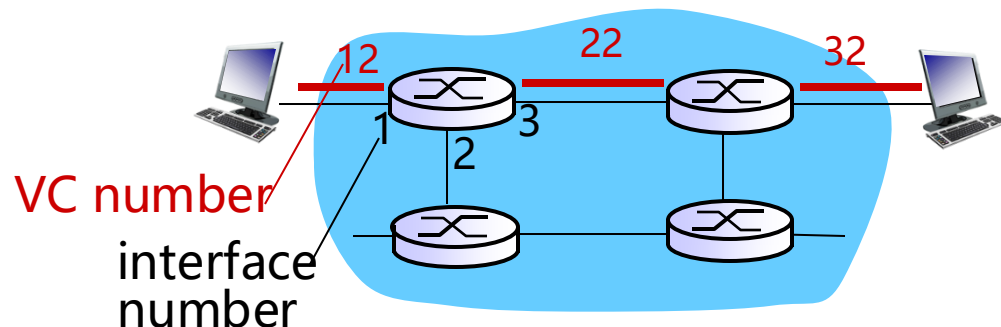


一个虚电路包含：

1. 从源到目的端的**路径**
2. **VC号**，沿着该路径的每段链路的一个号码
3. 沿着该路径的每台路由器的**转发表表项**

- ❖ VC的分组在它的首部携带一个VC号，而不是目的端地址
- ❖ 一条虚电路在每个链路上具有不同的VC号，新的VC号来自转发表

*forwarding table in  
northwest router:*

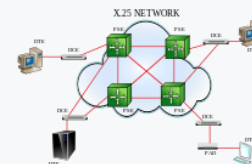


Incoming interface	Incoming VC #	Outgoing interface	Outgoing VC #
1	12	3	22
2	63	1	18
3	7	2	17
1	97	3	87
...	...	...	...

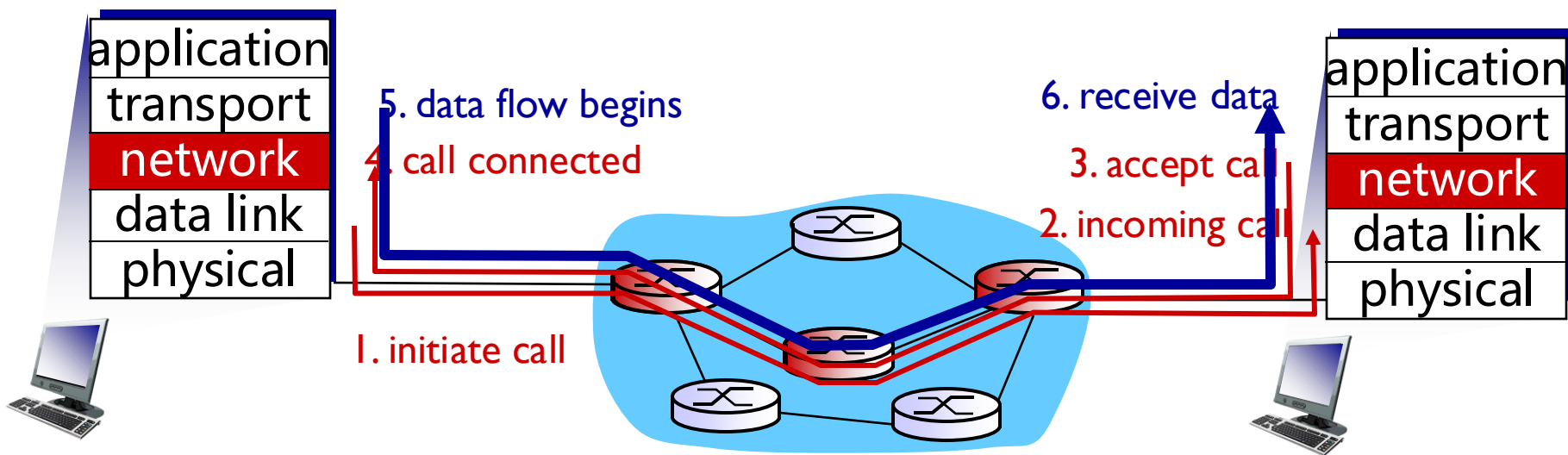
**VC 路由器维护连接状态信息!**

- 用来建立、维护、断开 VC
- 应用在 ATM, 帧中继, X.25 (电信级服务)
- 不应用于当今的 Internet

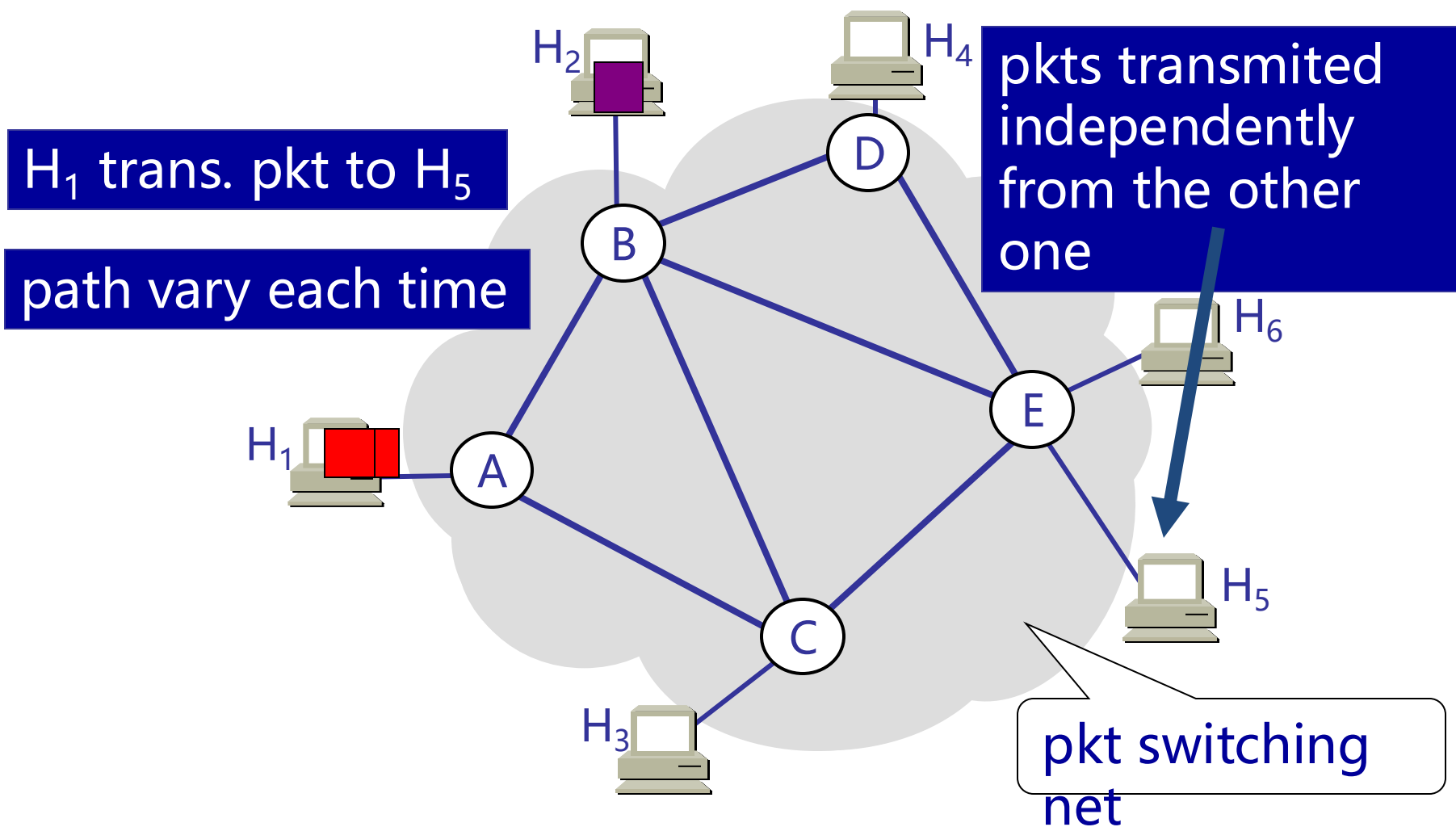
电话或者ISDN设备作为网络硬件设备来架构广域网的ITU-T网络协议



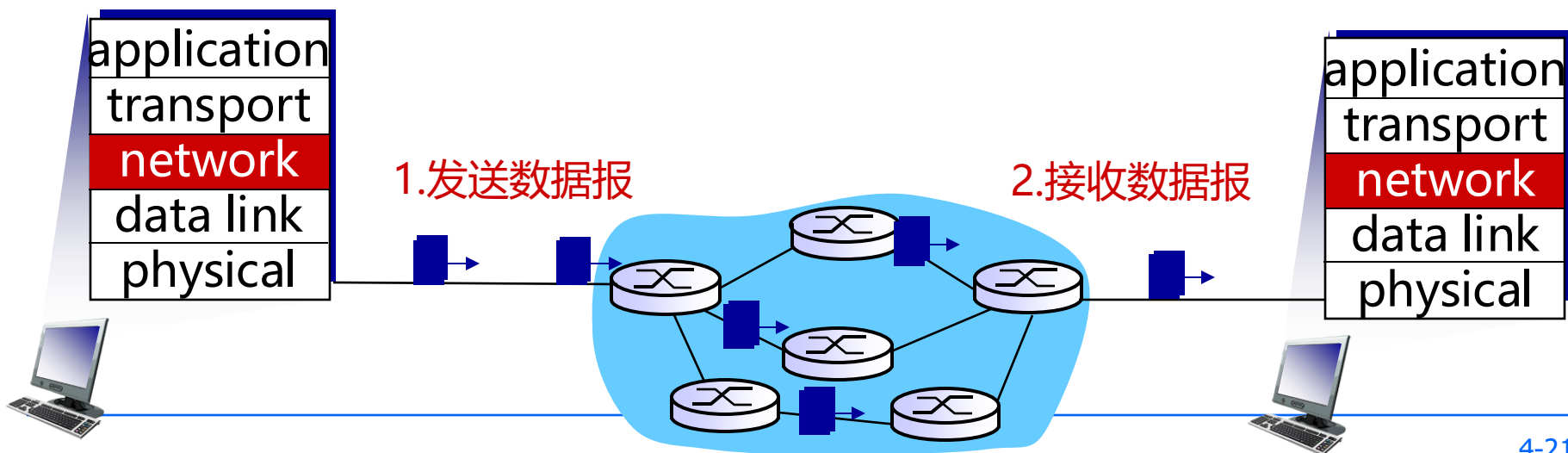
状态 已生效  
开始年 1976  
最新版本 (10/96)  
October 1996  
组织 ITU-T  
委员会 Study Group VII  
有关标准 X.400  
网站 <https://www.itu.int/rec/T-REC-X.25/>



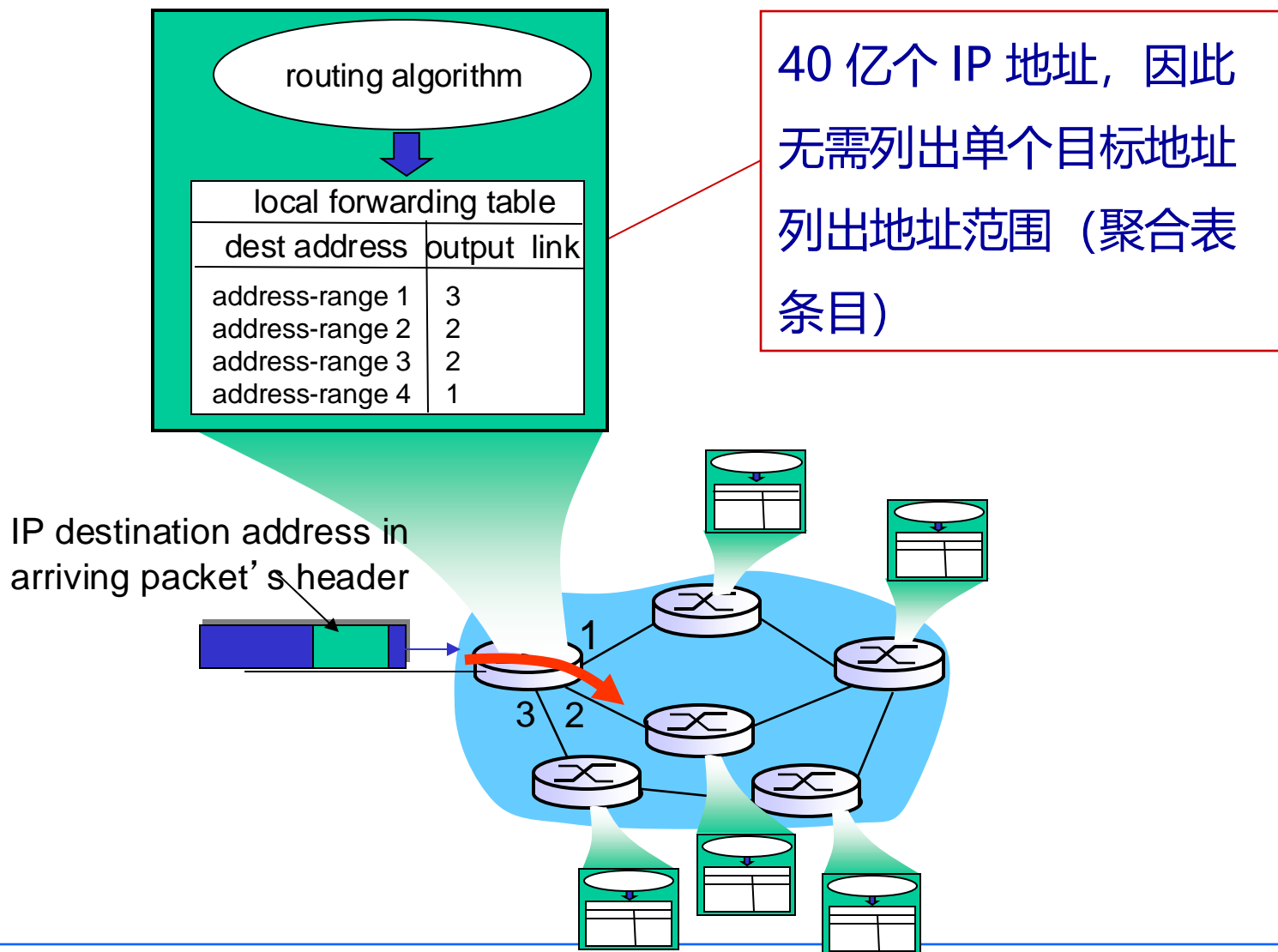
# 数据报：直观的视图



- 在网络层没有连接建立过程
- 路由器: 没有端对端的连接状态, 在网络层不存在“连接”的概念
- 分组使用目的主机的地址进行路由选择, 因此收发双方的不同分组所经由的路径可能不同



# 数据报路由转发表



## 最长前缀匹配

查找给定目标地址的转发表条目时，请使用与目标地址匹配的最长地址前缀。

Destination Address Range	Link interface
11001000 00010111 00010*** *****	0
11001000 00010111 00011000 *****	1
11001000 00010111 00011*** *****	2
otherwise	3

examples:

DA: 11001000 00010111 00010110 10100001    which interface?

DA: 11001000 00010111 00011000 10101010    which interface?

## Internet (数据报网络)

- 来源于计算机之间的数据交换的需求
  - “弹性” 服务，没有严格的时序要求
- 可利用各种不同类型的链路层技术
  - 具有不同的特性
  - 很难统一其服务标准
- “智能” 终端系统 (计算机)
  - 自适应，性能控制，容错
  - 网络内部架构简单，把复杂性留给了网络边缘

## ATM (VC)

- 源自于电话网络
- 人的会话交互：
  - 对实时性、可靠性有较高要求
  - 需要性能保障服务
- “dumb” 设备
- telephones
  - 电话
  - 网络内部架构复杂



1

4.1 概述

2

4.2 虚电路和数据报网络

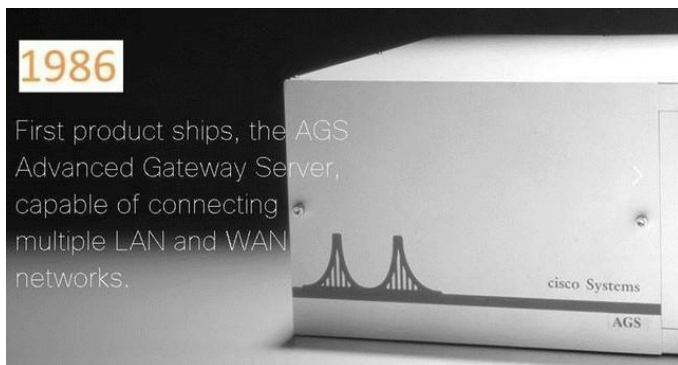
3

4.3 路由器工作原理

4

4.4 IP协议

# 路由器是什么样的？



思科 AGS – 全球第一台路由



您每天看到的路由器



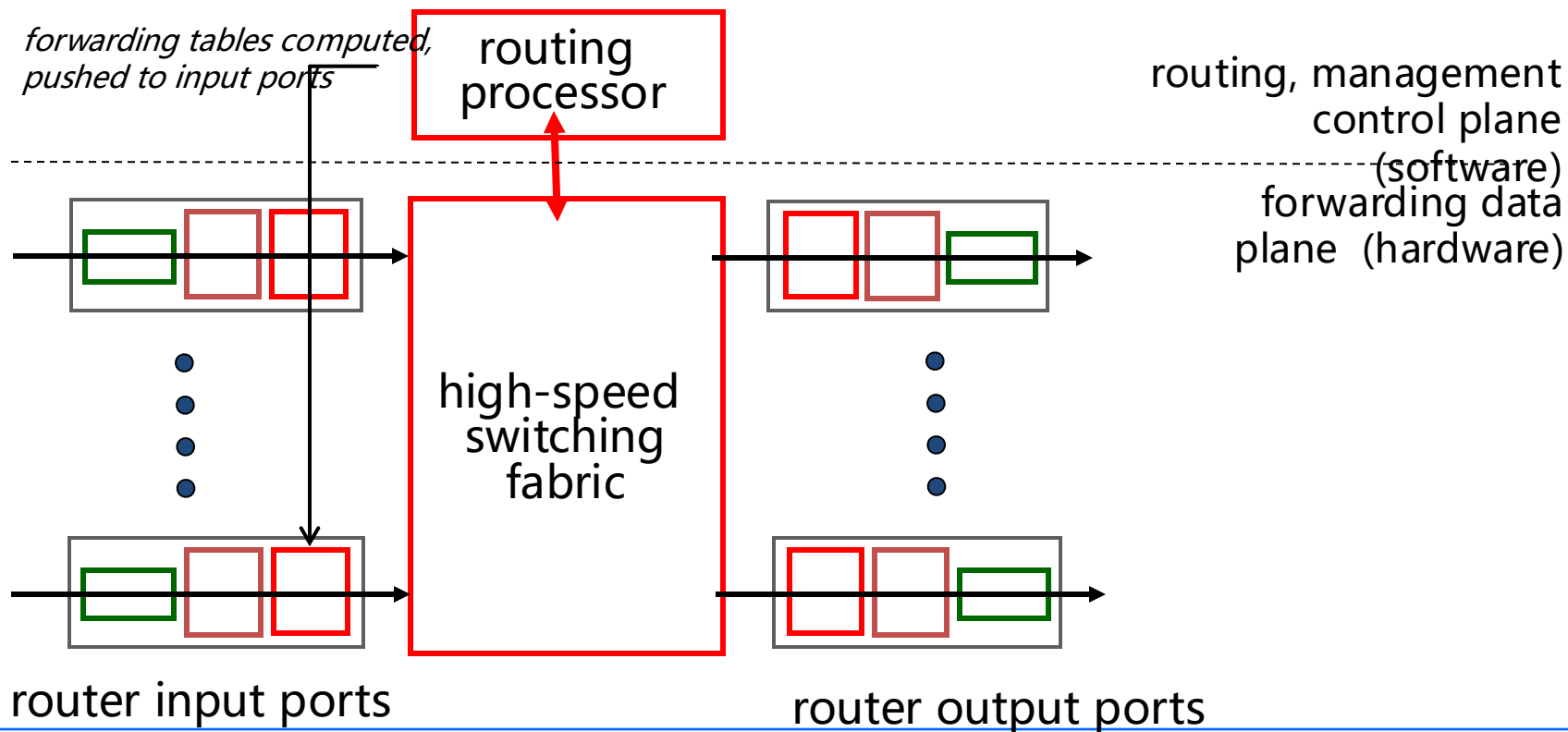
华为网络引擎8000系列路由器

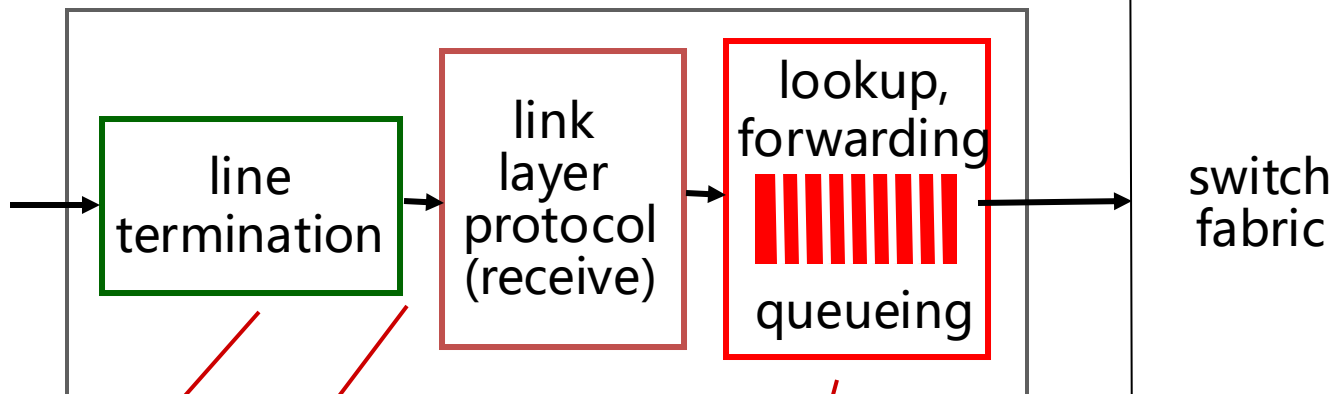


华为7900系列商用交换机

两个主要路由器功能:

- ❖ 运行路由算法/协议 (RIP, OSPF, BGP)
- ❖ 将数据报从入口链路转发到出口链路





物理层：接收物理信号  
，并转换为bit

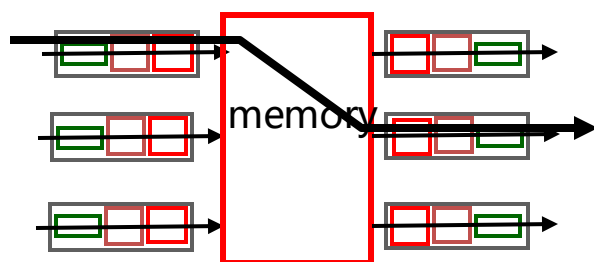
数据链路层：如以太网

网络层：查找，转发，  
排队

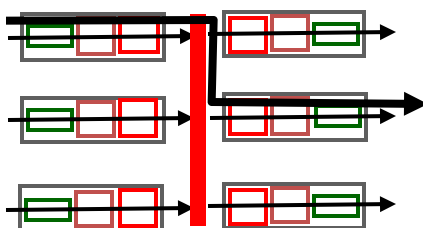
非集中式的交换：转发表的副本已发送给每个  
输入端口，因此，转发决策由输入端口在本地完成

- 根据数据报的目的地址，在输入端口内存的转发表中查找输出端口
- 目标：以线缆一级的速度完成输入端口处理
- 队列：如果数据报到达的速率大于转发处理的速率，则需要排队等待

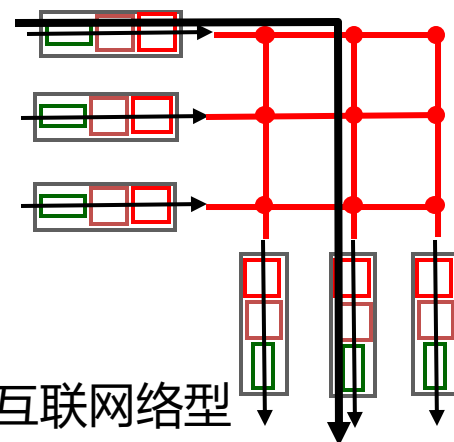
- ❖ 将分组从输入端口转发到适合的输出端口
- ❖ 交换速率：数据包从输入传输到输出的速率
  - $N$  输入：开关速率  $N$  倍所需线速
- ❖ 三种类型的交换结构



内存型



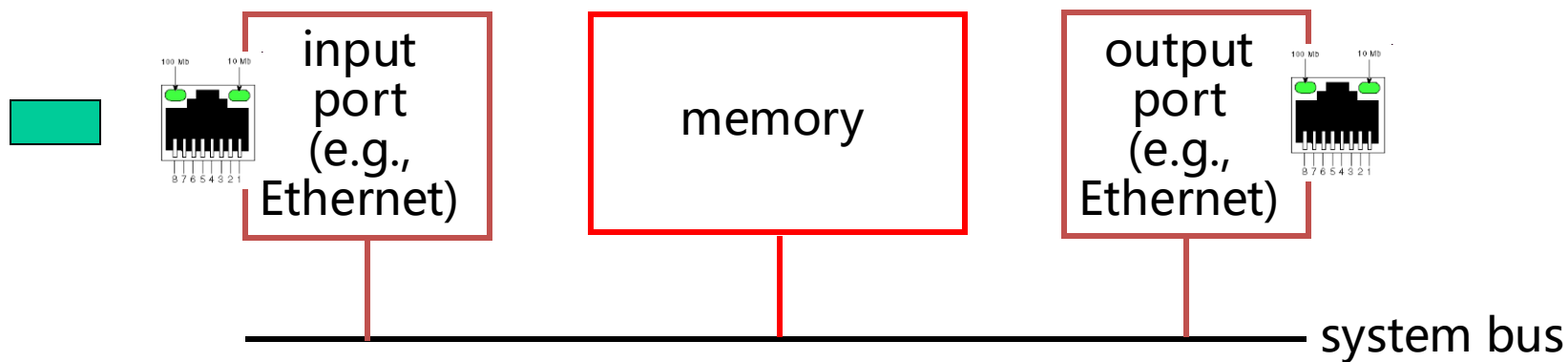
总线型



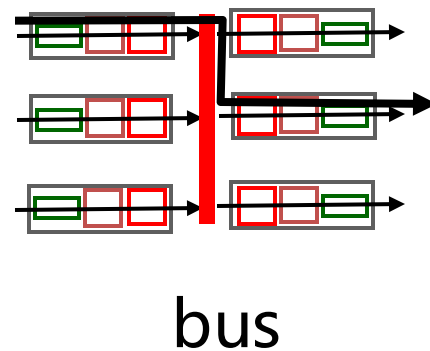
互连网络型

## 第一代路由器:

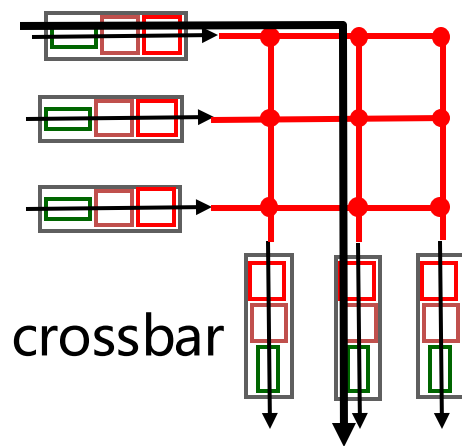
- 传统的计算机，由CPU直接控制完成交换
- 数据分组需要拷贝到系统的内存中
- 交换速度受限于内存带宽



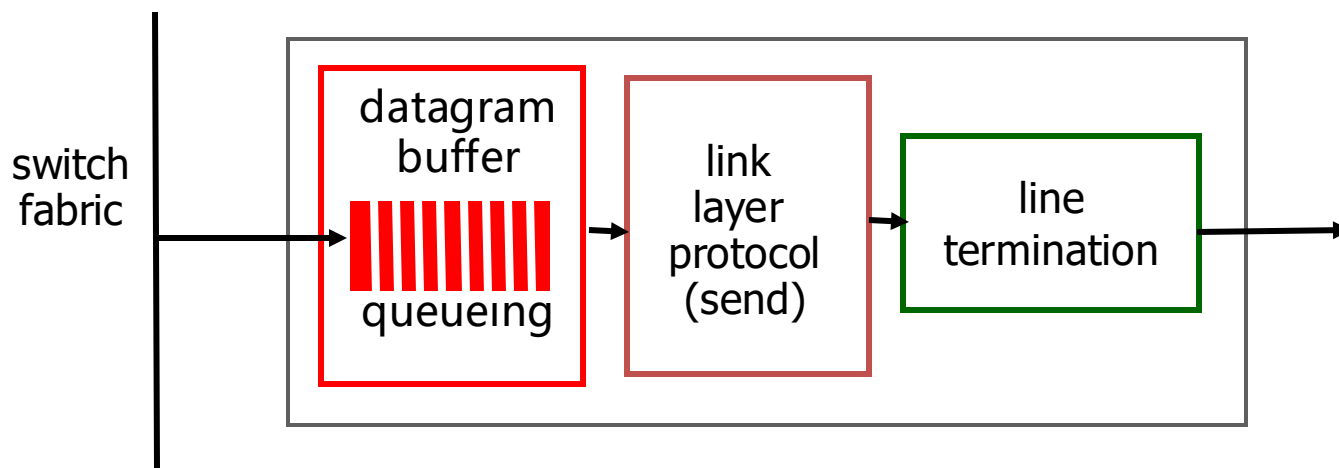
- ❖ 分组通过共享总线直接从输入端口的内存转发到输出端口的内存
- ❖ **总线竞争**（一次只有一个分组能够使用总线）：路由交换速度受限于总线带宽
- ❖ 32Gbps总线, Cisco5600：对于小型接入网和企业网其交换速度通常是足够的



- ❖ 克服了总线带宽的限制
- ❖ 早期发展的Banyan网络（ATM中一种交换串联网络），纵横式交换网络，以及其他一些互联网络，是在多处理器计算机体系结构中将处理器之间互联起来
- ❖ 高级设计：将数据报分割为固定长度的单元，每个单元通过交换结构完成交换
- ❖ Cisco12000：通过互联网络其交换带宽达到60Gbps





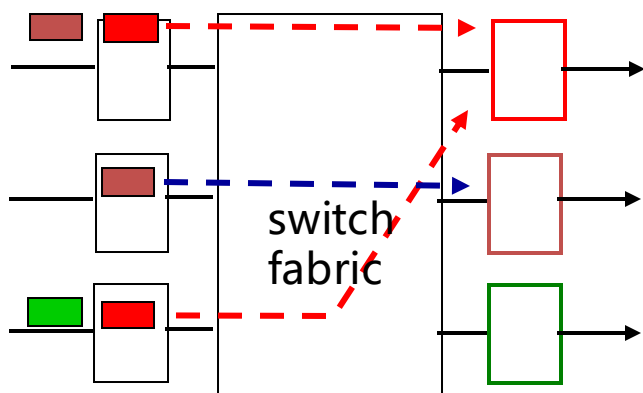


- ❖ **缓存区**：当交换结构速率较快时，就需要缓存，否则，在遇到拥塞时，会导致分组丢失
- ❖ **分组调度**：先来先服务（FCFS）规则，加权公平排队规则（WFQ），遵循“网络中立”原则，为服务质量保障（QoS）起到关键作用

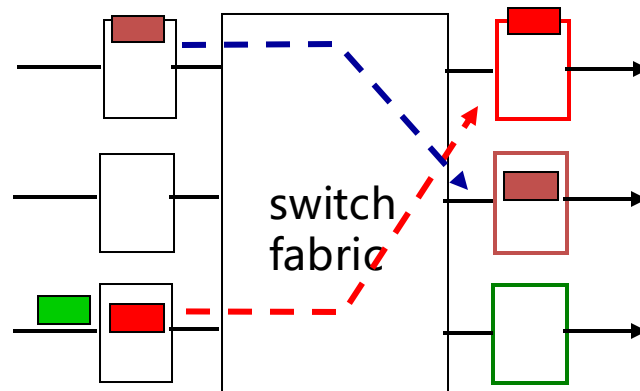
优先级调度 – 谁获得最佳性能，网络中立性

数据报（数据包）可能会因拥塞、缺少缓冲区而丢失

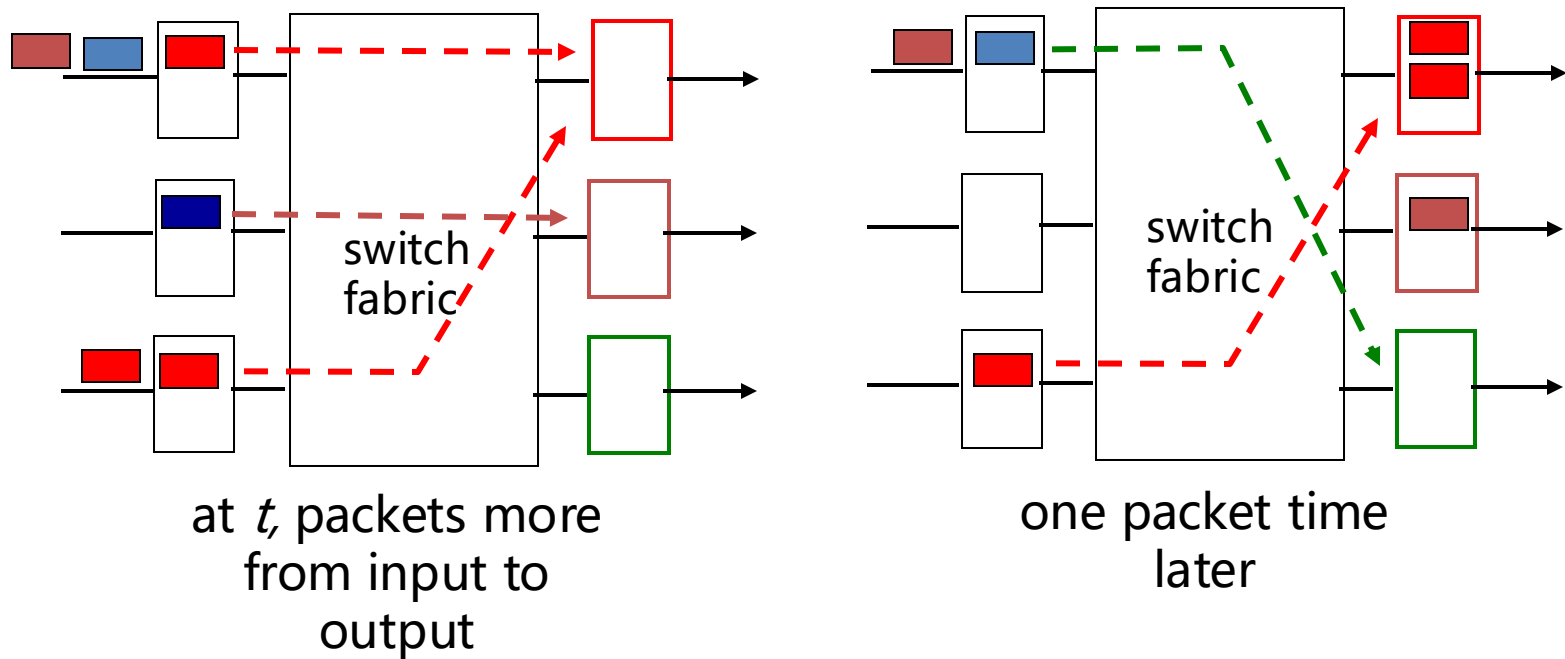
- 当交换结构的处理速度小于输入端口数据流入速度时，在输入端口就会出现排队等待
  - 当输入端口缓存溢出时，就会出现丢包！
- 队首阻塞：处在队首的分组阻碍了队列其他分组的转发



output port contention:  
only one red datagram can be  
transferred.  
*lower red packet is blocked*



one packet time later:  
green packet  
experiences HOL  
blocking



- 当交换速率大于输出端口链路速率时就需要缓存
- 当输出端口缓存溢出时，就会出现丢包

- RFC 3439规则：缓存大小=平均RTT\*链路容量C（这个结果是基于相对较少量的TCP数据流的排队动态分析得到的）
- 目前推荐的规则：当有大量的TCP流（N）流过一条链路时，  
缓存大小= 
$$\frac{RTT \cdot C}{\sqrt{N}}$$

1 4.1 概述

2 4.2 虚电路和数据报网络

3 4.3 路由器工作原理

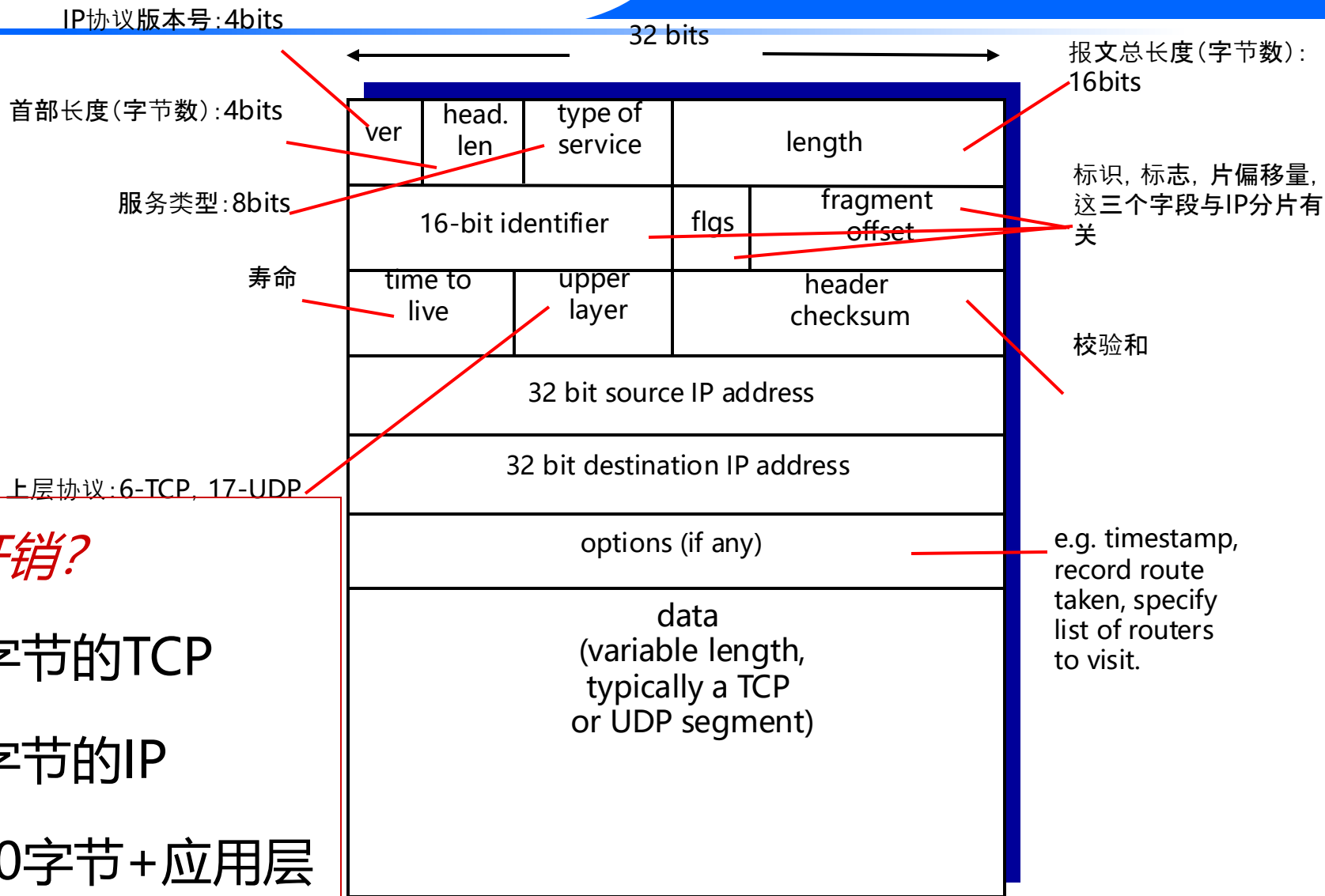
4 4.4 IP协议

数据报格式

IPv4 编址

IPv6

# IP数据报格式



## 多少开销?

- ❖ 20字节的TCP
- ❖ 20字节的IP
- ❖ = 40字节+应用层

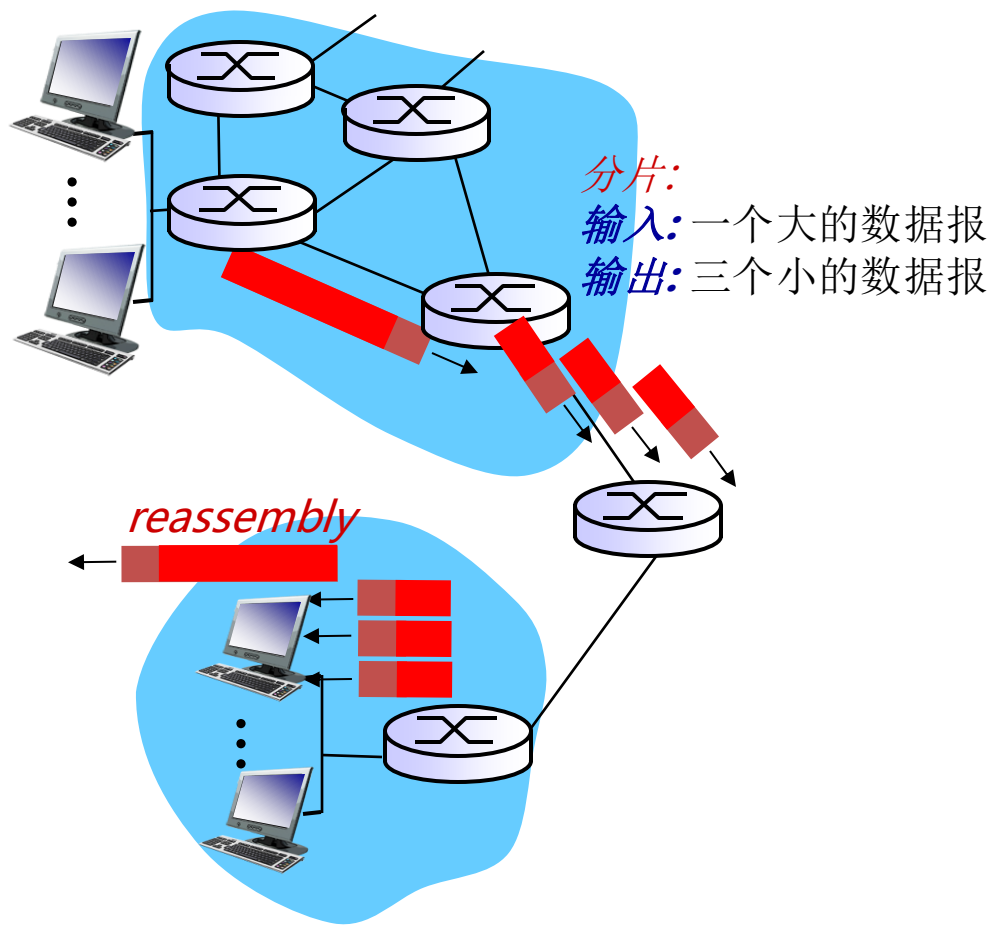
开销

■ 网络链路最大传输单元MTU—链路层帧能承载的最大数据量

- 不同链路类型，最大传输单元也不同，以太网：**1500字节**

■ 将较大的IP数据报分片

- 一个数据报分片为多个较小的数据报
- 在最终的目的端重组分片
- IP首部的标识、标志、片偏移用于分片排序、重组



例:

- ❖ 4000字节数据报
- ❖ MTU = 1500 bytes

	length	ID	fragflag	offset	
	=4000	=x	DF=0	=0	

一个大数据报变成几个较小的数据报

1480 bytes in  
data field  
offset =  
 $1480/8$

	length	ID	fragflag	offset	
	=1500	=x	MF=1	=0	

	length	ID	fragflag	offset	
	=1500	=x	MF=1	=185	

	length	ID	fragflag	offset	
	=1040	=x	MF=0	=370	



1 4.1 概述

2 4.2 虚电路和数据报网络

3 4.3 路由器工作原理

4 4.4 IP协议

数据报格式

IPv4 编址

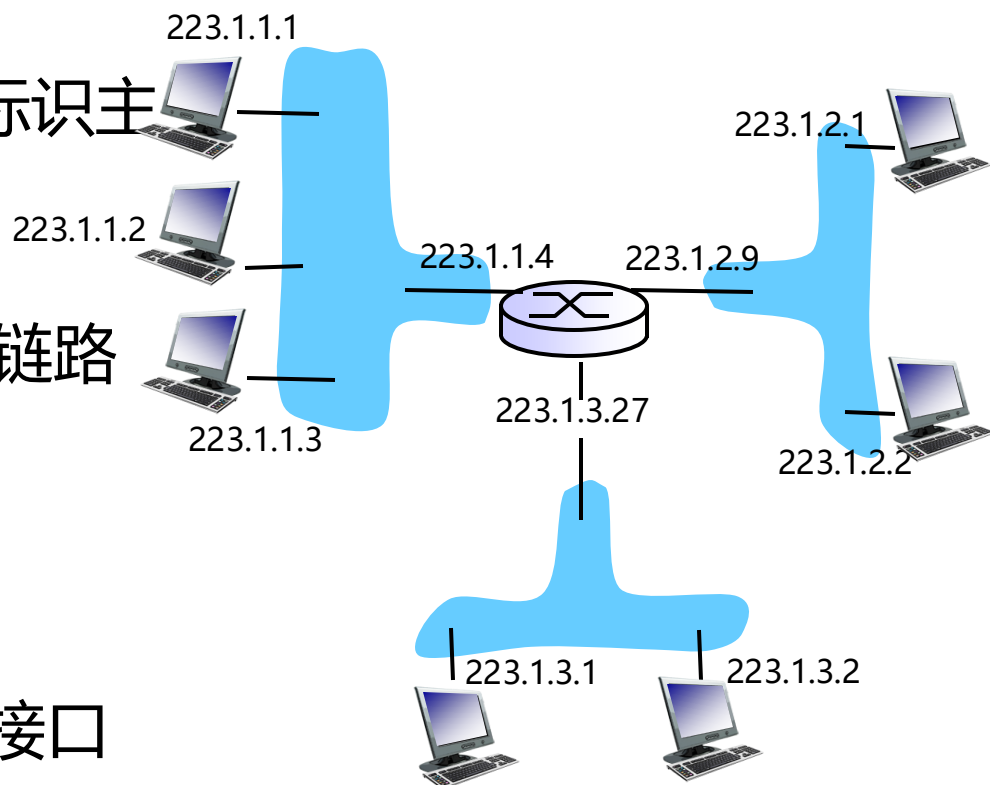
IPv6

# IP编址: IP 地址

**IP地址:** 32位, 用于唯一标识主机, 路由器接口

**接口:** 主机/路由器与物理链路之间的连接

- 路由器通常有多个接口
- 主机一般有一个或连个接口  
(有线以太网, 无线局域网)



802.11)

223.1.1.1 = 11011111 00000001 00000001 00000001  
223                    1                    1                    1

**IP地址与每个接口相关联**

# IP编址：IP地址

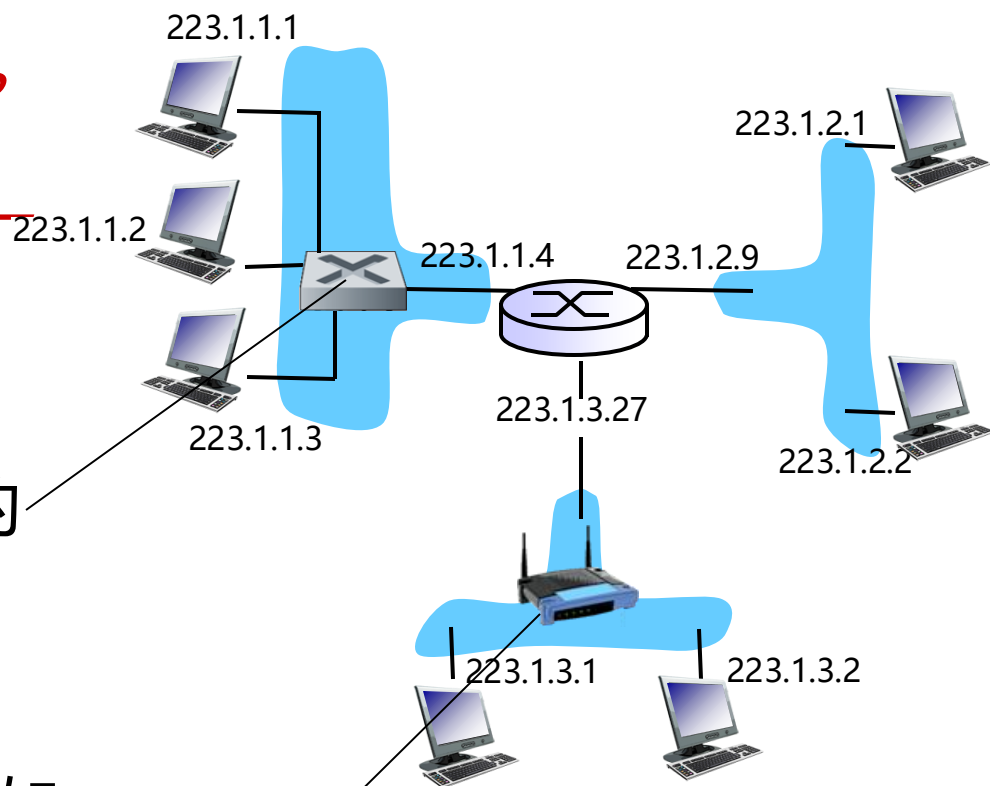
问：接口实际上是如何连接的？

答：我们将在第五章中了解这一点。

A：通过以太网交换机连接的有线以太网接口

目前：不需要担心一个接口如何连接到另一个接口(没有中间路由器)

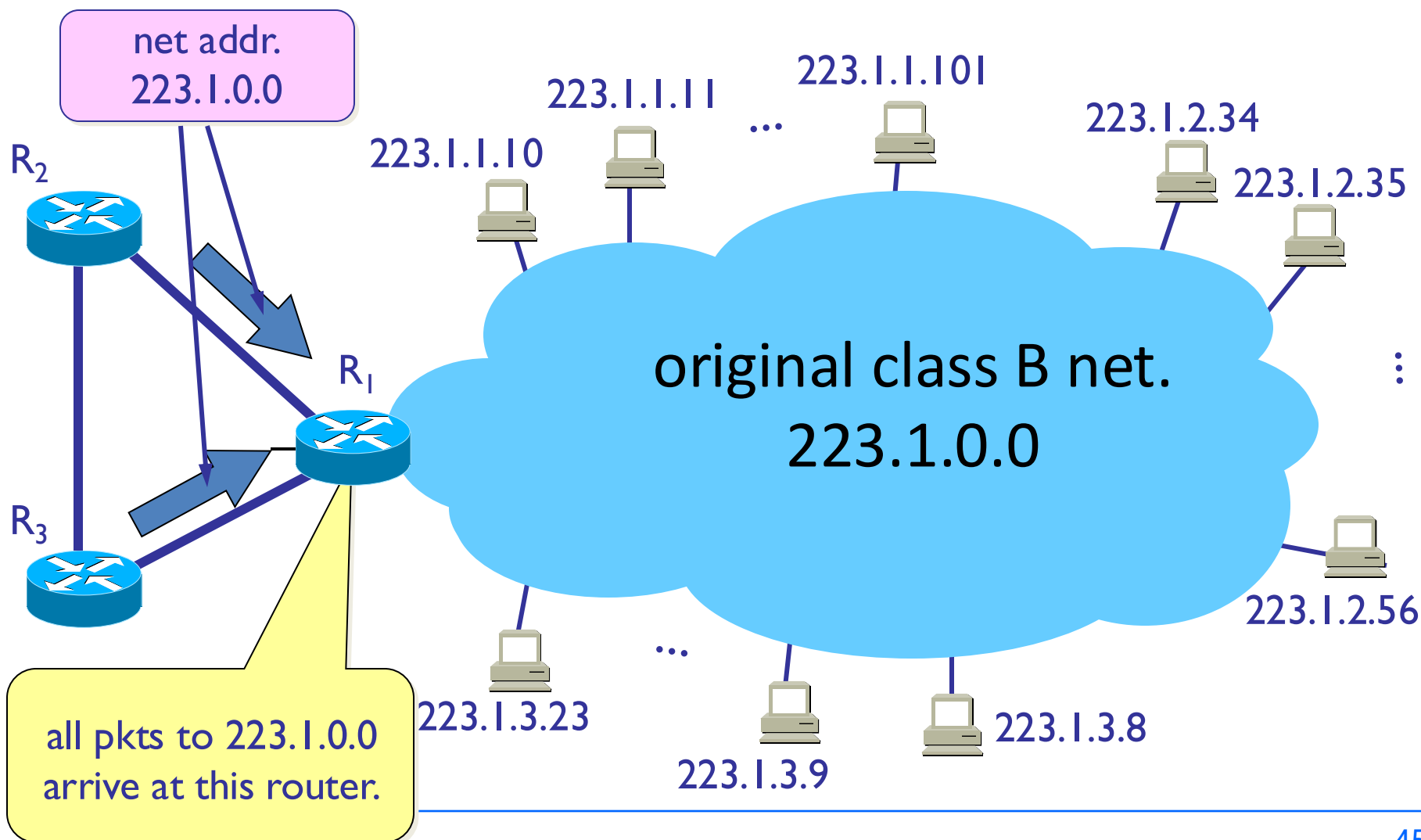
A：通过 WiFi 基站连接的无线 WiFi 接口



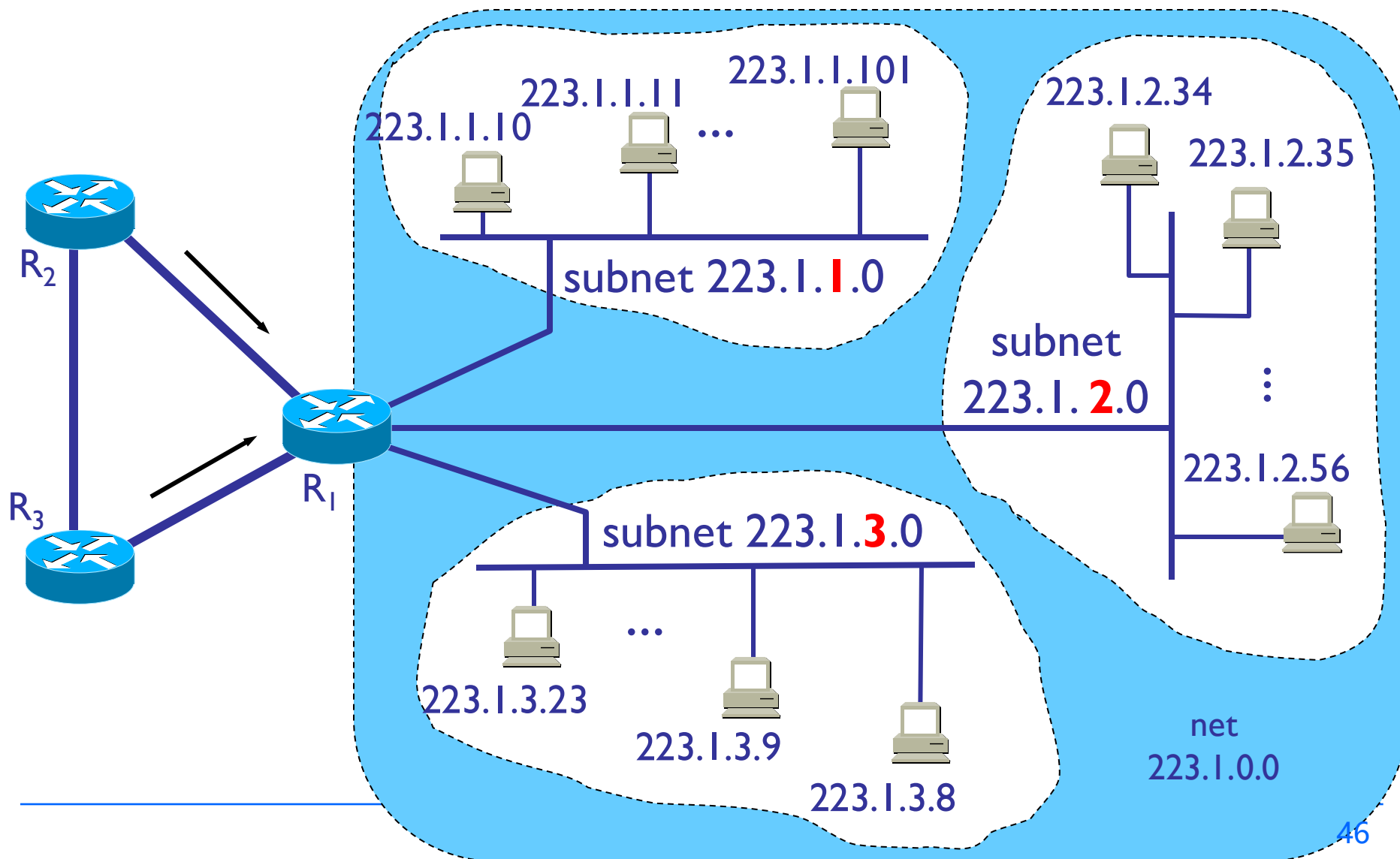
# IP 编址：分类式

Class	Leading bits	Size of network number bit field	Size of rest bit field	Number of networks	Addresses per network	IP address range	Private Address
Class A	0	8	24	128 ( $2^7-2$ )	16,777,216 ( $2^{24}-2$ )		
Class B	10	16	16	16,384 ( $2^{14}-1$ )	65,536 ( $2^{16}-2$ )		172.16.0.0 - 172.31.255.255
Class C	110	24	8	2,097,152 ( $2^{21}-1$ )	256 ( $2^8-2$ )		
Class D (multicast)	1110	not defined	not defined	not defined	not defined		not defined
Class E (reserved)	1111	not defined	not defined	not defined	not defined		

存在一些特殊的IP地址



# IP寻址：子网划分



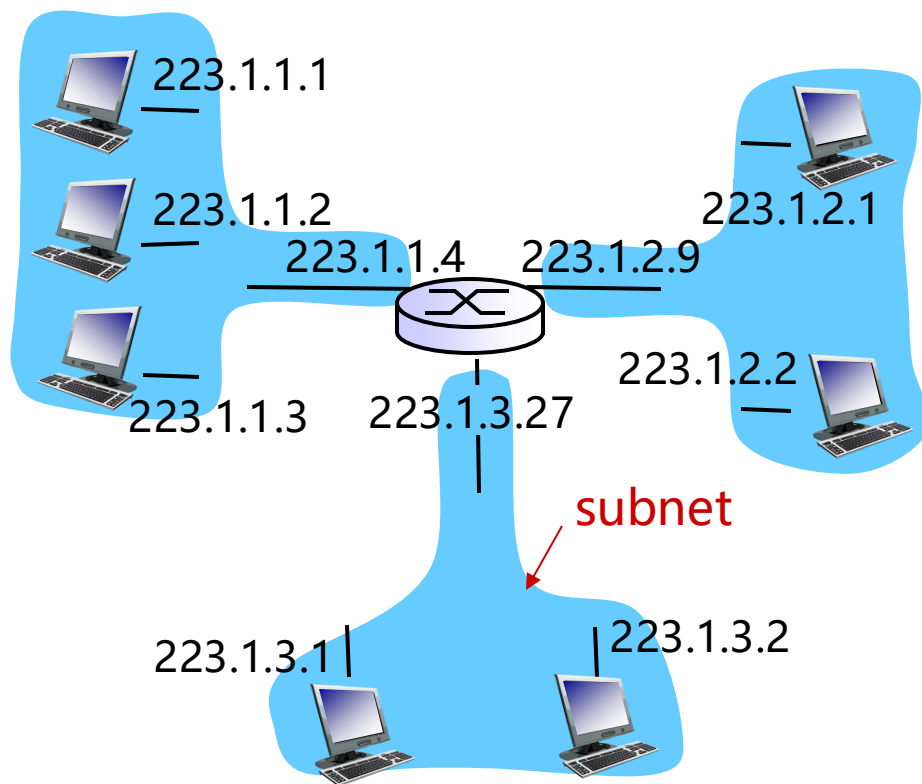
# IP编址：子网划分

## ■ IP地址:

- 子网部分 - 高阶层
- 主机部分 - 低阶层

## ■ 什么是子网?

- IP地址子网部分相同的设备接口
- 不通过路由器可直接物理互联



network consisting of 3 subnets

# IP编址：子网划分

❖ 将接口与其主机或路由器分离，每个隔离的网络称为一个子网

❖ 子网掩码

• 表达

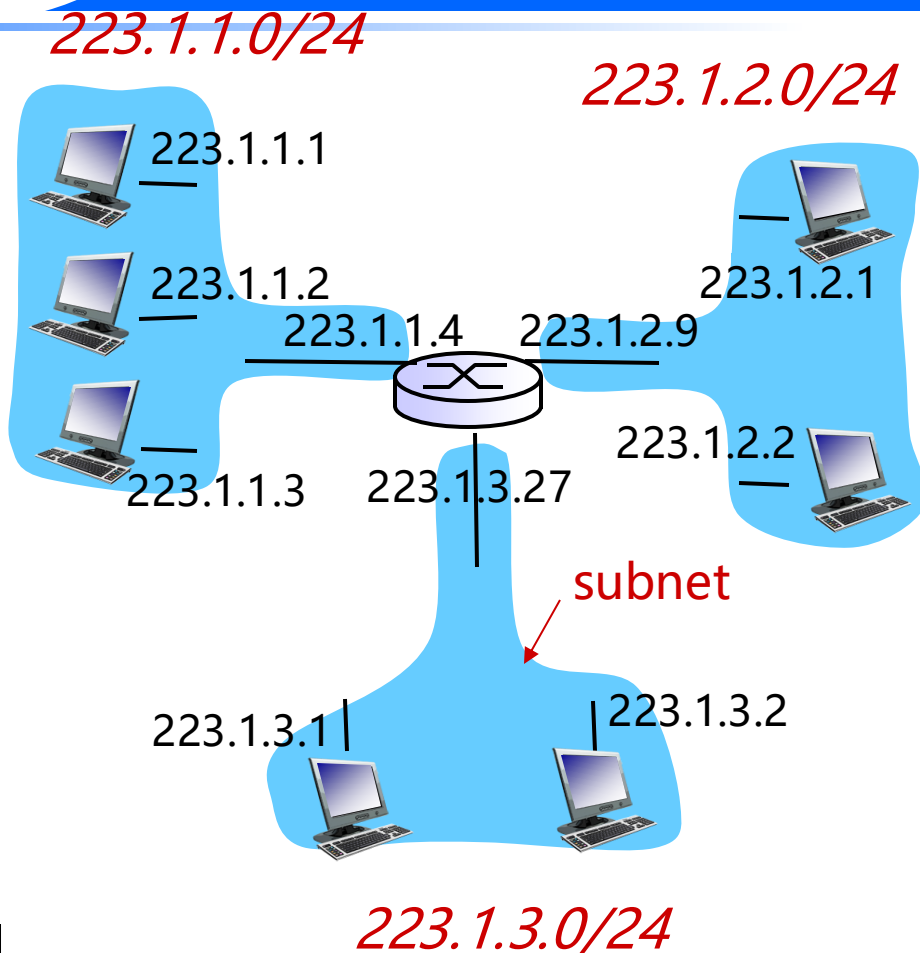
IP地址。 : 223.1.3.27

子网掩码: 255.255.255.0

斜杠表示法: 223.1.3.27/24

网络 ID 的计算

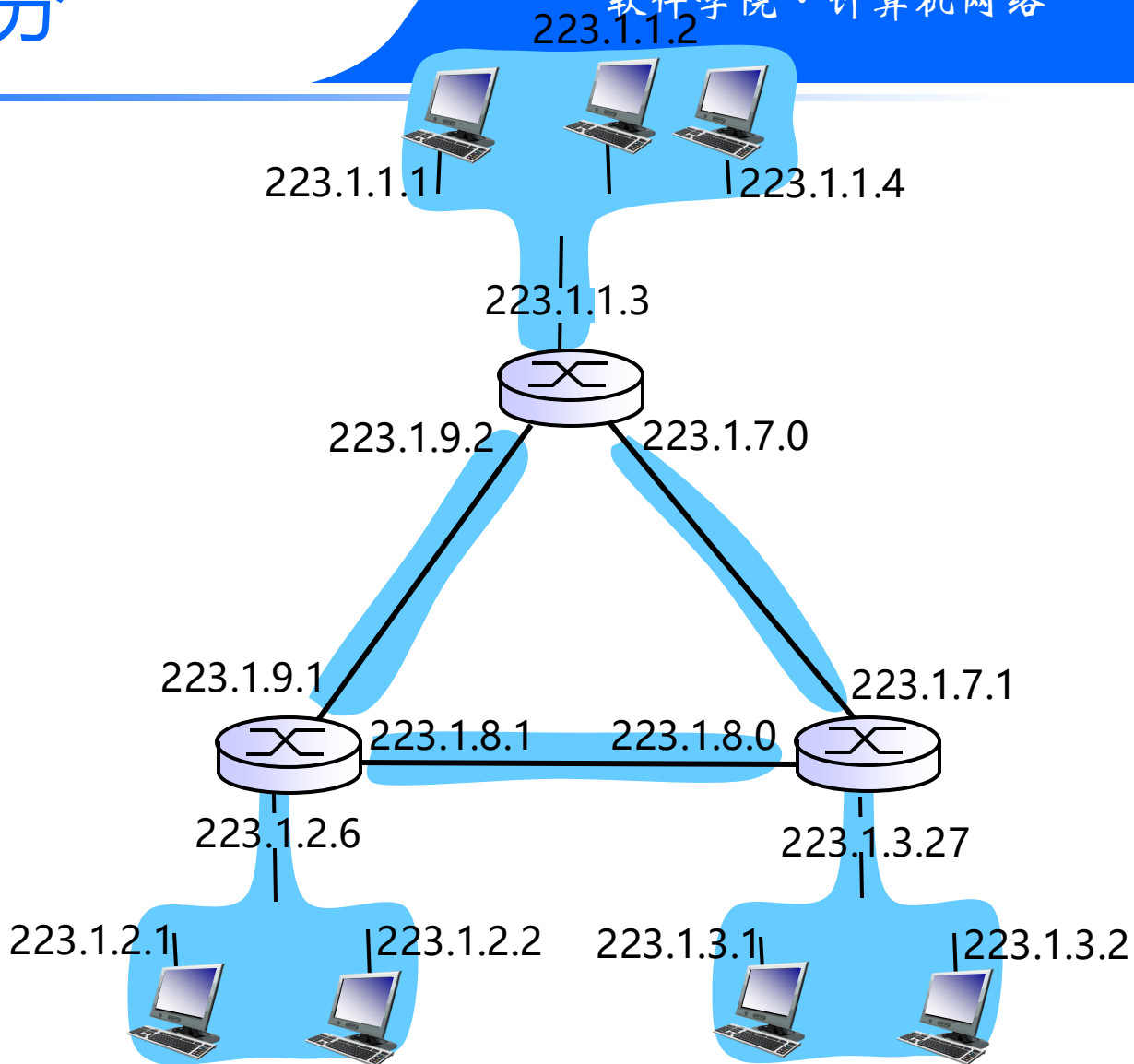
11011111	00000001	00000011	00011011
11111111	11111111	11111111	00000000
11011111	00000001	00000011	00000000





# IP编址：子网划分

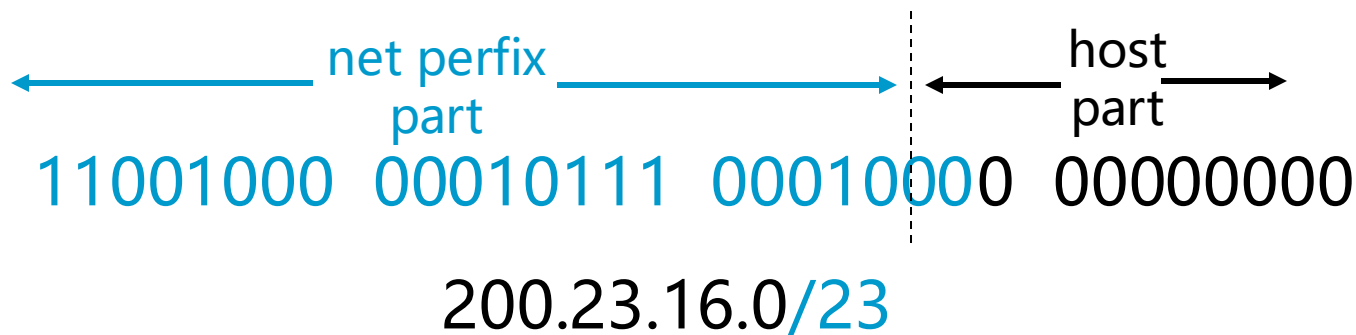
有多少个？



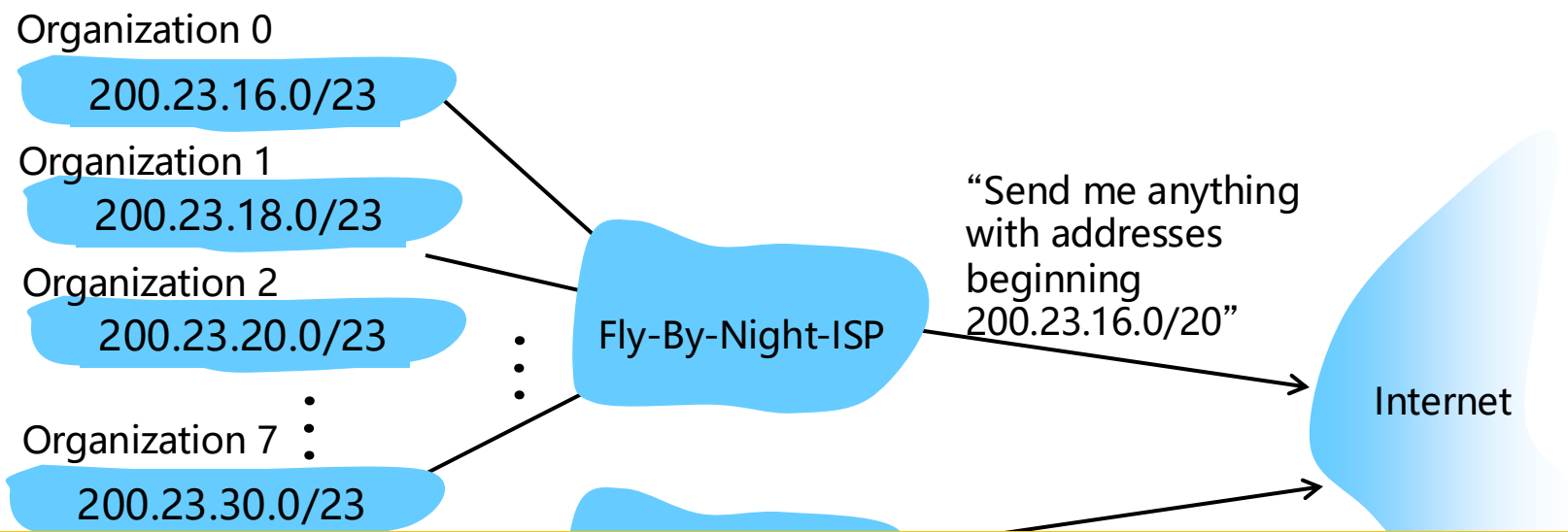
# IP编址: CIDR

## CIDR: 无类别域间路由

- IP地址的子网位可以是任意长度
- 地址格式: **a.b.c.d/x**, x是IP地址子网位的比特位数

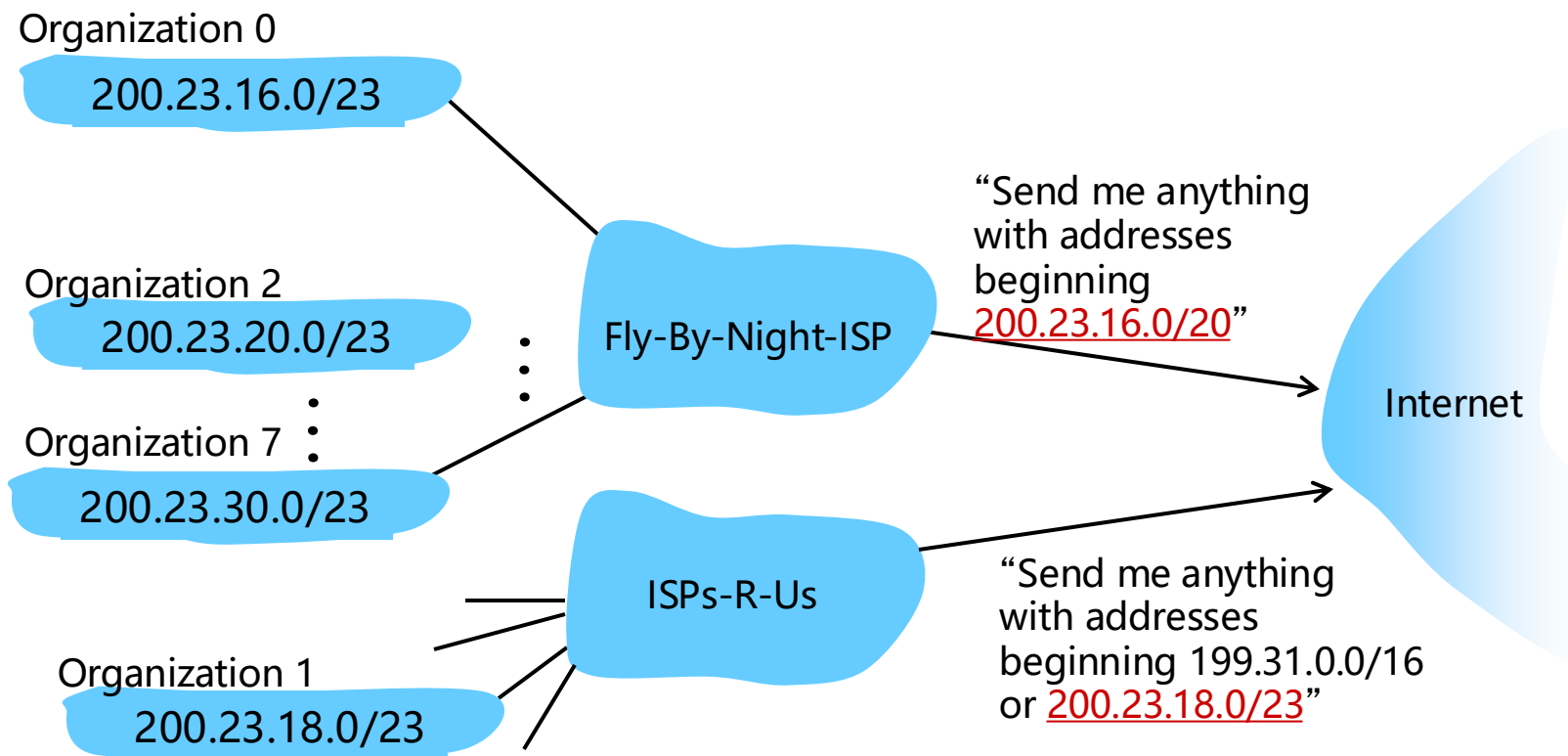


## 分层编址使得路由信息的宣告更加高效: 聚合构成超网



ISP's block	200.23.16.0/20	<u>11001000</u> <u>00010111</u> <u>00010000</u> 00000000
Organization 0	200.23.16.0/23	<u>11001000</u> <u>00010111</u> <u>00010000</u> 00000000
Organization 1	200.23.18.0/23	<u>11001000</u> <u>00010111</u> <u>00010010</u> 00000000
Organization 2	200.23.20.0/23	<u>11001000</u> <u>00010111</u> <u>00010100</u> 00000000

## ISPs-R-US 有更具体的路线到组织 1



# IP编址: CIDR

**CIDR: 为路由寻址转发创造了便利, 但是为子网划分带来了挑战。**

例: 课后习题P11, 考虑互联3个子网, 通过一台路由器互联。所有接口前缀223.1.17/24。子网1需要60个接口, 子网2需要90个, 子网3, 需要12个。要求子网划分方案, 满足需求。 **求解方法: 采用二叉树法。**

子网名称	所需的大小	分配的大小	地址	掩码	十进制掩码	划分范围	广播地址
B	90	126	223.1.17.0	/25	255.255.255.128	223.1.17.1 - 223.1.17.126	223.1.17.127
A	60	62	223.1.17.128	/26	255.255.255.192	223.1.17.129 - 223.1.17.190	223.1.17.191
C	12	14	223.1.17.192	/28	255.255.255.240	223.1.17.193 - 223.1.17.206	223.1.17.207

# IP编址: CIDR

**CIDR: 为路由寻址转发创造了便利, 但是为子网划分带来了挑战。**

例: 考虑互联4个子网, 通过一台路由器互联。所有接口前缀223.1.17/24。每个子网均需要16个接口。要求子网划分方案, 满足需求。

子网名称	所需的大小	分配的大小	地址	掩码	十进制掩码	划分范围	广播地址
A	16	30	223.1.17.0	/27	255.255.255.224	223.1.17.1 - 223.1.17.30	223.1.17.31
B	16	30	223.1.17.32	/27	255.255.255.224	223.1.17.33 - 223.1.17.62	223.1.17.63
C	16	30	223.1.17.64	/27	255.255.255.224	223.1.17.65 - 223.1.17.94	223.1.17.95
D	16	30	223.1.17.96	/27	255.255.255.224	223.1.17.97 - 223.1.17.126	223.1.17.127

**Q:** ISP如何获得IP地址段?

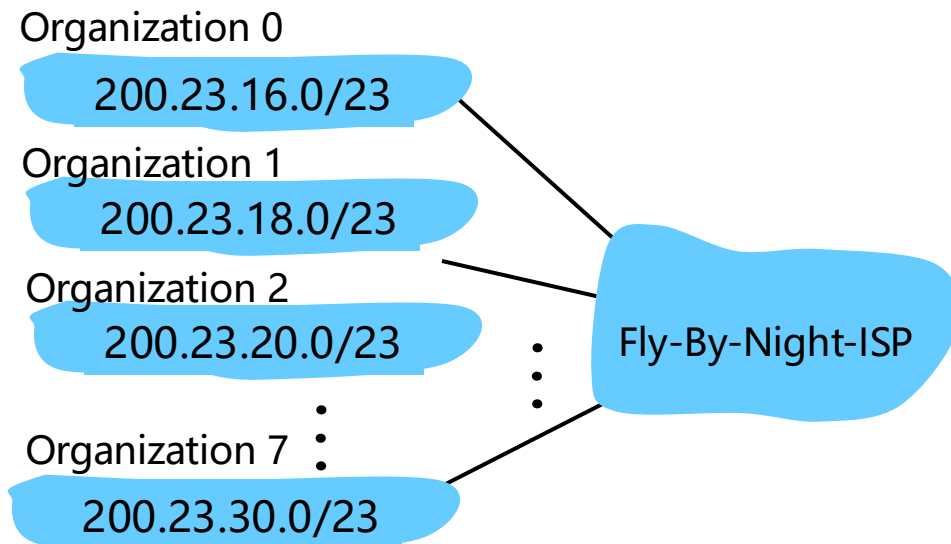
**A:** ICANN: 互联网分配公司

姓名和编号<http://www.icann.org/>

- 分配IP地址
- 管理NDS
- 分派域名, 处理争端

**Q:** 如何得到IP地址的网络位

**A:** 从ISP的地址空间里分配





**Q:**主机如何获取IP地址？

- 由系统管理员手工设置

- Windows：控制面板>网络>配置>tcp/ip->属性
- UNIX： /etc/rc.config
- 静态IP地址：专用服务器、专用电路计算机

- **DHCP**：动态主机配置协议：从DHCP服务器自动获取IP地址

- “即插即用”

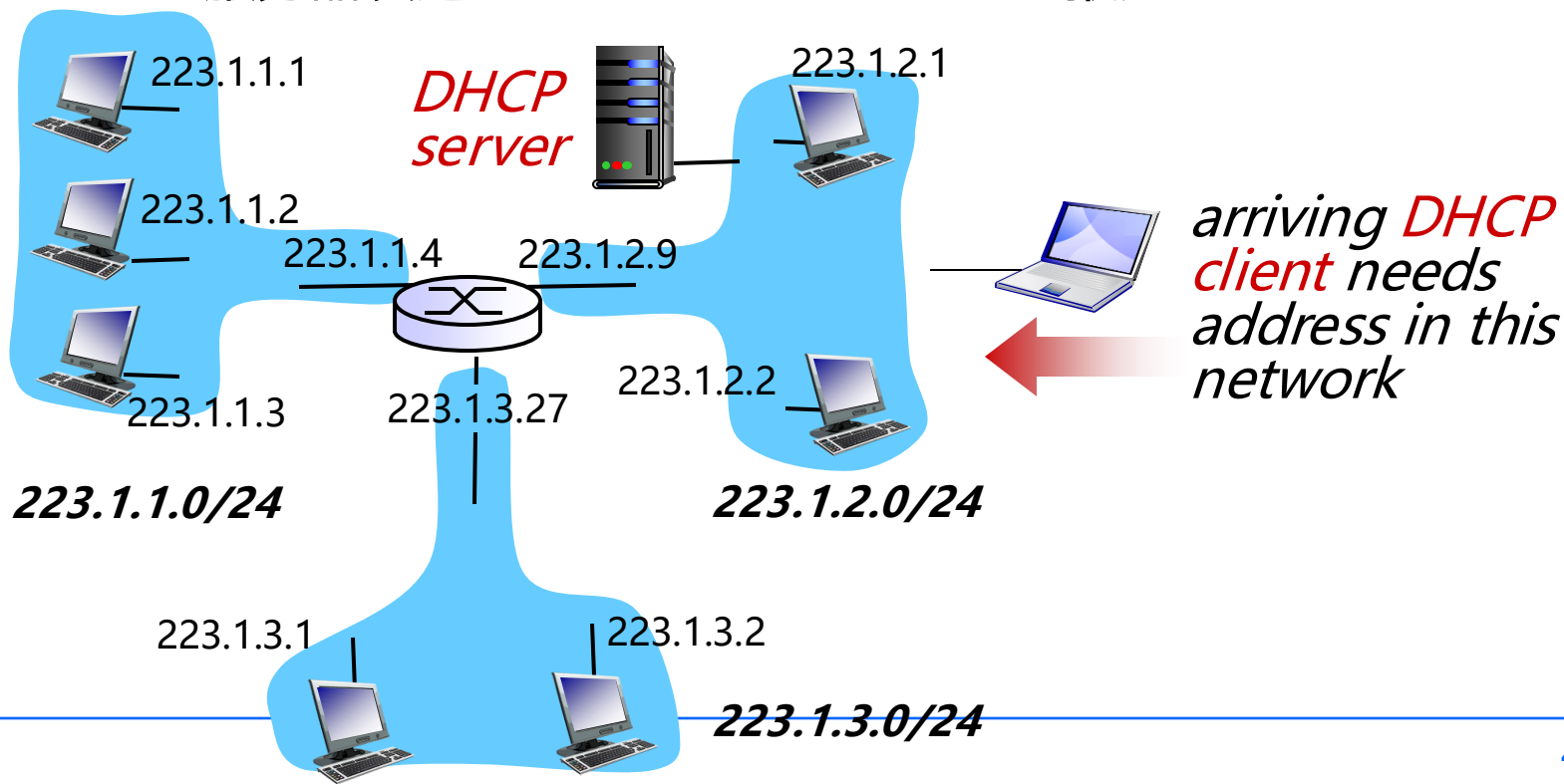
**目标:** 当一台主机加入网络时, 可以自动的从DHCP服务器获得IP地址

- 在用的地址可以更新租期
- 地址复用
- 支持移动用户随时加入网络

DHCP: 不仅是IP地址除了IP地址, DHCP服务同时返回**其他**信息:

- 第一条路由器的地址 (网关地址)
- DNS服务器地址
- 子网掩码 (区分地址中的网络位和主机位)

- 主机广播 “**DHCP discover**” 报文
- DHCP服务器响应 “**DHCP offer**” 报文
- 主机请求IP地址: “**DHCP request**” 报文
- DHCP服务器发送IP地址: “**DHCP ack**” 报文



# DHCP 概述

DHCP server: 223.1.2.5



**DHCP discover**

Broadcast: is there a  
DHCP server out there?

arriving  
client



**DHCP offer**

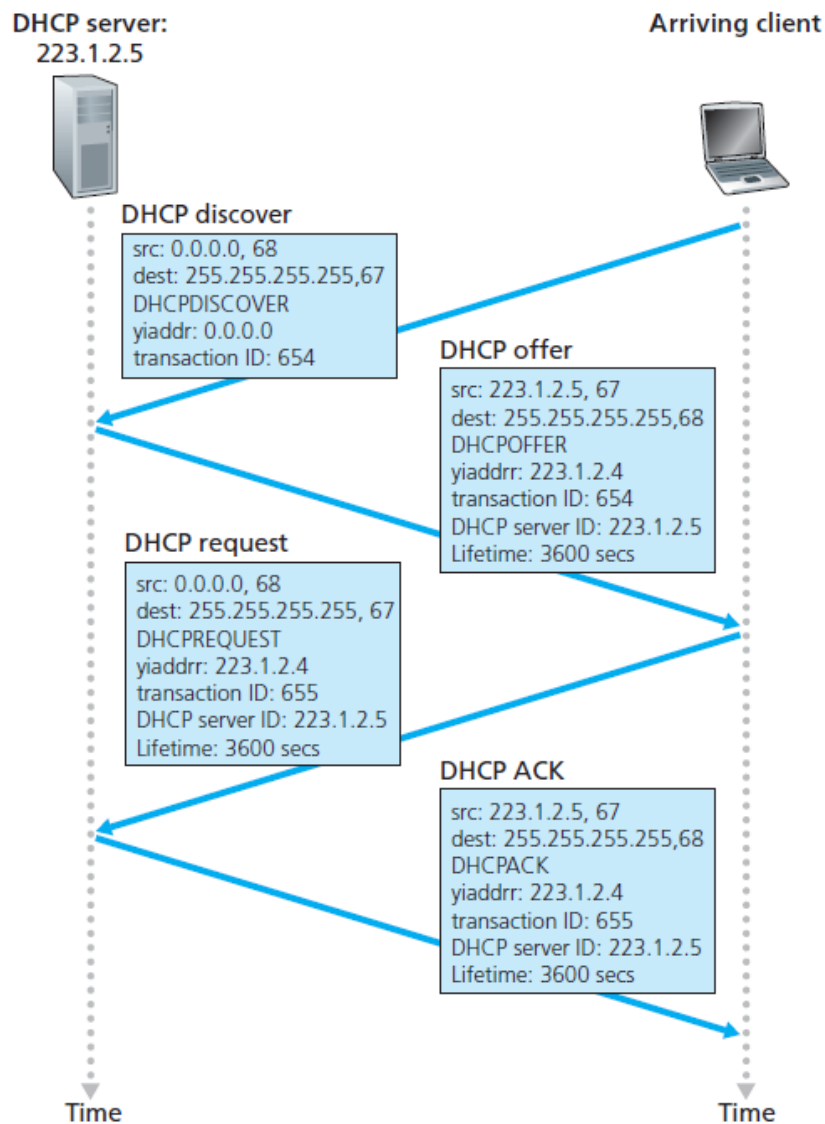
Broadcast: I'm a DHCP  
server! Here's an IP  
address you can use  
.....

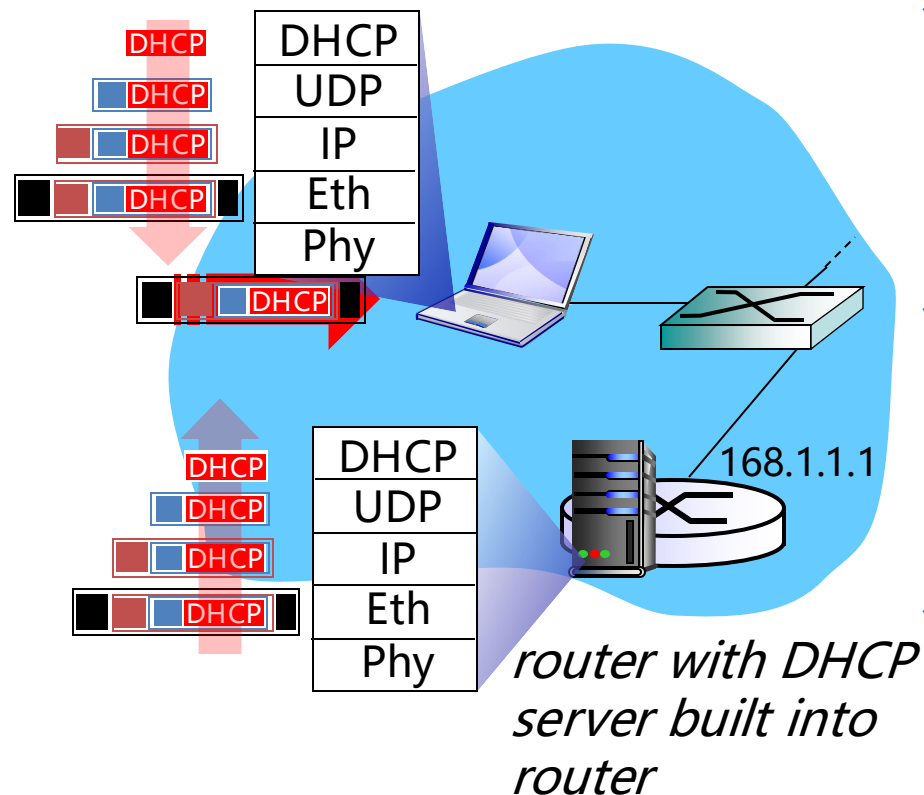
**DHCP request**

Broadcast: OK. I'll take  
that IP address!

**DHCP ACK**

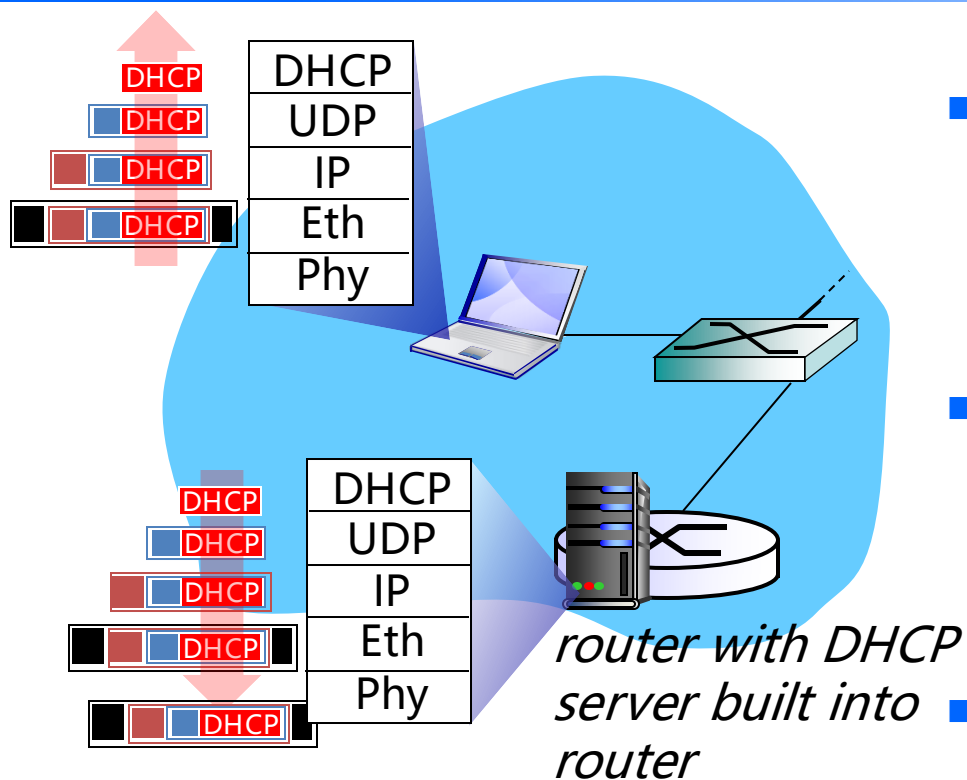
Broadcast: OK. You've  
got that IP address!  
.....





## DHCP是应用层协议

- ❖ 当一台笔记本接入网络时，它需要IP地址、网关地址、DNS服务器地址：此时可使用DHCP服务
- ❖ DHCP请求被封装在UDP协议中--》进而封装在IP协议中--》进而封装在802.1以太网帧中
- ❖ 局域网，以太网帧广播地址为：FF-FF-FF-FF-FF，运行DHCP的服务讲接收到该请求
- ❖ 接收到后，以太网帧会被解封得到IP数据包，再解封得到UDP报文段，在解封得到DHCP报文



- DHCP服务器生成“DHCP ACK”报文，包含了客户端IP地址，网关地址，DNS服务器地址
- DHCP服务器将响应报文封装为帧，并转发给客户端，客户端接收后解析出DHCP报文
- 客户端获得其分配的IP地址，DNS服务器地址，以及网关地址

# DHCP:Wireshark输出 (家庭局域网)

## request

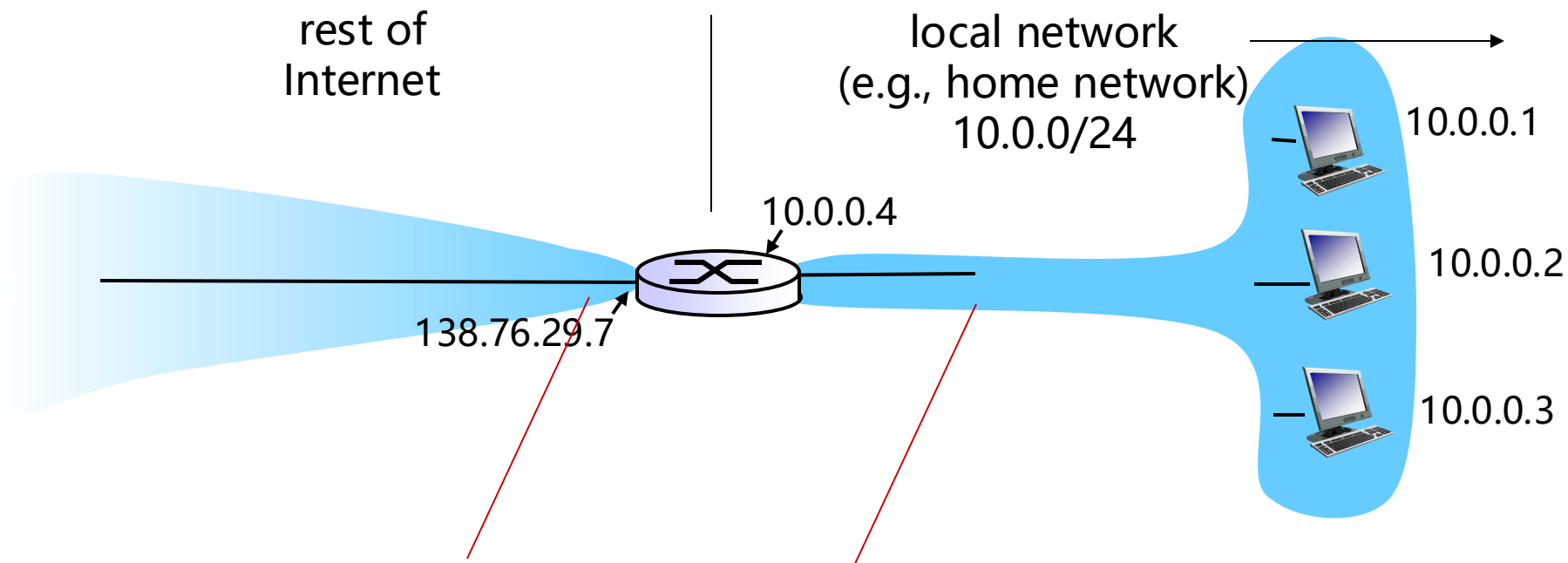
Message type: **Boot Request (1)**  
Hardware type: Ethernet  
Hardware address length: 6  
Hops: 0  
**Transaction ID: 0x6b3a11b7**  
Seconds elapsed: 0  
Bootp flags: 0x0000 (Unicast)  
Client IP address: 0.0.0.0 (0.0.0.0)  
Your (client) IP address: 0.0.0.0 (0.0.0.0)  
Next server IP address: 0.0.0.0 (0.0.0.0)  
Relay agent IP address: 0.0.0.0 (0.0.0.0)  
**Client MAC address: Wistron\_23:68:8a (00:16:d3:23:68:8a)**  
Server host name not given  
Boot file name not given  
Magic cookie: (OK)  
Option: (t=53,l=1) **DHCP Message Type = DHCP Request**  
Option: (61) Client identifier  
    Length: 7; Value: 010016D323688A;  
    Hardware type: Ethernet  
    Client MAC address: Wistron\_23:68:8a (00:16:d3:23:68:8a)  
Option: (t=50,l=4) Requested IP Address = 192.168.1.101  
Option: (t=12,l=5) Host Name = "nomad"  
**Option: (55) Parameter Request List**  
    Length: 11; Value: 010F03062C2E2F1F21F92B  
    **1 = Subnet Mask; 15 = Domain Name**  
    **3 = Router; 6 = Domain Name Server**  
    44 = NetBIOS over TCP/IP Name Server  
    .....

## reply

Message type: **Boot Reply (2)**  
Hardware type: Ethernet  
Hardware address length: 6  
Hops: 0  
**Transaction ID: 0x6b3a11b7**  
Seconds elapsed: 0  
Bootp flags: 0x0000 (Unicast)  
**Client IP address: 192.168.1.101 (192.168.1.101)**  
Your (client) IP address: 0.0.0.0 (0.0.0.0)  
**Next server IP address: 192.168.1.1 (192.168.1.1)**  
Relay agent IP address: 0.0.0.0 (0.0.0.0)  
Client MAC address: Wistron\_23:68:8a (00:16:d3:23:68:8a)  
Server host name not given  
Boot file name not given  
Magic cookie: (OK)  
**Option: (t=53,l=1) DHCP Message Type = DHCP ACK**  
**Option: (t=54,l=4) Server Identifier = 192.168.1.1**  
**Option: (t=1,l=4) Subnet Mask = 255.255.255.0**  
**Option: (t=3,l=4) Router = 192.168.1.1**  
**Option: (6) Domain Name Server**  
    Length: 12; Value: 445747E2445749F244574092;  
    IP Address: 68.87.71.226;  
    IP Address: 68.87.73.242;  
    IP Address: 68.87.64.146  
**Option: (t=15,l=20) Domain Name = "hsd1.ma.comcast.n**



# NAT: 网络地址转换



所有数据报离开本地网络时均使用  
同一个源端的NAT IP地址：  
138.76.29.7，而源端口号各不相同

本地网络内部相互传输的数据报其源  
地址和目的地址直接使用10.0.0.0/24  
的私有地址

**目的：**对外部网络而言，本地网络只需一个IP地址就足够了：

- ISP不再需要一个地址区间：只需一个IP地址就可以满足所有的设备了
- 本地网络的内部设备的地址变更无需通告外部网络
- 变更ISP对本地网络内部设备无需做任何修改
- 本地网络内部设备没有公网IP地址，外网无法直接访问（增强了安全性）

**实现：** NAT路由器必须

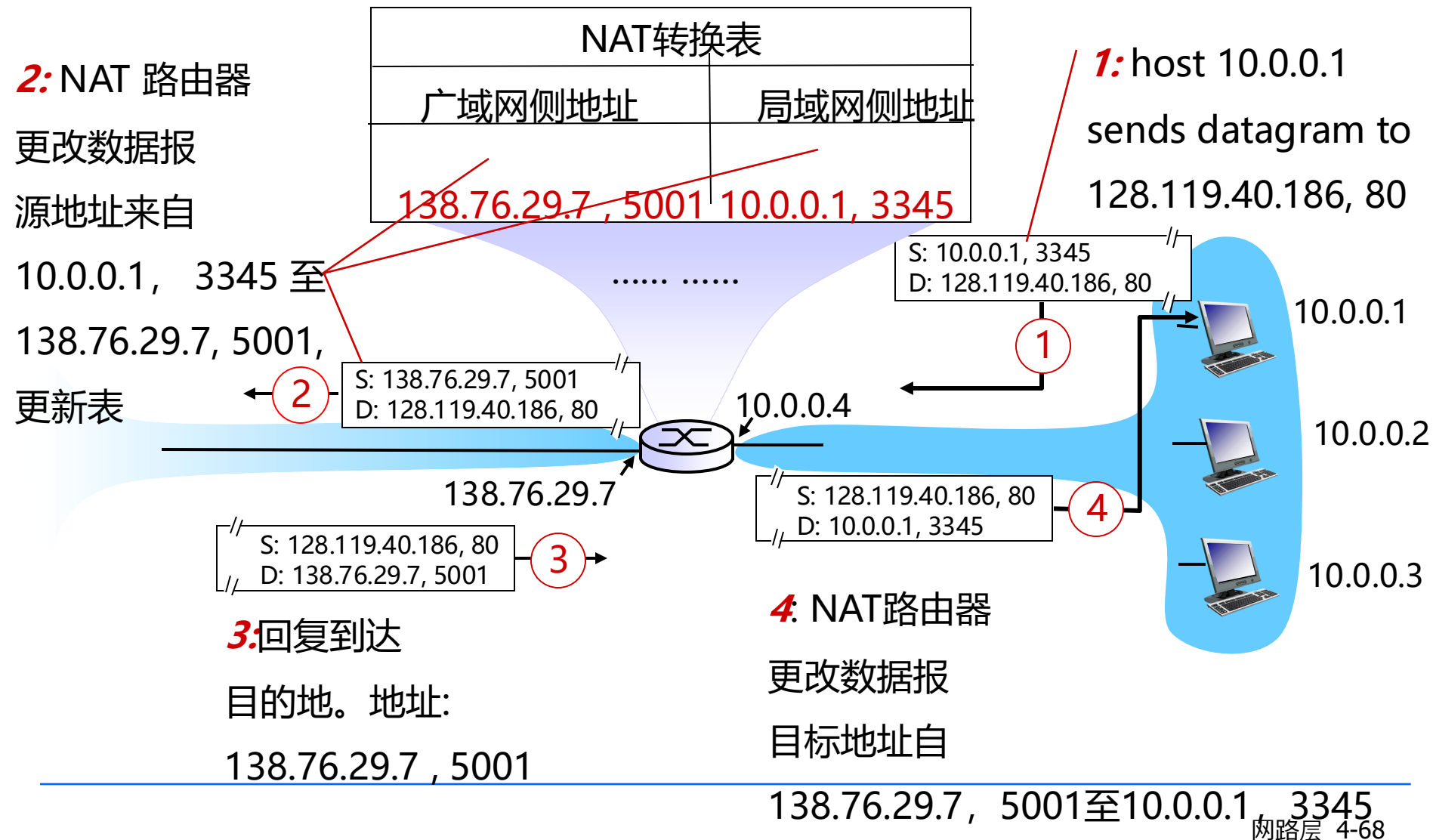
*向外转发数据报时：* 将所有向外转发的数据报的（源IP地址，端口号）替换为（NAT IP地址，新的端口号）

远端的客户端或服务器响应时将使用（NAT IP地址，新的端口号）作为目的地址

- *通过NAT转换表，* 记录每个（源IP地址，源端口号）到（NAT IP地址，新的端口号）的映射关系
- *外部进入的数据报：* 每一个外部进入的数据报，根据存储在NAT转换表里的映射关系将其目的地址字段从（NAT IP地址，新的端口号）替换回（源IP地址，端口号）

# NAT: 网络地址转换

软件学院 · 计算机网络



- 端口号为16bit (65535) :
  - 在一个局域网内部可同时满足60000多个网络连接!

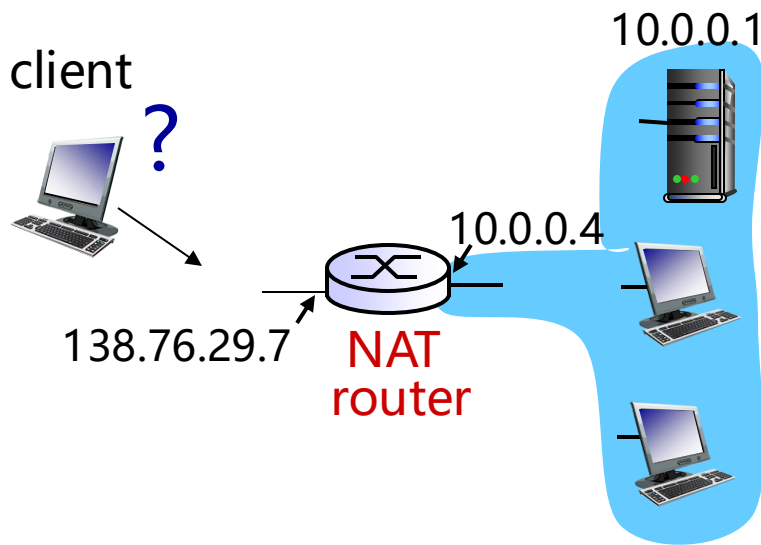
- NAT 备受争议:

- 路由器必须在网络层处理地址转换
- 违反了端到端的通信原则

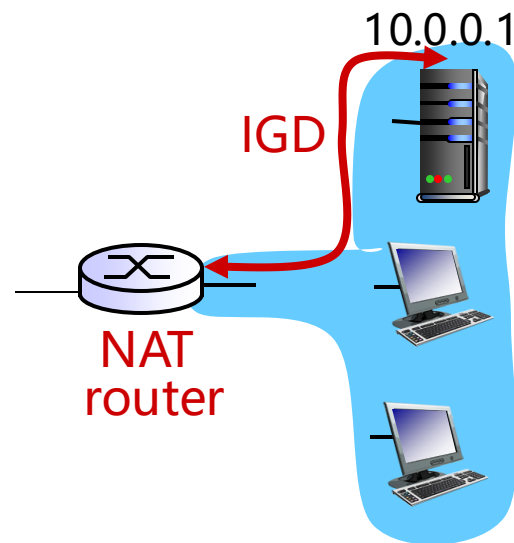
NAT的使用，很可能使得应用开发人员必须将其考虑在内，如P2P应用

- IPv4地址资源枯竭的终极解决方案：IPv6

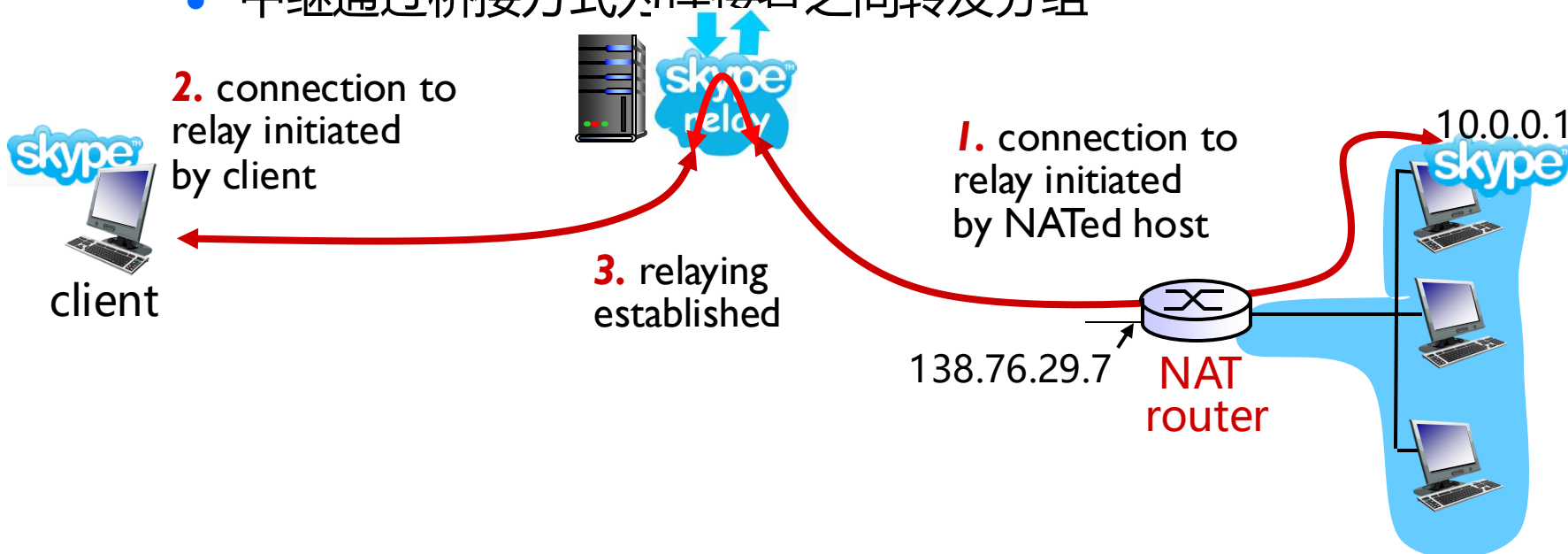
- client想要连接到地址为10.0.0.1的服务器
  - 服务器地址10.0.0.1为局域网内部地址，client不能直接使用该地址作为目的地址
- **方案1:** 配置静态NAT转换信息，将一个特定的端口号映射为服务器
  - 例如 (123.76.29.7: 2500) 将始终转发到10.0.0.1:25000



- **方案2:** 通用即插即用 (UPnP) 互联网网关设备 (IGD) 协议, 可以使得通过NAT转换的主机:
  - ❖ 动态配置
  - ❖ 添加/删除端口映射 (有租约时间)
  - ❖ 例如: 自动静态NAT端口映射配置



- **方案3: 中继 (Skype采用该方案)**
  - NAT内部主机与中继设备建立连接
  - 外部主机也连接到中继设备
  - 中继通过桥接方式为连接者之间转发分组





1 4.1 概述

2 4.2 虚电路和数据报网络

3 4.3 路由器工作原理

4 4.4 IP协议

数据报格式

IPv4 编址

IPv6

- **最初起因:** 32bit的IPv4地址空间很快将被全部分配
- 其他原因:
  - 首部格式可以帮助加快处理/转发的速度
  - 改变首部信息来促进QoS的发展

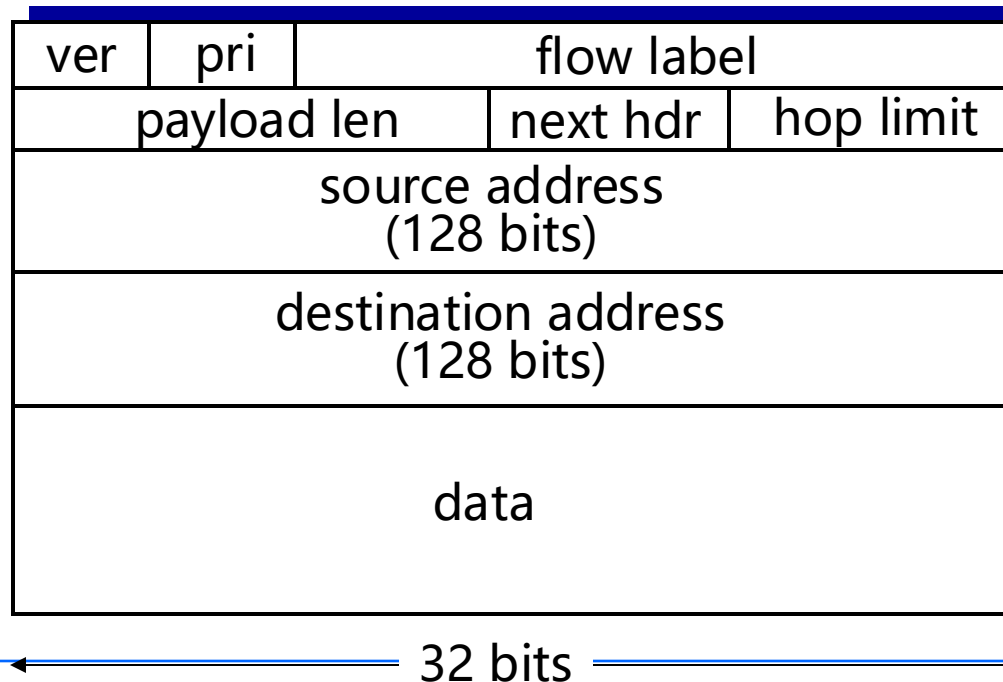
## **IPv6数据报格式:**

- 首部长度固定为40byte
- 不再允许分片

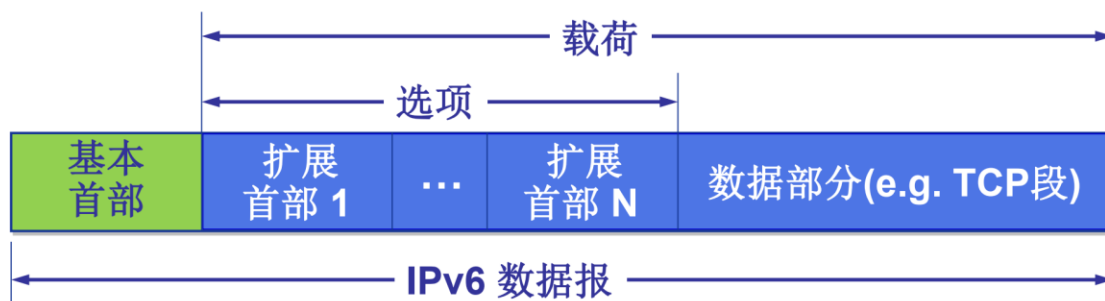
*priority*: 定义数据报在流中的优先级

*flow Label*: 在同一流中唯一的标识数据报

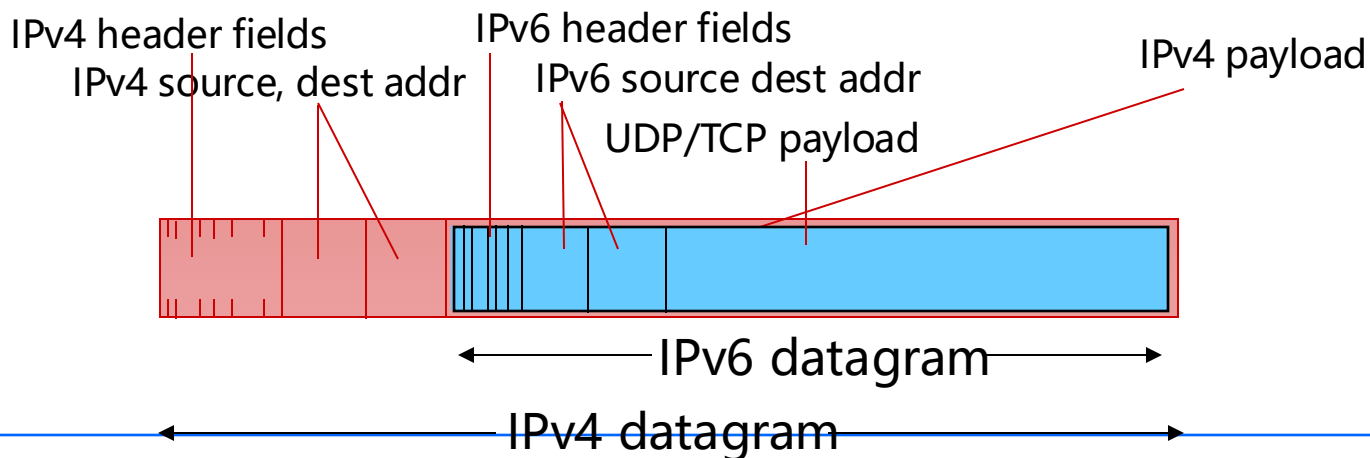
*next header*: 标识所承载数据的上层协议



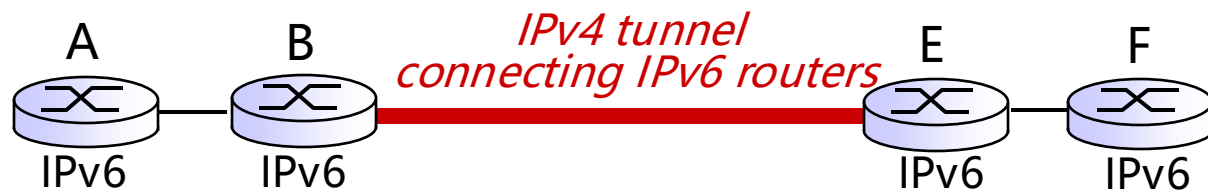
- *identifier, flags, offset*: 删除. 由于复杂性, IPv6 中不允许片段。
- *checksum*: 取消, 减少了每一条的处理时间
- *options*: 不再作为独立字段包含在首部中, 需要时, 可在 “next header” 字段中标示
- *ICMPv6*: 新版本的ICMP
  - 增加了报文类型
  - 增加了组播组的管理功能



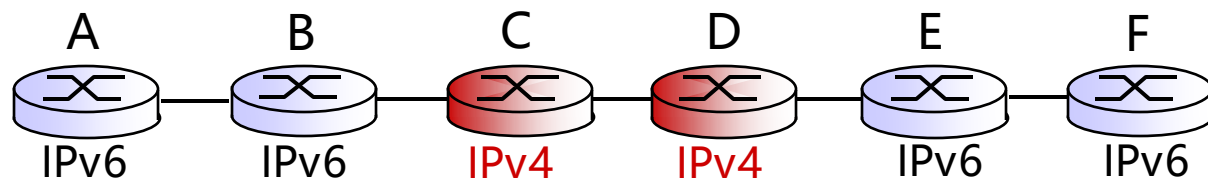
- 不是所有的路由器都可以同时升级来支持IPv6
  - 没有一个明确的时间节点
  - 如何处理网络中IPv4和IPv6的混用
- **隧道**: 在IPv4的路由器之间将IPv6数据报作为IPv4数据报的负载



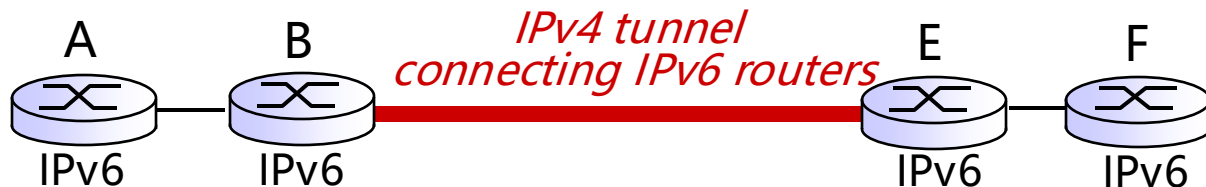
逻辑视图:



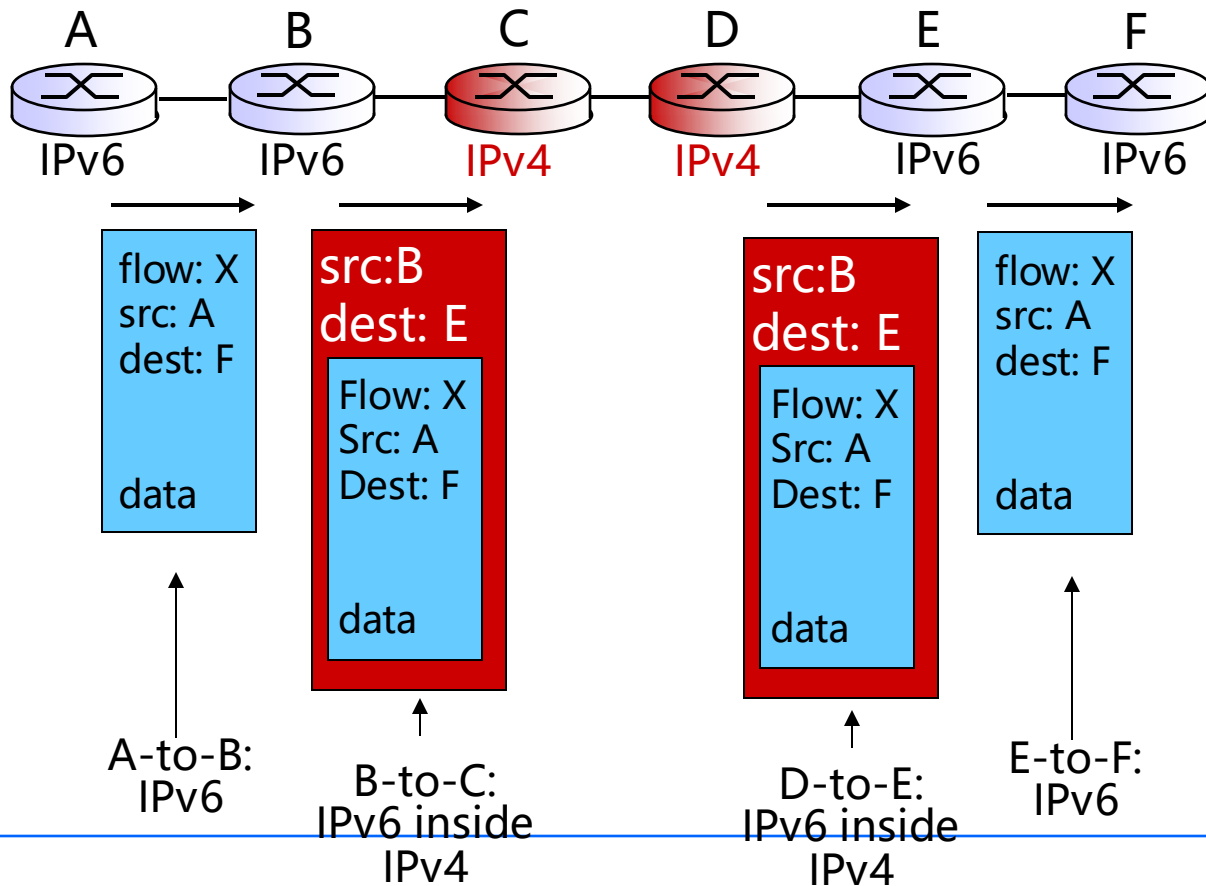
物理视图:



逻辑视图:



物理视图:



# IPv6的使用情况

- ❖ 谷歌：8%的客户端通过IPv6访问服务
- ❖ NIST：1/3的美国政府域名支持IPv6
- ❖ *还需很久很久来完善和使用*
  - 20年甚至更久
  - 想想过去20年应用层的变化：WWW, Facebook
  - *这是为什么呢？*



本章重点：

1. 路由器转发原理；
2. 网络层提供不同的服务类型；
3. IP编址：分配、使用，演进（IPV6）
4. NAT

作业：

1. R3、R10、R22、R24
2. P5、P11、P14、P15