

主讲人: 李全龙

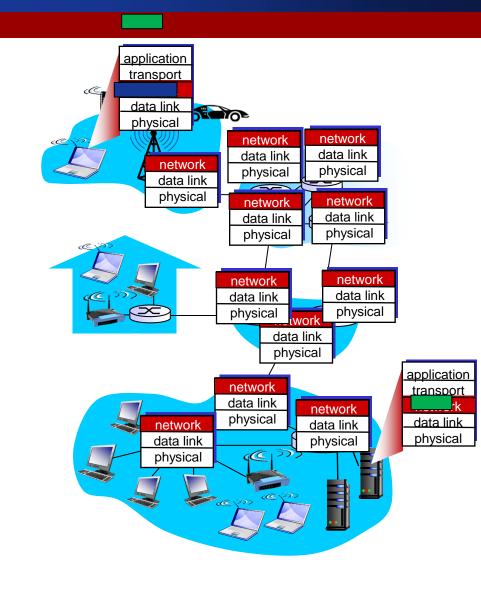
# 本讲主题

## 网络层服务



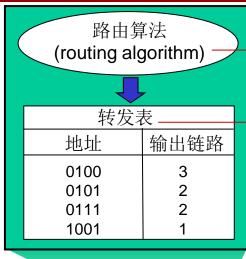
### 网络层

- ❖ 从发送主机向接收主机传 送数据段(segment)
- ❖ 发送主机:将数据段封装 到数据报(datagram)中
- ❖ 接收主机:向传输层交付数据段(segment)
- ❖ 每个主机和路由器都运行 网络层协议
- ❖ 路由器检验所有穿越它的 IP数据报的头部域
  - 决策如何处理IP数据报



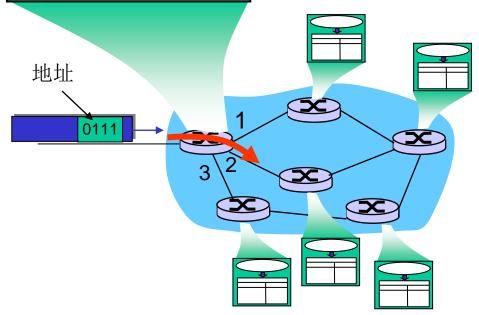
### 网络层核心功能-转发与路由

- \* 转发(forwarding): 将分组从路由器的输入端口转移到合适的输出端口
- ❖ 路由(routing): 确定 分组从源到目的经 过的路径
  - 路由算法 (routing algorithms)



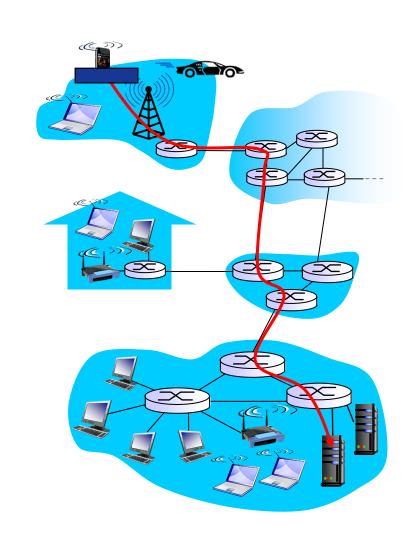
路由算法(协议)确定 通过网络的端到端路径

\_转发表确定在本路 由器如何转发分组



### 网络层核心功能-连接建立

- ❖某些网络的重要功能:
  - ATM, 帧中继, X.25
- ❖ 数据分组传输之前两端主机需要首先建立虚拟/逻辑连接
  - 网络设备(如路由器)参与连接的建立
- ❖ 网络层连接与传输层连接的 对比:
  - 网络层连接:两个主机之间(路 径上的路由器等网络设备参与 其中)
  - 传输层连接: 两个应用进程之 间(对中间网络设备透明)



### 网络层服务模型

Q: 网络层为发送端(主机)到接收端(主机)的数据报传送"通道(channel)"提供什么样的服务模型(service model)?

N	Network		Guarantees ?				Congestion
Arch	nitecture		Bandwidth	Loss	Order	Timing	
	Internet	best effort	none	no	no	no	no (inferred via loss)
	ATM	CBR	constant	yes	yes	yes	no
			rate				congestion
	ATM	VBR	guaranteed	yes	yes	yes	no
			rate				congestion
	ATM	ABR	guaranteed	no	yes	no	yes
			minimum				<u> </u>
	ATM	UBR	none	no	yes	no	no





### 网络层服务模型

### ❖无连接服务(connection-less service):

- 不事先为系列分组的传输确定传输路径
- 每个分组独立确定传输路径
- 不同分组可能传输路径不同
- 数据报网络(datagram network)

### ❖连接服务(connection service):

- 首先为系列分组的传输确定从源到目的经过的路径 (建立连接)
- 然后沿该路径(连接)传输系列分组
- 系列分组传输路径相同
- 传输结束后拆除连接
- 虚电路网络(virtual-circuit network)









主讲人: 李全龙

# 本讲主题

## 虚电路网络

# 连接服务与无连接服务

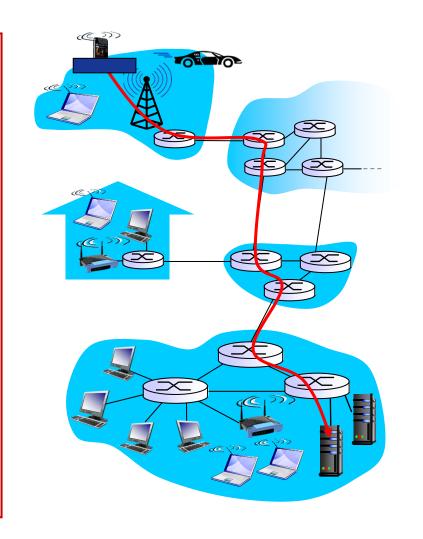
- ❖数据报(datagram)网络与虚电路(virtual-circuit)网络是典型两类分组交换网络
- \*数据报网络提供网络层无连接服务
- ❖虚电路网络提供网络层连接服务
- ❖类似于传输层的无连接服务(UDP)和面向连接服务(TCP),但是网络层服务:
  - 主机到主机服务
  - 网络核心实现



# 虚电路(Virtual circuits)

虚电路: 一条从源主机 到目的主机,类似于 电路的路径(逻辑连接)

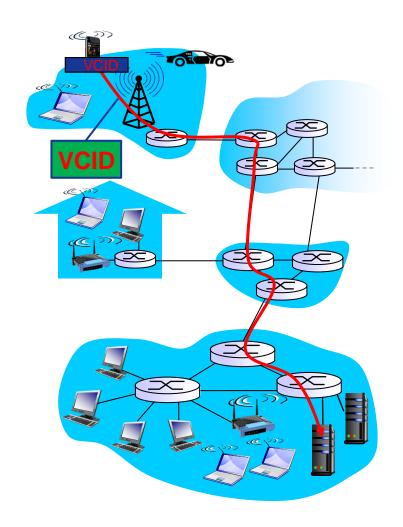
- 分组交换
- 每个分组的传输利用链 路的全部带宽
- 源到目的路径经过的网络层设备共同完成虚电路功能



# 虚电路(Virtual circuits)

#### ❖ 通信过程:

- 呼叫建立(call setup)→数据传输 →拆除呼叫
- ❖ 每个分组携带虚电路标识(VC ID),而不是目的主机地址
- ※虚电路经过的每个网络设备 (如路由器),维护每条经过 它的虚电路连接状态
- \*链路、网络设备资源(如带宽、缓存等)可以面向VC进行预分配
  - 预分配资源=可预期服务性能
  - 如ATM的电路仿真(CBR)



## VC的具体实现

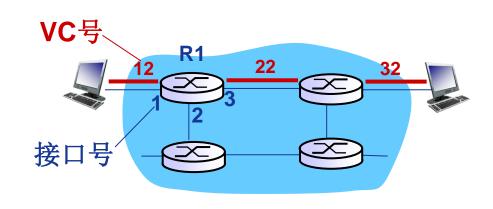
#### 每条虚电路包括:

- 1. 从源主机到目的主机的一条路径
- 2. 虚电路号(VCID),沿路每段链路一个编号
- 3. 沿路每个网络层设备(如路由器),利用转发表记录经过的每条虚电路
- ❖ 沿某条虚电路传输的分组,携带对应虚电路的 VCID,而不是目的地址
- ❖ 同一条VC,在每段链路上的VCID通常不同
  - 路由器转发分组时依据转发表改写/替换虚电路号

主讲人: 李全龙



## VC转发表



#### 路由器R1的VC转发表:

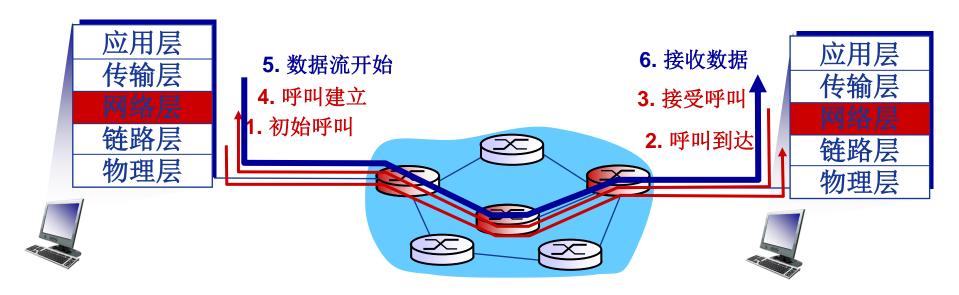
输入接口	输入VC#	输出接口	输出VC#
1	12	3	22
2	63	1	18
3	7	2	17
1	97	3	87

VC路径上每个路由器都需要维护VC连接的状态信息!



### 虚电路信令协议(signaling protocols)

- ❖用于VC的建立、维护与拆除
  - 路径选择
- ❖应用于虚电路网络
  - 如ATM、帧中继(frame-relay)网络等
- ❖目前的Internet不采用







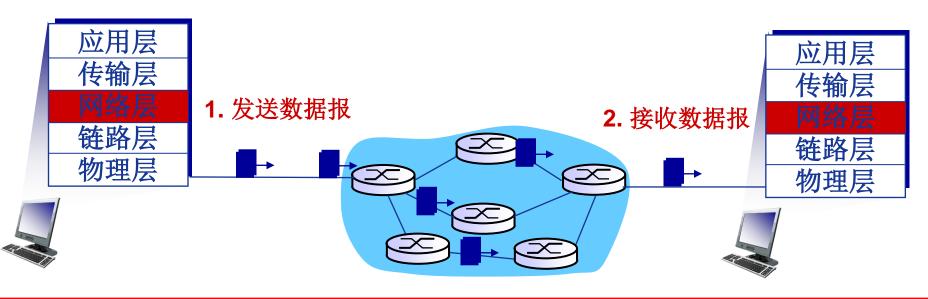
主讲人: 李全龙

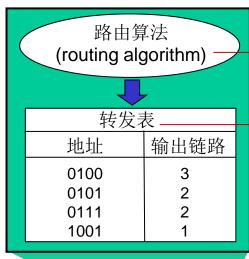
# 本讲主题

## 数据报网络

# 数据报网络

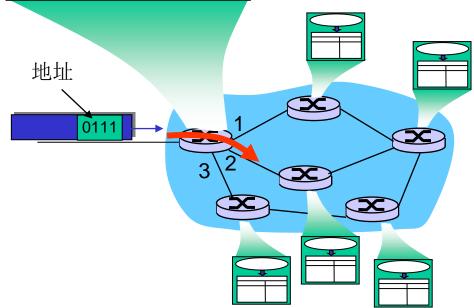
- \* 网络层无连接
- \*每个分组携带目的地址
- \* 路由器根据分组的目的地址转发分组
  - 基于路由协议/算法构建转发表
  - 检索转发表
  - 每个分组独立选路

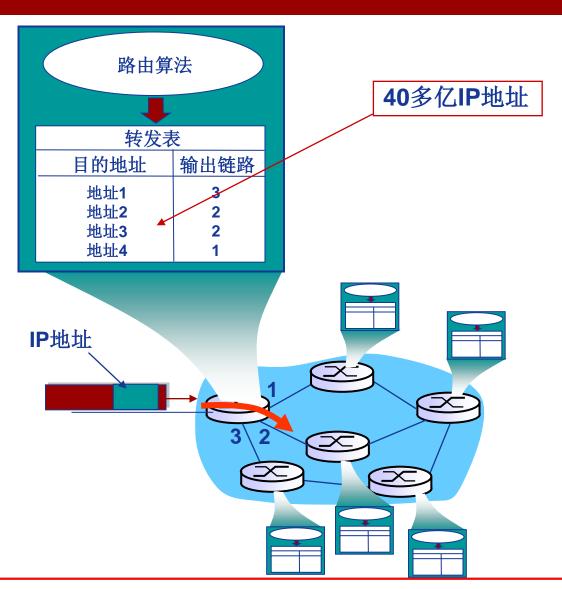


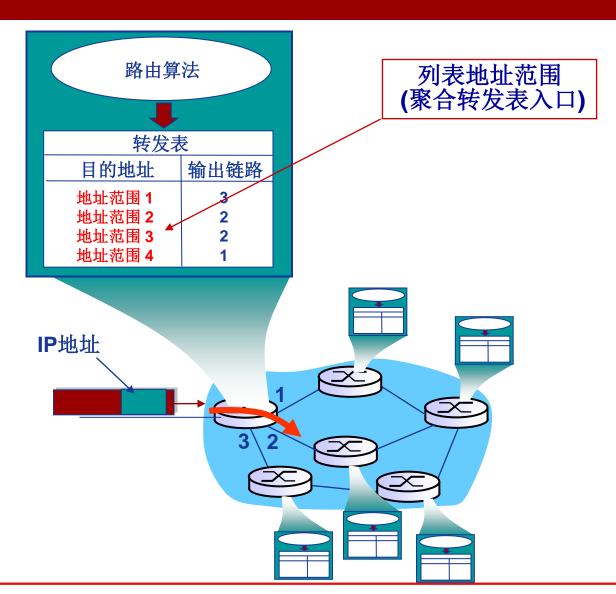


路由算法(协议)确定 通过网络的端到端路径

\_转发表确定在本路 \_由器如何转发分组







目的地址范围	链路接口
11001000 00010111 00010 <mark>000 000000000</mark> 至 11001000 00010111 00010 <mark>111 11111111</mark>	0
11001000 00010111 000110 <mark>00 00000000</mark> 至 11001000 00010111 000110 <mark>11 11111111</mark>	1
11001000 00010111 000111 <mark>00 00000000</mark> 至 11001000 00010111 000111 <mark>11 11111111</mark>	2
其他	3

# 好据报转发表

11001000 00010111 00010000 00000000 11001000 00010111 00011000 00000000 11001000 00010111 00010111 11111111 11001000 00010111 00011011 11111111 11001000 00010111 00011100 00000000 其他 11001000 00010111 00011111 11111111

Q: 如果地址范围划分的不是这么"完美"会怎么样?



# 最长前缀匹配优先

#### 例如:

	目的地	链路接口		
11001000	00010111	00010***	*****	0
11001000	00010111	00011000	*****	1
11001000	00010111	00011***	*****	2
其他				3

DA: 11001000 00010111 00010<mark>110 10100001</mark>

从哪个接口转发? A:0

DA: 11001000 00010111 00011000 10101010

从哪个接口转发? A:1

#### 最长前缀匹配优先

在检索转发表时,优先选择与分组目的地址匹配<mark>前缀最长</mark>的入口(entry)。



## 数据报网络 or VC网络?

#### Internet (数据报网络)

- \* 计算机之间的数据交换
  - "弹性"服务,没有严格 时间需求
- \* 链路类型众多
  - 特点、性能各异
  - 统一服务困难
- \* "智能"端系统(计算机)
  - 可以自适应、性能控制、 差错恢复
- ❖ 简化网络, 复杂"边缘"

#### ATM (VC网络)

- \* 电话网络演化而来
- \*核心业务是实时对话:
  - 严格的时间、可靠性需 求
  - 需要有保障的服务
- ❖ "哑(dumb)"端系统 (非智能)
  - 电话机
  - 传真机
- ❖ 简化"边缘", 复杂网络









主讲人: 李全龙

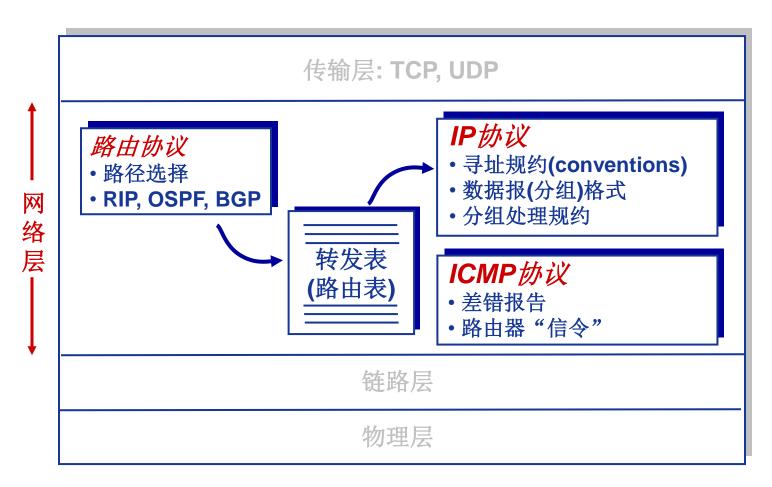
## 本讲主题

### IP协议(1)-IP数据报

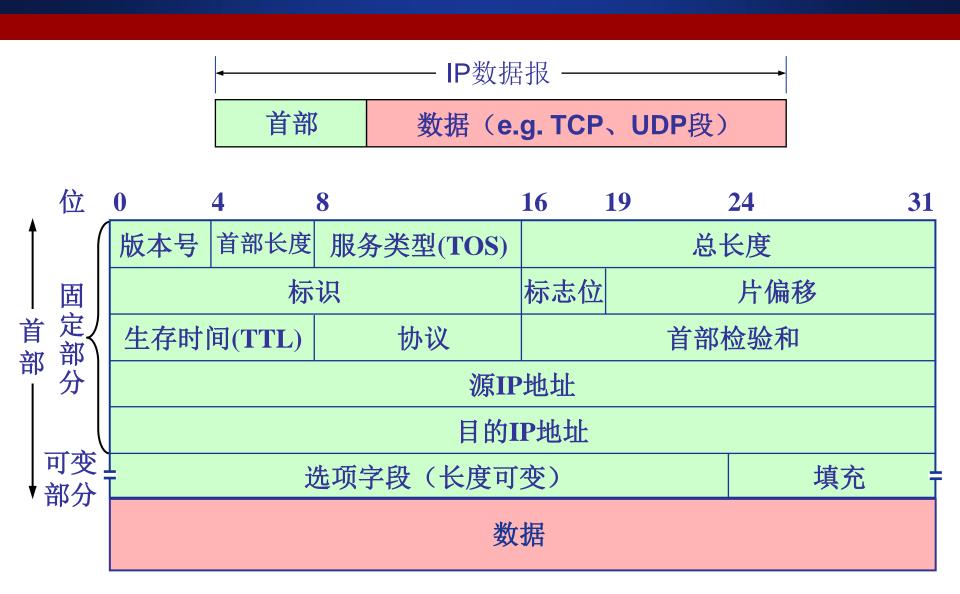


## Internet网络层

主机、路由器网络层主要功能:



# IP数据报(分组)格式





位 8 **19** 0 16 24 31 版本号 首部长度 服务类型(TOS) 总长度 标志位 标识 片偏移 生存时间(TTL) 协议 首部检验和 源IP地址 目的IP地址 选项字段(长度可变) 填充 数据

- **※ 版本号**字段占4位: IP协议的版本号
  - E.g. 4→IPv4, 6 → IPv6



位 8 0 16 **19** 24 31 版本号 首部长度 服务类型(TOS) 总长度 标志位 标识 片偏移 生存时间(TTL) 协议 首部检验和 源IP地址 目的IP地址 选项字段(长度可变) 填充 数据

- ❖ 首部长度字段占4位: IP分组首部长度
  - 以4字节为单位
  - E.g. 5→IP首部长度为20(5×4)字节



位 8 16 **19** 24 31 版本号 首部长度 服务类型(TOS) 总长度 标志位 标识 片偏移 生存时间(TTL) 协议 首部检验和 源IP地址 目的IP地址 选项字段(长度可变) 填充 数据

- ❖ 服务类型(TOS)字段占8位: 指示期望获得哪种类型的服务
  - 1998年这个字段改名为区分服务
  - 只有在网络提供区分服务(DiffServ)时使用
  - 一般情况下不使用,通常IP分组的该字段(第2字节)的值为00H



位 8 16 **19** 24 31 版本号 首部长度 服务类型(TOS) 总长度 标志位 标识 片偏移 生存时间(TTL) 协议 首部检验和 源IP地址 目的IP地址 选项字段(长度可变) 填充 数据

- ❖ 总长度字段占16位: IP分组的总字节数(首部+数据)
  - 最大IP分组的总长度: 65535B
  - 最小的IP分组首部: 20B
  - IP分组可以封装的最大数据: 65535-20=65515B



位 8 16 **19** 24 31 首部长度 版本号 服务类型(TOS) 总长度 标志位 片偏移 标识 生存时间(TTL) 协议 首部检验和 源IP地址 目的IP地址 选项字段(长度可变) 填充 数据

- ❖ 生存时间(TTL)字段占8位: IP分组在网络中可以通过的 路由器数 (或跳步数)
  - 路由器转发一次分组,TTL减1
  - 如果TTL=0,路由器则丢弃该IP分组



位 8 0 16 **19** 24 31 版本号 首部长度 服务类型(TOS) 总长度 标志位 片偏移 标识 生存时间(TTL) 协议 首部检验和 源IP地址 目的IP地址 选项字段(长度可变) 填充 数据

- ❖ 协议字段占8位: 指示IP分组封装的是哪个协议的数据包
  - 实现复用/分解
  - E.g. 6为TCP,表示封装的为TCP段;17为UDP,表示封装的是UDP数据报



位 8 0 16 **19** 24 31 首部长度 版本号 服务类型(TOS) 总长度 标志位 标识 片偏移 生存时间(TTL) 协议 首部检验和 源IP地址 目的IP地址 选项字段(长度可变) 填充 数据

- ❖ 首部校验和字段占16位:实现对IP分组首部的差错检测
  - 计算校验和时,该字段置全0
  - 采用反码算数运算求和,和的反码作为首部校验和字段
  - 逐跳计算、逐跳校验



位 8 0 16 **19** 24 31 版本号 首部长度 服务类型(TOS) 总长度 标志位 片偏移 标识 生存时间(TTL) 协议 首部检验和 源IP地址 目的IP地址 选项字段(长度可变) 填充 数据

※源IP地址、目的IP地址字段各占32位:分别标识发送分组的源主机/路由器(网络接口)和接收分组的目的主机/路由器(网络接口)的IP地址



位 8 0 16 **19** 24 31 版本号 首部长度 服务类型(TOS) 总长度 标志位 片偏移 标识 生存时间(TTL) 协议 首部检验和 源IP地址 目的IP地址 选项字段(长度可变) 填充 数据

- ❖ 选项字段占长度可变,范围在1~40B之间:携带安全、源 选路径、时间戳和路由记录等内容
  - 实际上很少被使用



位 8 0 16 **19** 24 31 首部长度 版本号 服务类型(TOS) 总长度 标志位 片偏移 标识 生存时间(TTL) 协议 首部检验和 源IP地址 目的IP地址 选项字段(长度可变) 填充 数据

❖填充字段占长度可变,范围在0~3B之间:目的是补齐整个首部,符合32位对齐,即保证首部长度是4字节的倍数







主讲人: 李全龙

## 本讲主题

### IP协议(2)-IP分片

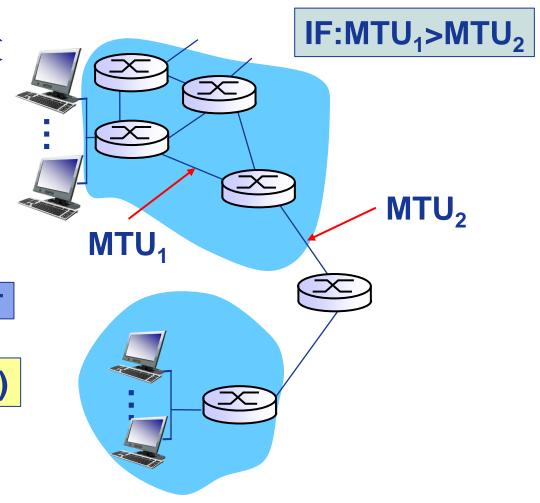
# 最大传输单元(MTU)

❖ 网络链路存在MTU (最大传输单元)—链路层数据帧可封装数据的上限

■ 不同链路的MTU不同

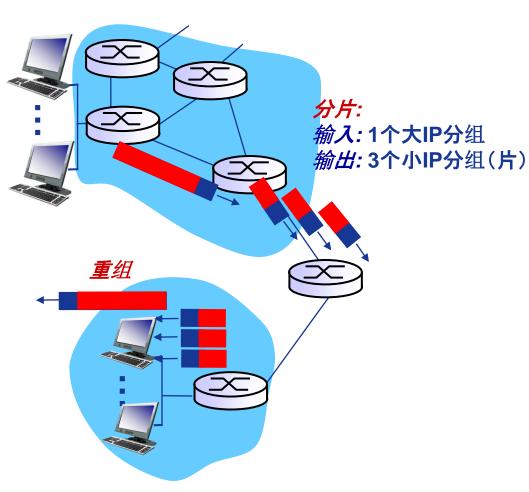


MTU=Max(Data)



## IP分片与重组

- ❖ 大IP分组向较小MTU链路转发时,可以被"分片" (fragmented)
  - 1个IP分组分为多片IP分组
  - IP分片到达目的主机后进 行"重组" (reassembled)
- ❖IP首部的相关字段用 于标识分片以及确定 分片的相对顺序
  - 总长度、标识、标志 位和片偏移



位 8 **19** 0 16 24 31 版本号 首部长度 服务类型(TOS) 总长度 标志位 片偏移 标识(ID) 生存时间(TTL) 协议 首部检验和 源IP地址 目的IP地址 选项字段(长度可变) 填充 数据

- ❖ 标识字段占16位:标识一个IP分组
  - IP协议利用一个计数器,每产生IP分组计数器加1,作为该IP分组的标识



位 8 0 16 **19** 24 31 版本号 首部长度 服务类型(TOS) 总长度 标志位 片偏移 标识(ID) 生存时间(TTL) 协议 首部检验和 源IP地址 目的IP地址 选项字段(长度可变) 填充 数据



- DF (Don't Fragment)
- MF (More Fragment)



- **DF** =1: 禁止分片;
  - **DF** =0: 允许分片
- MF =1: 非最后一片;
  - MF =0: 最后一片(或未分片)



位 8 0 16 **19** 24 31 首部长度 版本号 服务类型(TOS) 总长度 标志位 片偏移 标识(ID) 生存时间(TTL) 协议 首部检验和 源IP地址 目的IP地址 选项字段(长度可变) 填充 数据

- ❖ 片偏移字段占13位: 一个IP分组分片封装原IP分组数据的 相对偏移量
  - 片偏移字段以8字节为单位



## IP分片过程

- ❖ 假设原IP分组总长度为L,待转发链路的MTU为M
- **❖** 若*L>M*,且DF=0,则可以/需要分片
- \* 分片时每个分片的标识复制原IP分组的标识
- ❖ 通常分片时,除最后一个分片,其他分片均分为MTU允许的最大分片
- ❖ 一个最大分片可封装的数据应该是8的倍数,因此,一个 最大分片可封装的数据为:

$$d = \left| \frac{M - 20}{8} \right| \times 8$$

\* 需要的总片数为:

$$n = \left\lceil \frac{L - 20}{d} \right\rceil$$





### IP分片过程

❖ 每片的片偏移字段取值为:

$$F_i = \frac{d}{8} \times (i-1), \qquad 1 \le i \le n$$

\*每片的总长度字段为:

$$L_i = \begin{cases} d + 20 & 1 \le i < n \\ L - (n-1)d & i = n \end{cases}$$

❖ 每片的MF标志位为:

$$\mathbf{MF}_{i} = \begin{cases} 1 & 1 \le i < n \\ 0 & i = n \end{cases}$$

## IP分片过程

#### 例如:

- \* 4000B数据报
- 输出链路MTU = 1500B
- **⋄** DF=0



#### 1个大数据报分片为3个小数据报(片)

片偏移 = 1480/8 | length | ID | MF=1 | offset | =185

length	ID	MF=0	offset	
=1040	= <b>x</b>		=370	





主讲人: 李全龙

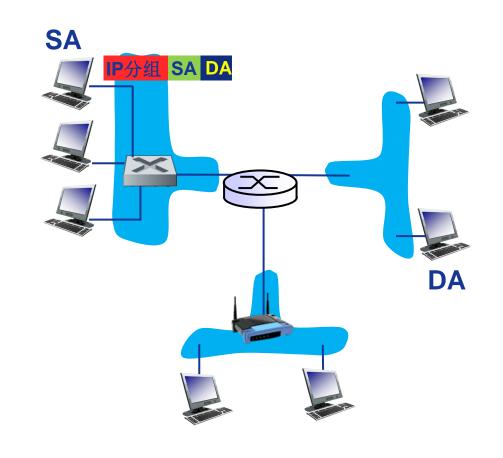
## 本讲主题

### IP协议(3)-IP编址

# IP编址(addressing)

#### **❖ IP**分组:

- 源地址(SA)-从哪儿来
- 目的地址(DA)-到哪儿去
- ❖ 接口(interface): 主机/路 由器与物理链路的连接
  - 实现网络层功能
  - 路由器通常有多个接口
  - 主机通常只有一个或两个 接口 (e.g.,有线的以太网 接口,无线的802.11接口)



# IP编址(addressing)

223.1.1.1

❖ IP地址: 32比特(IPv4) 编号标识主机、路由 器的接口

11011111 00000001 00000001 =223.1.1.1 223 1 1 1

❖ IP地址与每个接口关联

223.1.2.1 223.1.1.2 223.1.2.9 223.1.3.27 223.1.1.3 223.1.3.1 223.1.3.2

❖ 怎样为接口分配IP地址 呢?

### IP子网 (Subnets)

#### ❖IP地址:

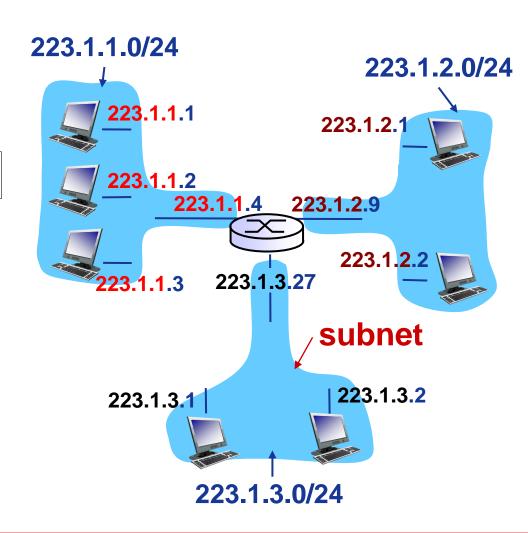
- 网络号(NetID) 高位比特
- ■主机号(HostID) 低位比特

**NetID** 

**HostID** 

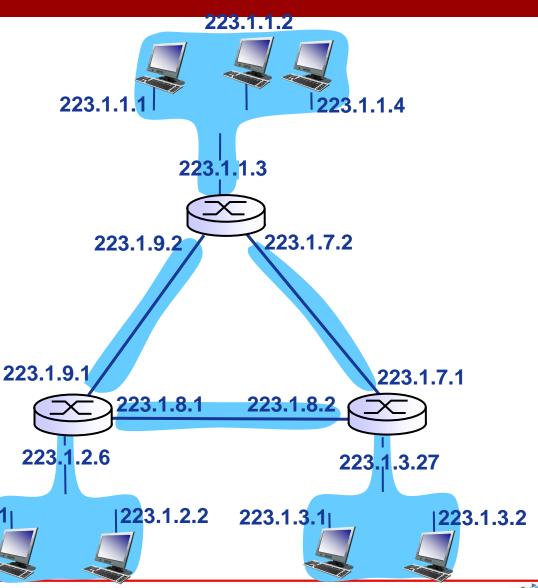
#### ❖IP子网:

- IP地址具有相同网络号的设备接口
- ■不跨越路由器(第三及以 上层网络设备)可以彼此 物理联通的接口



## IP子网 (Subnets)

图中网络有多少个IP子网?





223.1.2.1





主讲人: 李全龙

## 本讲主题

### IP协议(4)-有类IP地址



### IP子网 (Subnets)

#### ❖IP地址:

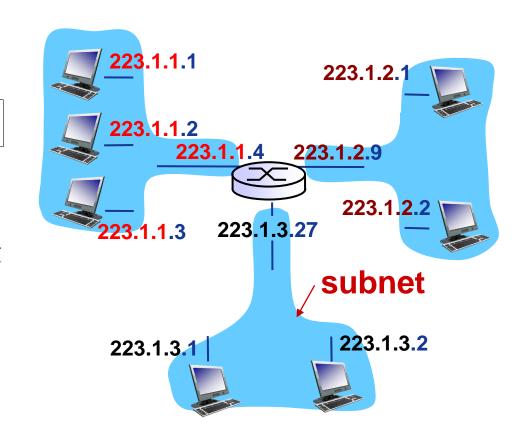
- 网络号(NetID) 高位比特
- ■主机号(HostID) 低位比特

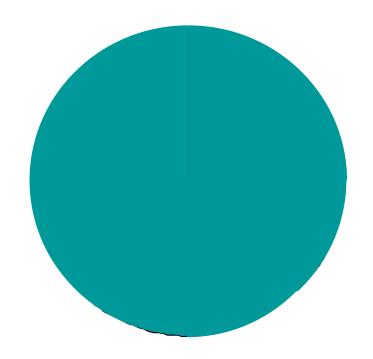
**NetID** 

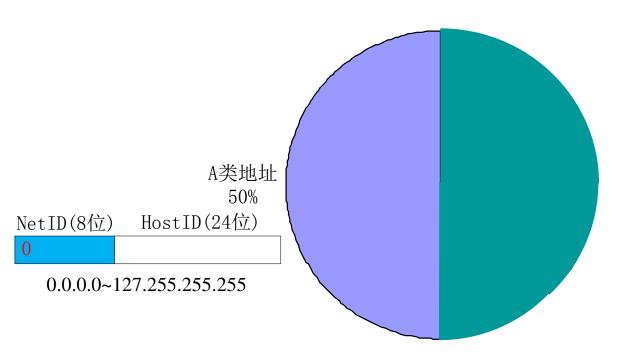
**HostID** 

#### ❖IP子网:

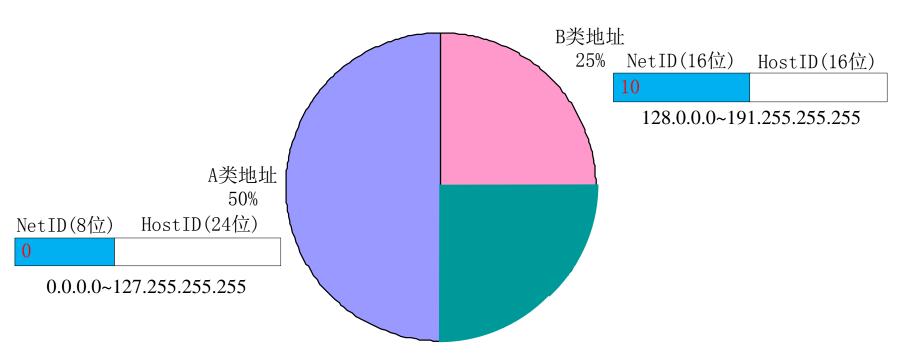
- IP地址具有相同网络号的设备接口
- ■不跨越路由器(第三及以 上层网络设备)可以彼此 物理联通的接口

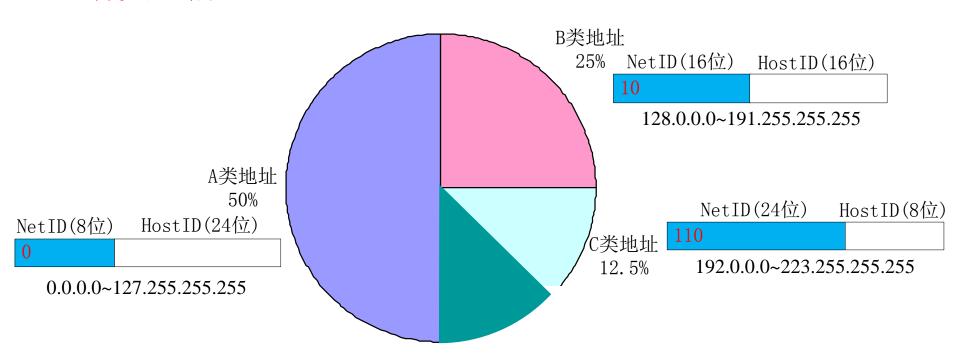


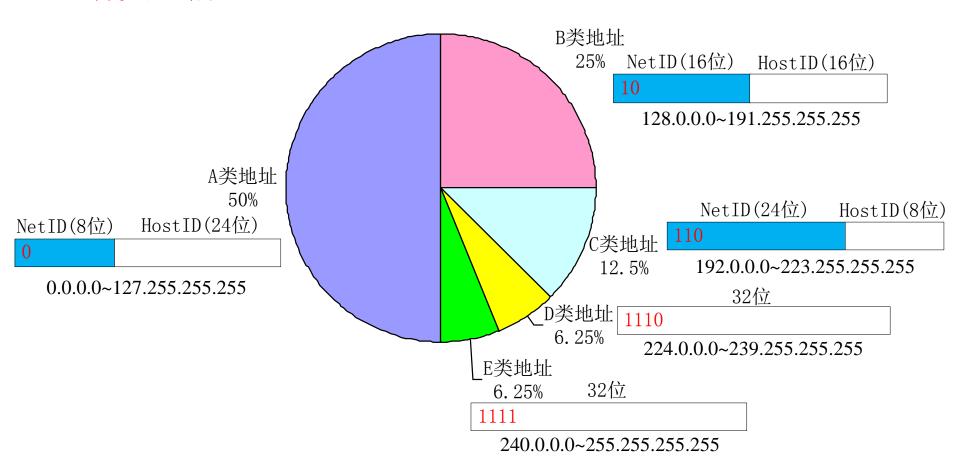














# 特殊IP地址

NetID	HostID	作为IP分组 源地址	作为IP分组 目的地址	用途
全0	全0	可以	不可以	在本网范围内表示本机;在 路由表中用于表示默认路由 (相当于表示整个Internet网络)
全0	特定值	不可以	可以	表示本网内某个特定主机
全1	全1	不可以	可以	本网广播地址(路由器不转发)
特定值	全0	不可以	不可以	网络地址,表示一个网络
特定值	全1	不可以	可以	直接广播地址,对特定网络上的所有主机进行广播
127	非全 <b>0</b> 或 非全 <b>1</b> 的 任何数	可以	可以	用于本地软件环回测试,称为环回地址



## 私有(Private)IP地址

Class	NetIDs	Blocks
A	10	1
В	172.16 to 172.31	16
C	192.168.0 to 192.168.255	256





主讲人: 李全龙

### 本讲主题

IP协议(5)-IP子网划分与子网掩码



#### IP子网 (Subnets)

#### ❖IP地址:

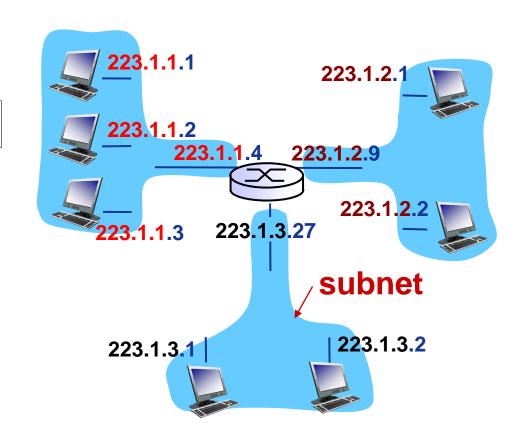
- 网络号(NetID) 高位比特
- ■主机号(HostID) 低位比特

**NetID** 

**HostID** 

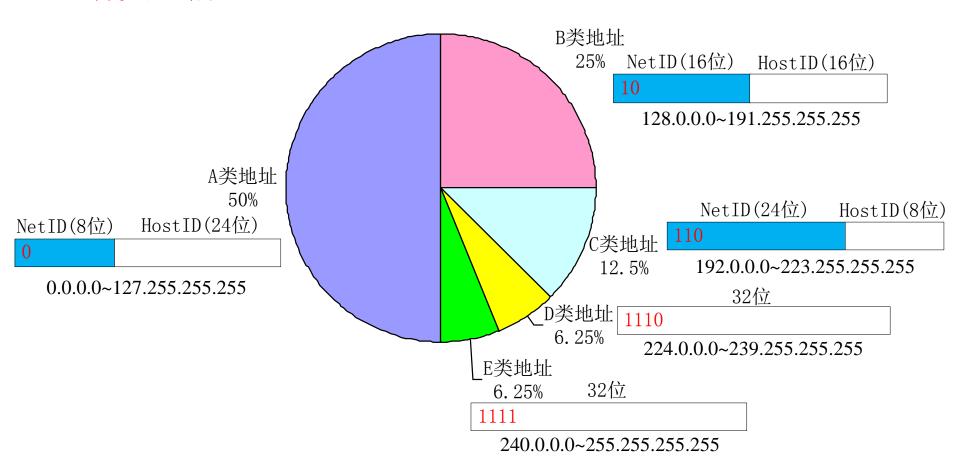
#### ❖IP子网:

- IP地址具有相同网络号的设备接口
- ■不跨越路由器(第三及以 上层网络设备)可以彼此 物理联通的接口



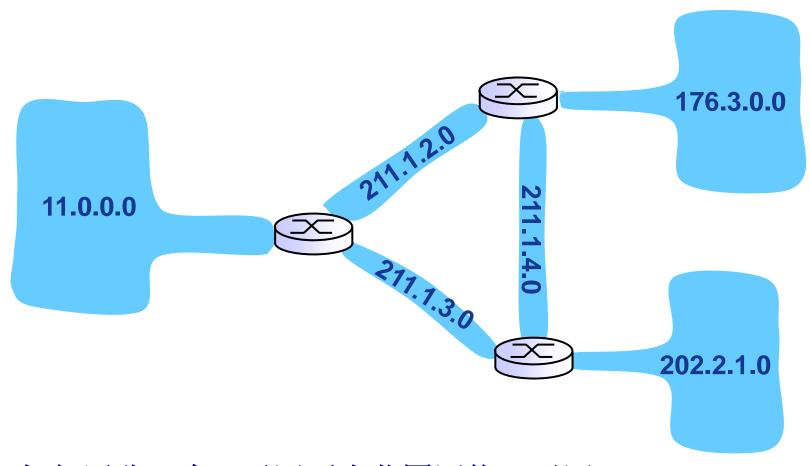
# IP地址(Addresses)

#### "有类"编址:





# 子网划分(Subnetting)?



- ❖如何区分一个IP子网更小范围网络(子网)?
  - 子网划分

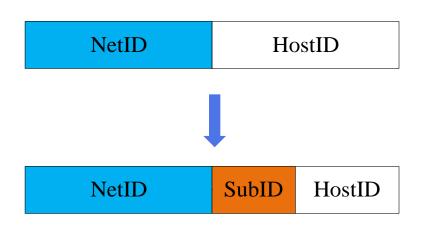


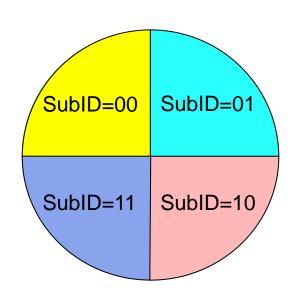


# 子网划分?

#### ❖IP地址:

- 网络号(NetID) 高位比特
- ■子网号(SubID) 原网络主机号部分比特
- ■主机号(HostID) 低位比特

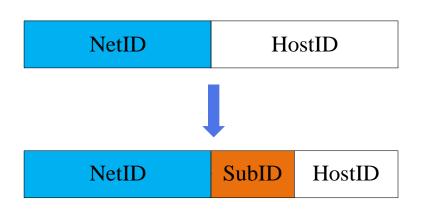




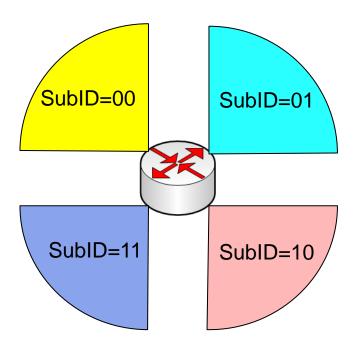
### 子网划分?

#### ❖IP地址:

- 网络号(NetID) 高位比特
- ■子网号(SubID) 原网络主机号部分比特
- 主机号(HostID) 低位比特



- \* 如何确定是否划分了子网? 利用多 少位划分子网?
  - 子网掩码



### 子网掩码

**NetID** 

- ❖形如IP地址:
  - 32位
  - ■点分十进制形式
- 取值:
  - NetID、SubID位全取1
  - HostID位全取0

# 子网地址+子网掩码

**HostID** 

SubID

→准确确定子网大小

#### ❖例如:

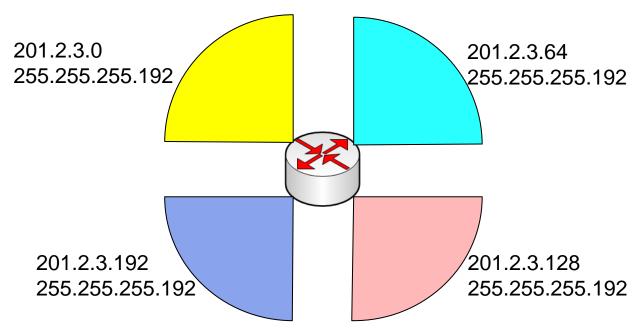
- ■A网的默认子网掩码为: 255.0.0.0
- ■B网的默认子网掩码为: 255.255.0.0
- ■C网的默认子网掩码为: 255.255.255.0
- 借用3比特划分子网的B网的子网掩码为: 255.255.224.0



### 子网划分

#### ❖例如:

■子网201.2.3.0, 255.255.255.0, 划分为等长的4个子网



❖路由器如何确定应该将IP分组转发到哪个子网?



### 子网掩码的应用

\*将IP分组的目的IP地址与子网掩码按位与运算, 提取子网地址

#### ❖例如:

■目的IP地址: 172.32.1.112, 子网掩码: 255.255.254.0

```
172. 32. 1. 112= 10101100 00100000 00000001 01110000
10101100 00100000 00000000 00000000
            172
                  32
```

- ■子网地址: 172.32.0.0(子网掩码: 255.255.254.0)
- 地址范围: 172.32.0.0~172.32.1.255
- ■可分配地址范围: 172.32.0.1~172.32.1.254
- ■广播地址: 172.32.1.255



# 一个C类网络划分子网举例

子网	SubID (二进制)	HostID取值范围 (二进制)	第4八位组取值范围 (十进制)
1#	000	00000 thru 11111	.0 thru .31
2#	001	00000 thru 11111	.32 thru .63
3#	010	00000 thru 11111	.64 thru .95
4#	011	00000 thru 11111	.96 thru .127
5#	100	00000 thru 11111	.128 thru .159
6#	101	00000 thru 11111	.160 thru .191
7#	110	00000 thru 11111	.192 thru .223
8#	111	00000 thru 11111	.224 thru .255





# 一个C类网络划分子网举例

子网	SubID (二进制)	HostID取值范围 (二进制)	第4八位组取值范围 (十进制)
1#	000	00000 thru 11111	. <mark>0</mark> thru . <b>31</b>
2#	001	00000 thru 11111	.32 thru .63
3#	010	00000 thru 11111	.64 thru .95
4#	011	00000 thru 11111	.96 thru .127
5#	100	00000 thru 11111	.128 thru .159
6#	101	00000 thru 11111	.160 thru .191
7#	110	00000 thru 11111	.192 thru .223
8#	111	00000 thru 11111	.224 thru .255







主讲人: 李全龙

### 本讲主题

### IP协议(6)-CIDR与路由聚合



#### CIDR

无类域间路由(CIDR: Classless InterDomain Routing)

- 消除传统的A类、B类和C类地址界限
  - NetID+SubID→Network Prefix (Prefix)可以任意长度
- 融合子网地址与子网掩码,方便子网划分
  - 无类地址格式: a.b.c.d/x,其中x为前缀长度
- 例如



11001000 00010111 00010000 00000000

200.23.16.0/23

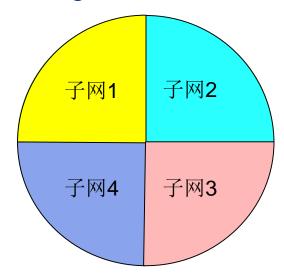
 $\blacksquare$   $\exists$   $\exists$   $\boxtimes$  201.2.3.64, 255.255.255.192 $\rightarrow$  201.2.3.64/26



#### CIDR与路由聚合

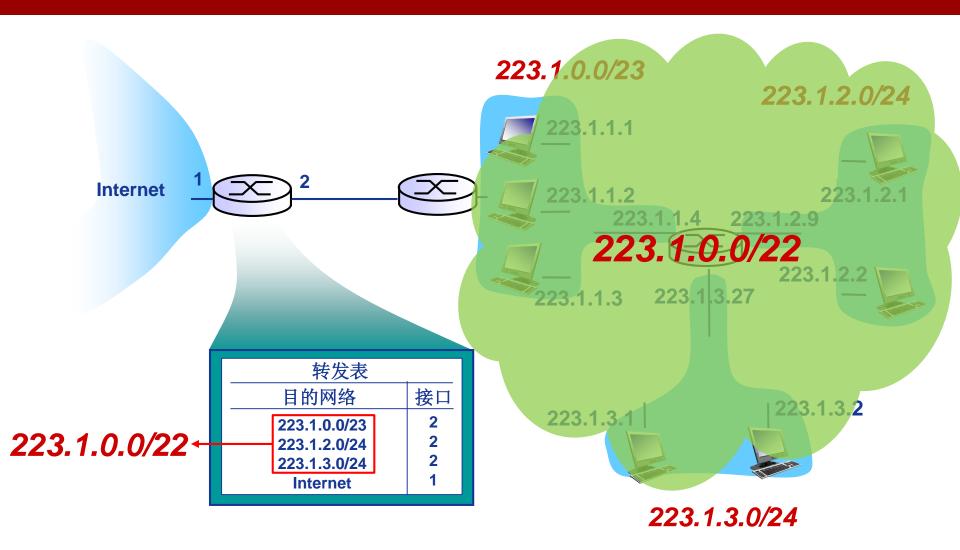
#### 无类域间路由(CIDR: Classless InterDomain Routing)

- 提高IPv4 地址空间分配效率
- 提高路由效率
  - 将多个子网聚合为一个较大的子网
  - 构造超网(supernetting)

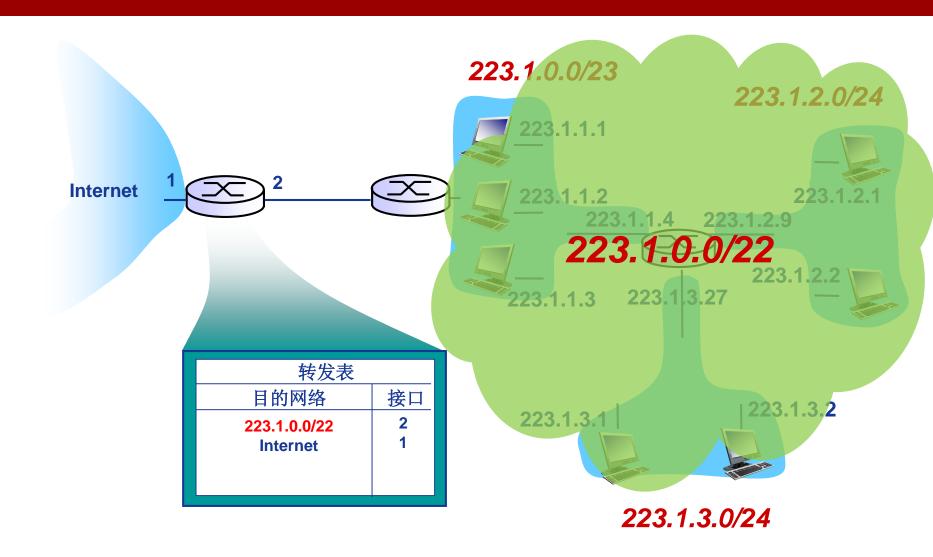


路由聚合(route aggregation)

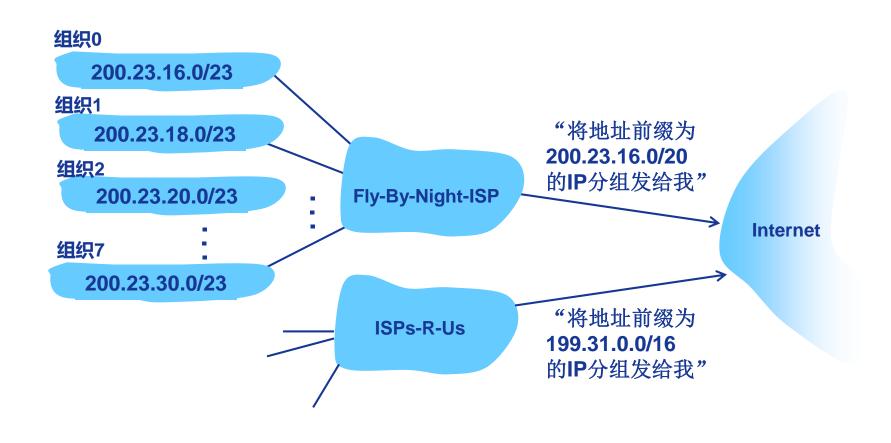






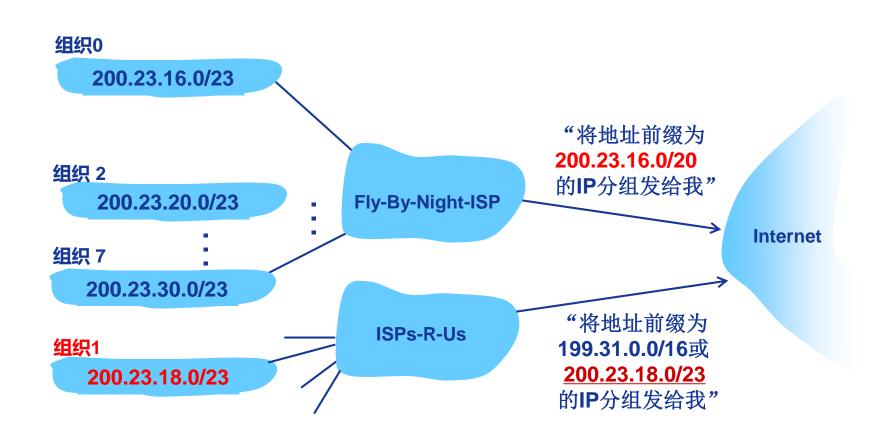


#### 层级编址使得路由信息通告更高效:





#### 选用更具体的路由: 最长前缀匹配优先!









主讲人: 李全龙

# 本讲主题

### DHCP协议

#### 如何获得IP地址?

- Q: 一个主机如何获得IP地址?
- ❖ "硬编码"
  - 静态配置



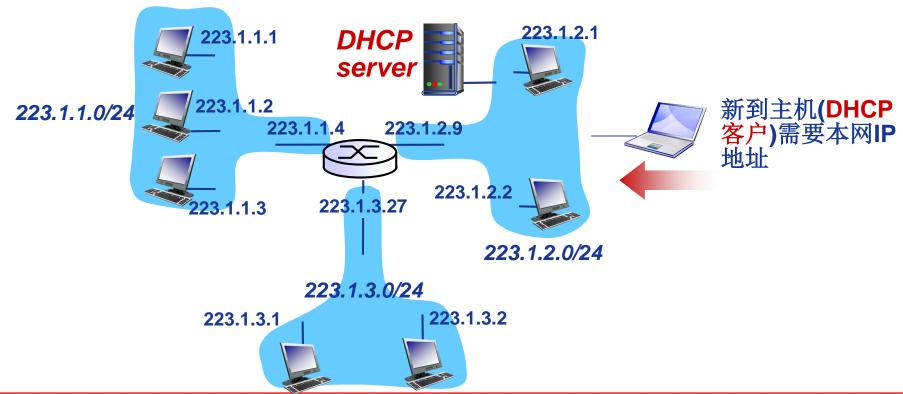
#### 如何获得IP地址?

- Q: 一个主机如何获得IP地址?
- ❖ "硬编码"
  - 静态配置
- ❖ 动态主机配置协议-DHCP: Dynamic Host Configuration Protocol
  - 从服务器动态获取:
    - IP 地址
    - 子网掩码
    - 默认网关地址
    - · DNS服务器名称与IP地址
  - "即插即用"
  - 允许地址重用
  - 支持在用地址续租
  - 支持移动用户加入网络



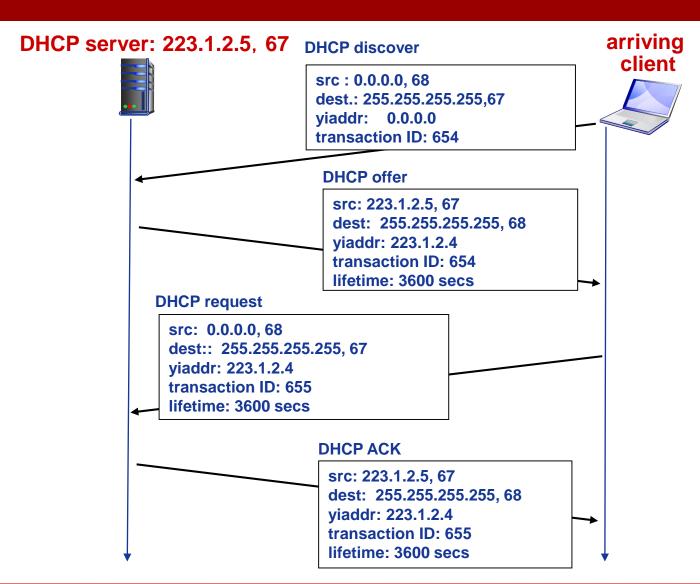
### 动态主机配置协议(DHCP)

- ❖ 主机广播 "DHCP discover" (发现报文)
- ❖ DHCP服务器利用 "DHCP offer" (提供报文) 进行响应
- ❖ 主机请求IP地址: "DHCP request" (请求报文)
- ❖ DHCP服务器分配IP地址: "DHCP ack" (确认报文)



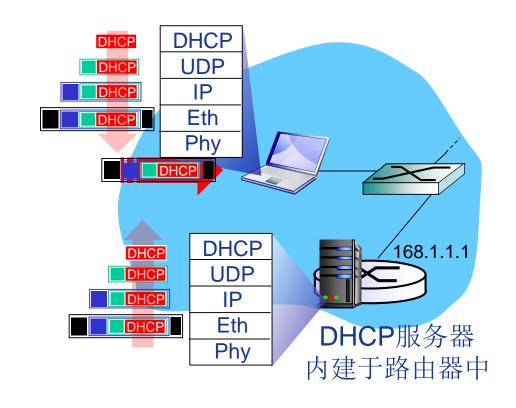


### DHCP工作过程示例



### DHCP工作过程示例

- ❖ DHCP协议在应用层实现
  - 请求报文封装到 UDP数据报中
  - IP广播
  - 链路层广播 (e.g. 以太网广播)



### DHCP工作过程示例

- ❖ DHCP服务器构造 ACK报文
  - 包括分配给客户的 IP地址、子网掩码、默认网关、DNS 服务器地址

