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

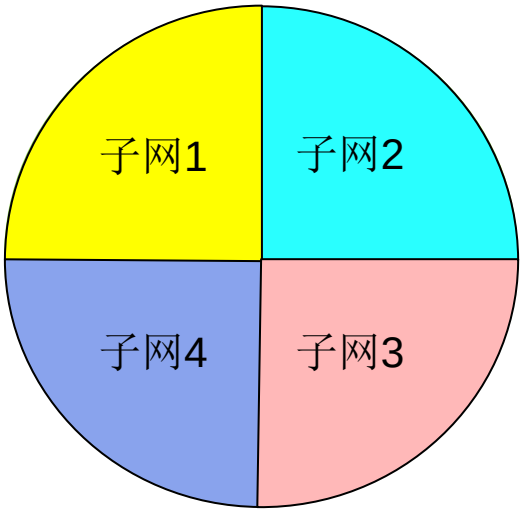
200.23.16.0/23

- 子网201.2.3.64，255.255.255.192→201.2.3.64/26

CIDR与路由聚合

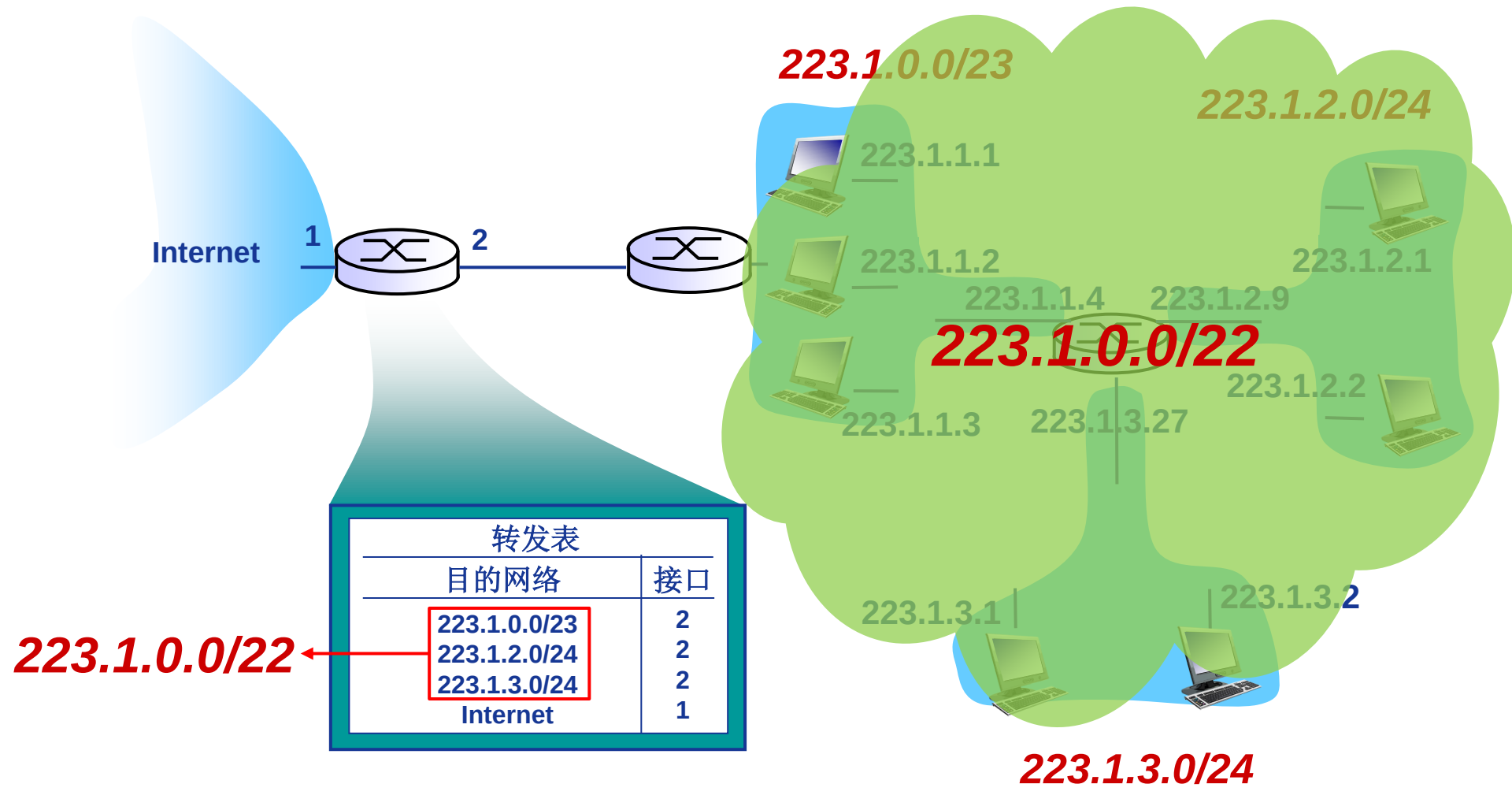
无类域间路由(CIDR: Classless InterDomain Routing)

- 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）

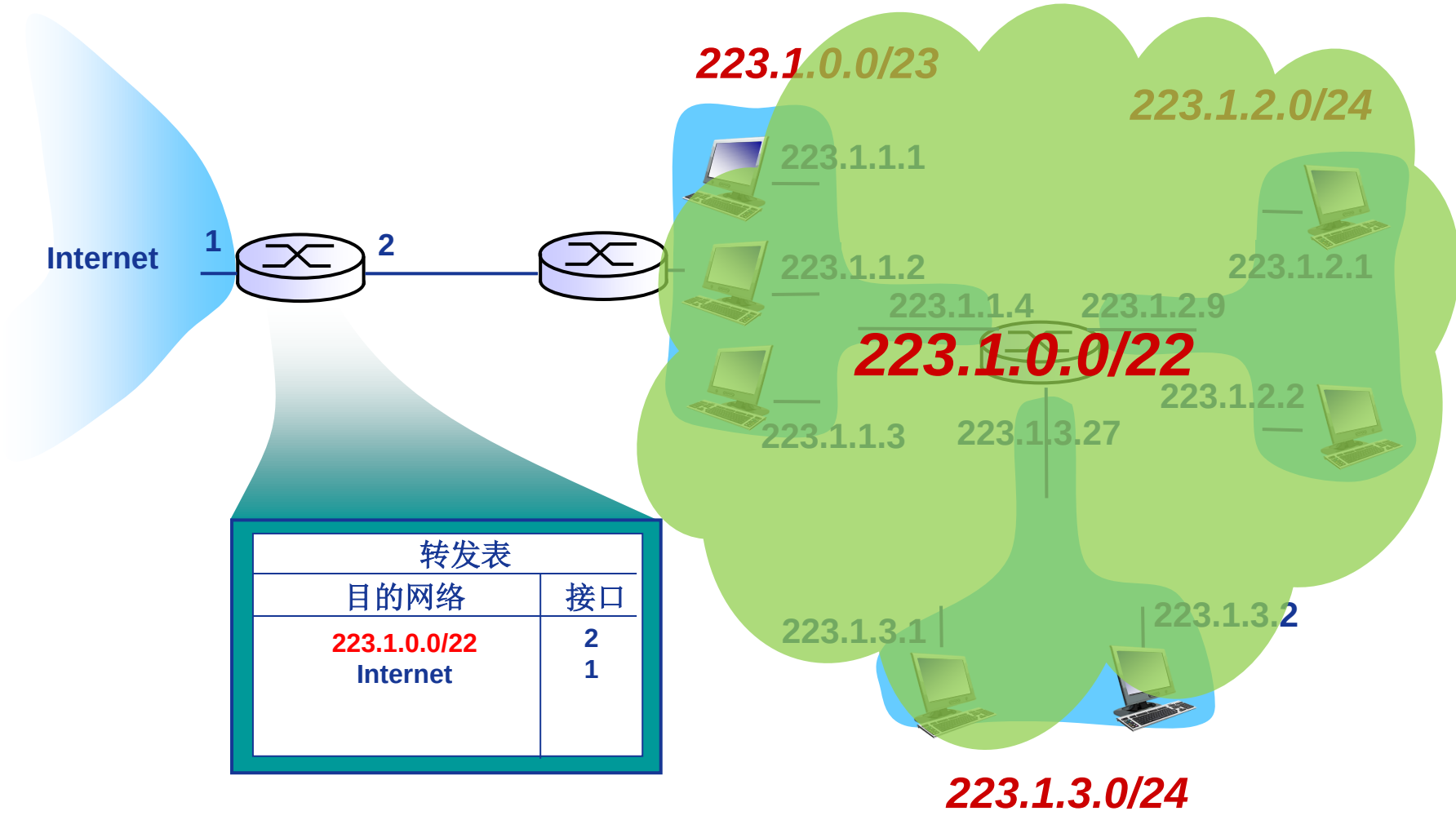


  - 路由聚合（route aggregation）

路由聚合

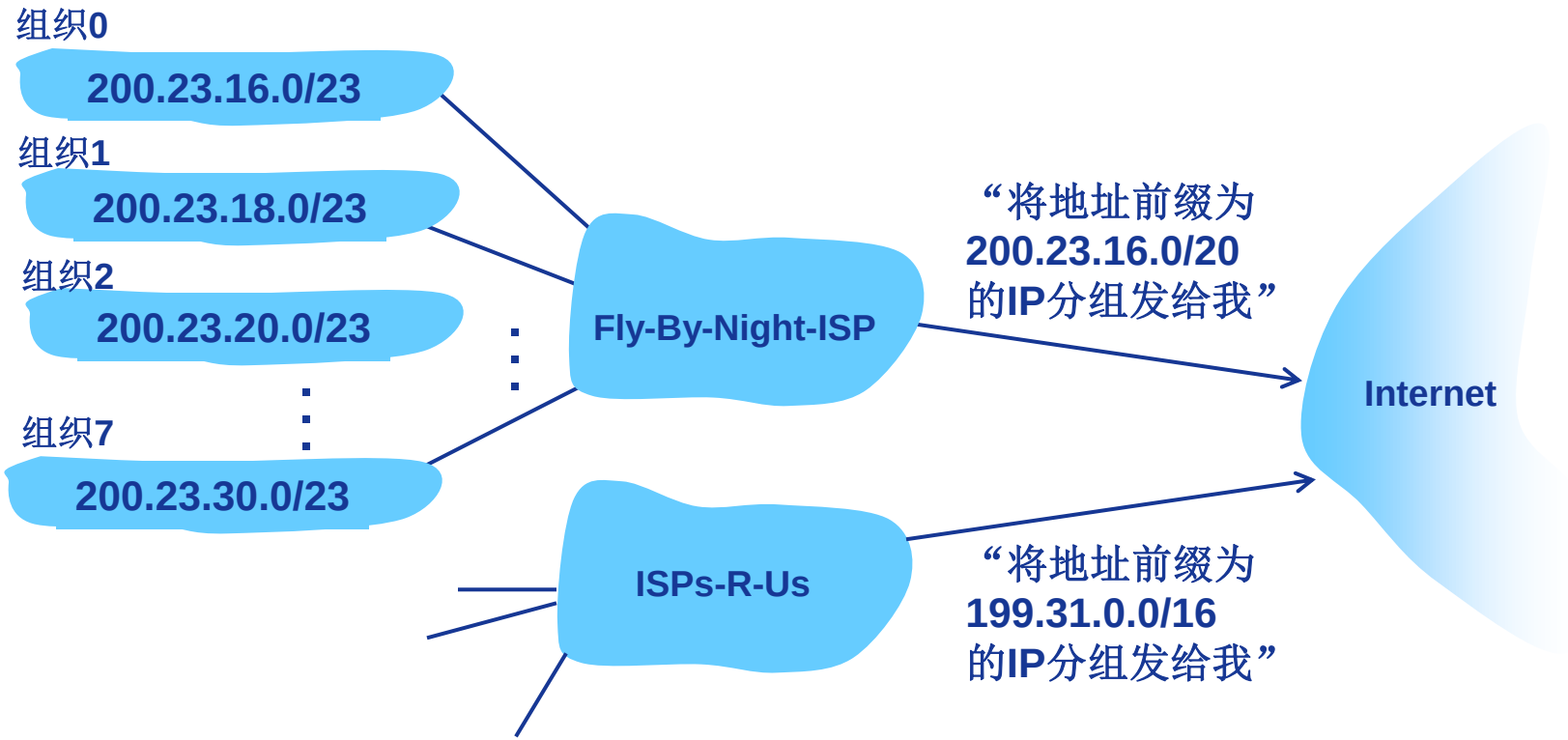

223.1.0.0/23
223.1.2.0/24
223.1.1.1
223.1.1.2
223.1.2.1
223.1.1.4
223.1.2.9
223.1.0.0/22
223.1.2.2
223.1.1.3
223.1.3.27
Internet
1
2
223.1.3.1
223.1.3.2
223.1.3.0/24
转发表
目的网络
接口
223.1.0.0/23
223.1.2.0/24
223.1.3.0/24
Internet
2
2
2
1
223.1.0.0/22

路由聚合


223.1.0.0/23
223.1.2.0/24
223.1.1.1
223.1.1.2
223.1.2.1
223.1.1.4
223.1.2.9
223.1.0.0/22
223.1.2.2
223.1.1.3
223.1.3.27
Internet
1
2
转发表
目的网络
接口
223.1.0.0/22
2
Internet
1
223.1.3.1
223.1.3.2
223.1.3.0/24

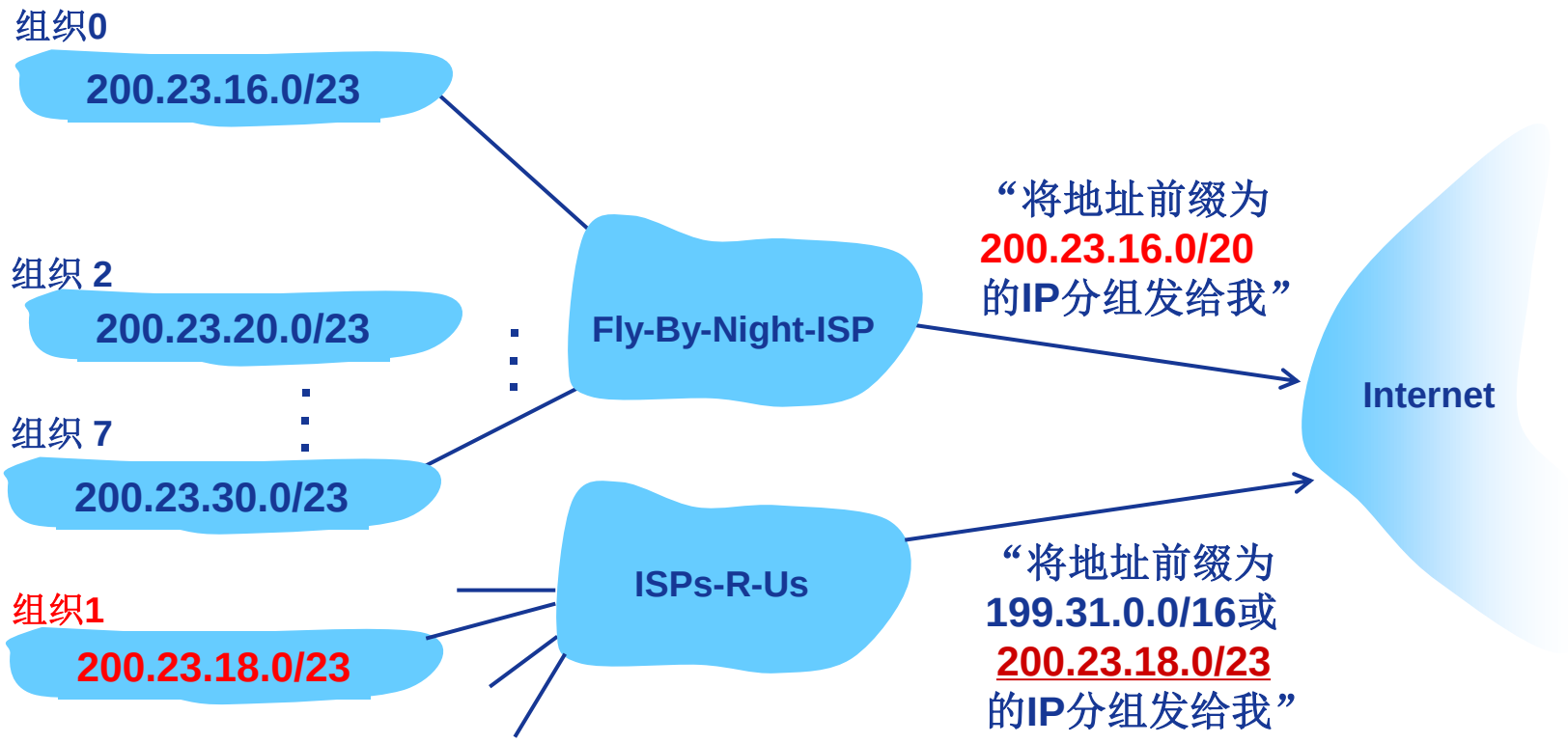
路由聚合

层级编址使得路由信息通告更高效：



路由聚合

选用更具体的路由：最长前缀匹配优先！



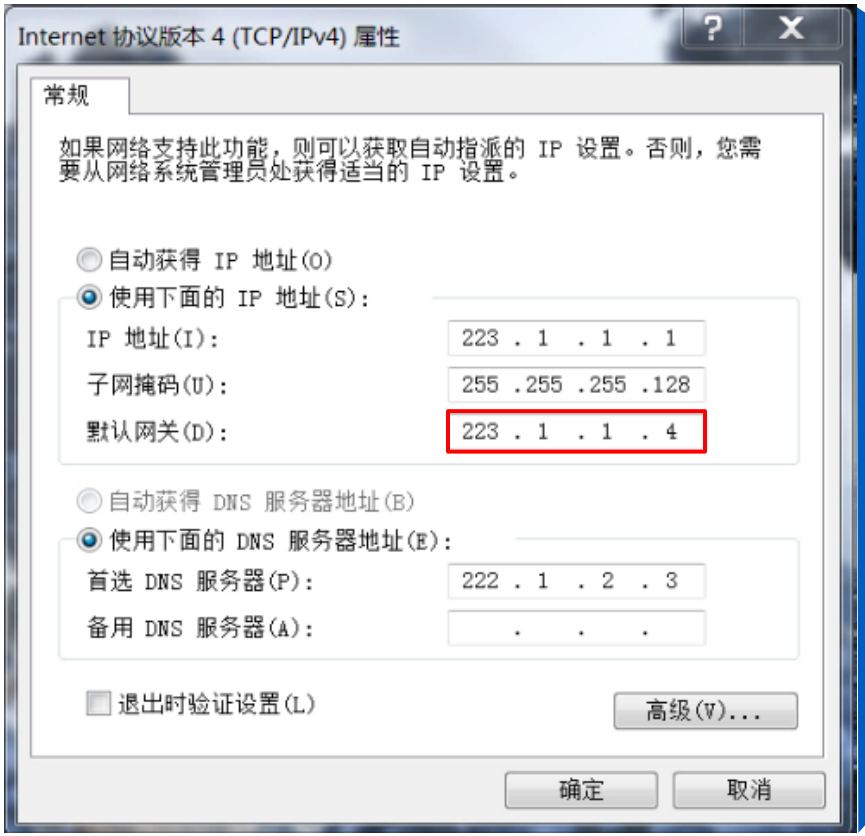
本讲主题

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



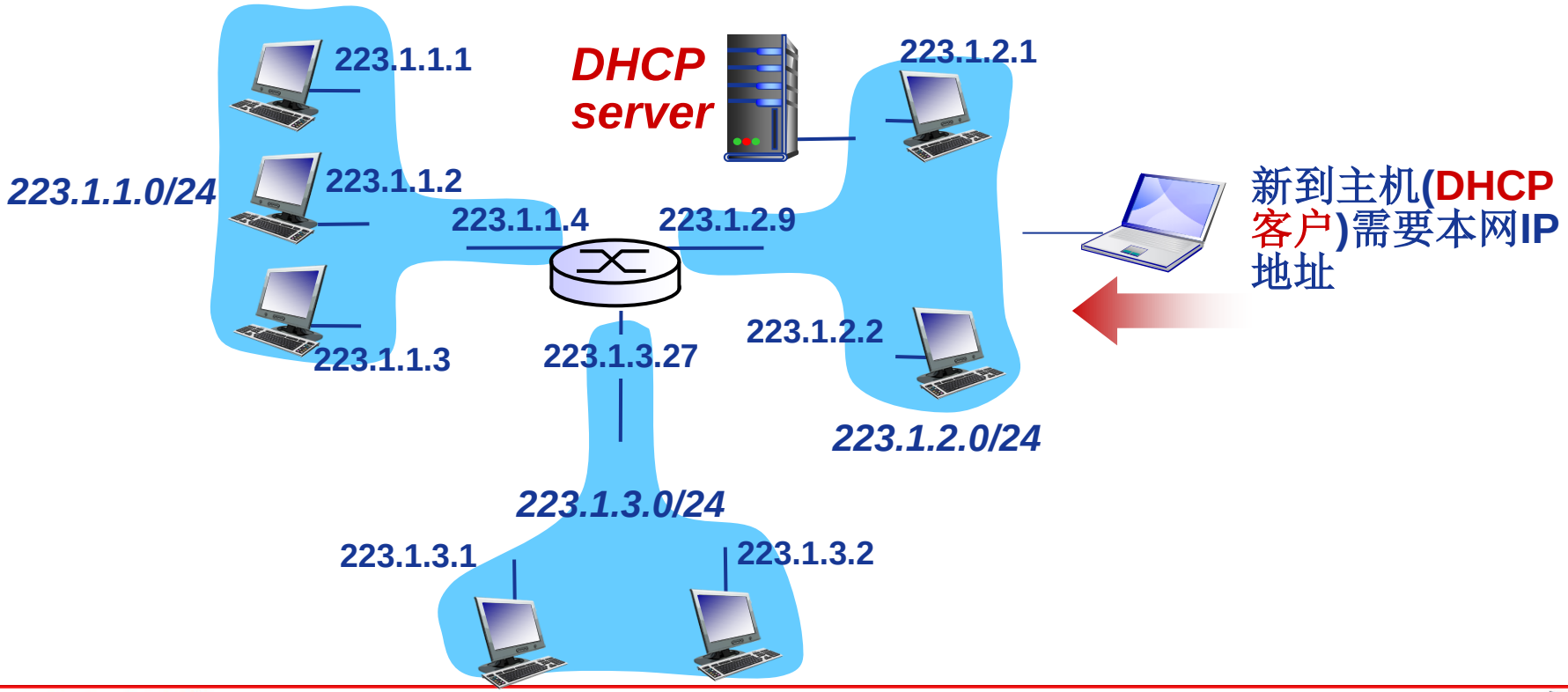
如何获得IP地址?

Q: 一个主机如何获得IP地址？

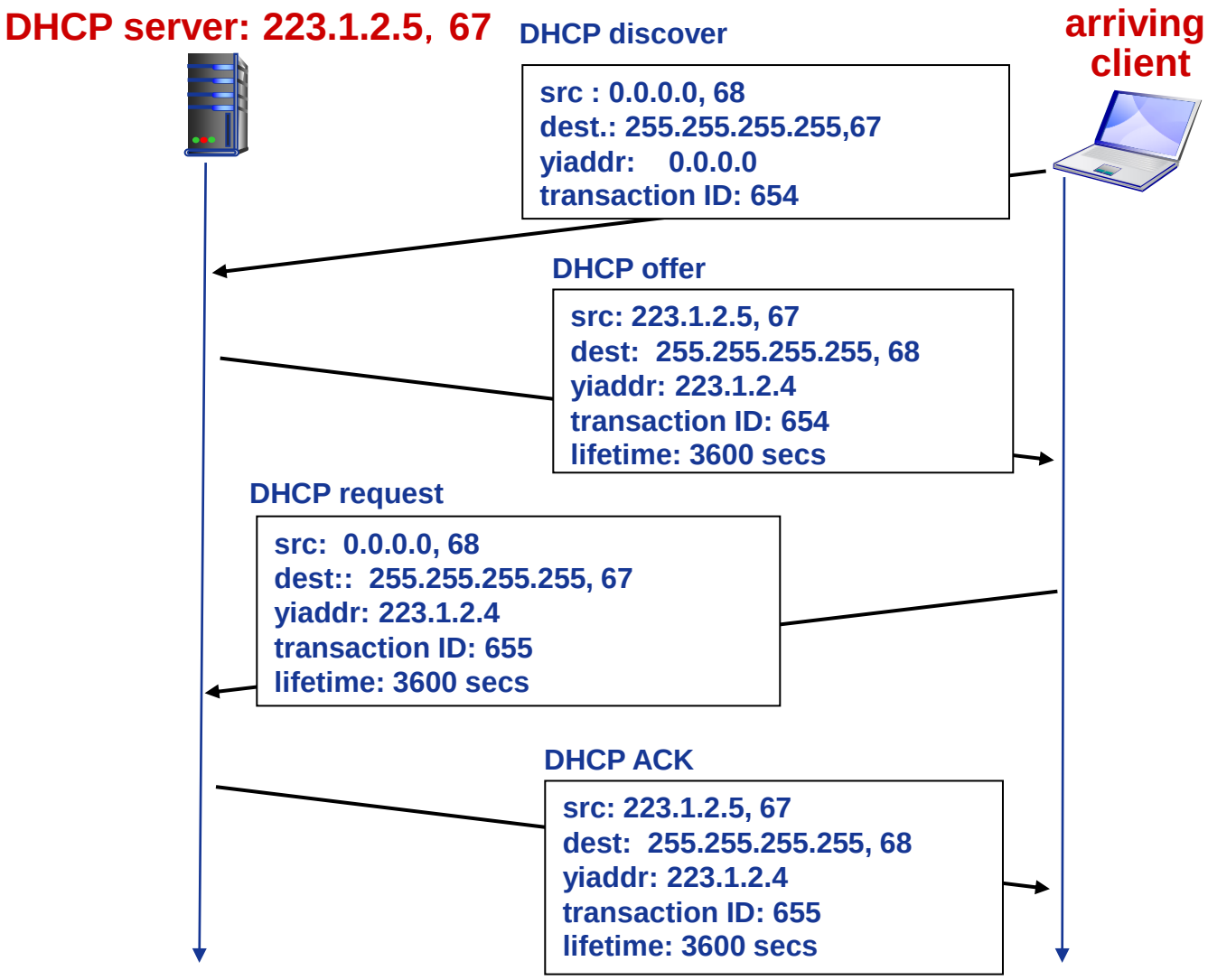
- “硬编码”
  - 静态配置
- 动态主机配置协议-DHCP: Dynamic Host Configuration Protocol
  - 从服务器动态获取：
    - **IP地址**
    - **子网掩码**
    - **默认网关地址**
    - **DNS服务器名称与IP地址**
  - “即插即用”
  - 允许地址重用
  - 支持在用地址续租
  - 支持移动用户加入网络

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)

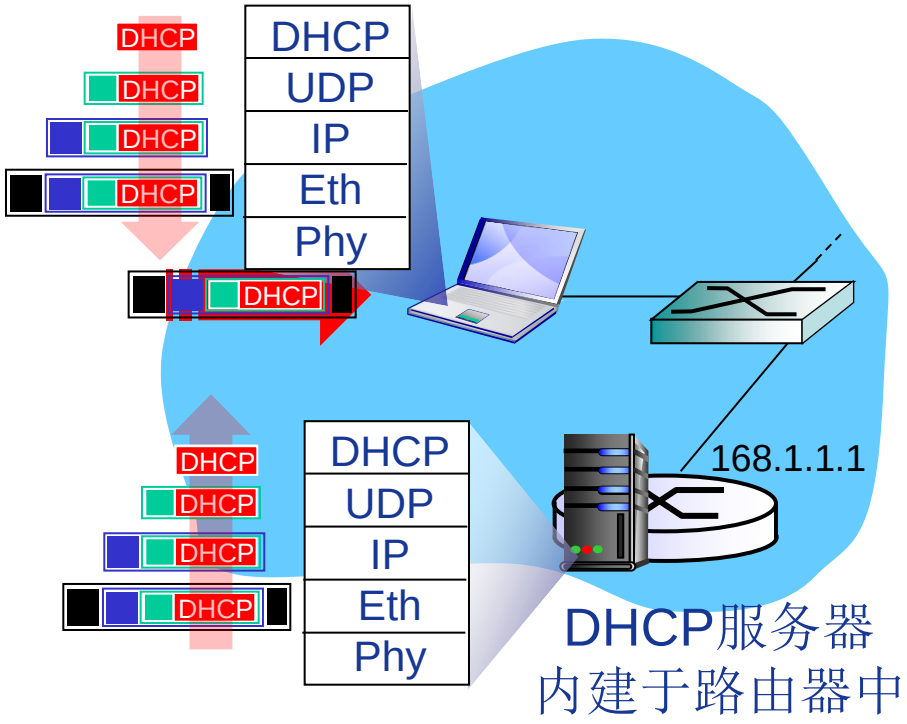


DHCP工作过程示例


DHCP server: 223.1.2.5, 67
DHCP discover
arriving client
src : 0.0.0.0, 68
dest.: 255.255.255.255,67
yiaddr: 0.0.0.0
transaction ID: 654
DHCP offer
src: 223.1.2.5, 67
dest: 255.255.255.255, 68
yiaddr: 223.1.2.4
transaction ID: 654
lifetime: 3600 secs
DHCP request
src: 0.0.0.0, 68
dest:: 255.255.255.255, 67
yiaddr: 223.1.2.4
transaction ID: 655
lifetime: 3600 secs
DHCP ACK
src: 223.1.2.5, 67
dest: 255.255.255.255, 68
yiaddr: 223.1.2.4
transaction ID: 655
lifetime: 3600 secs

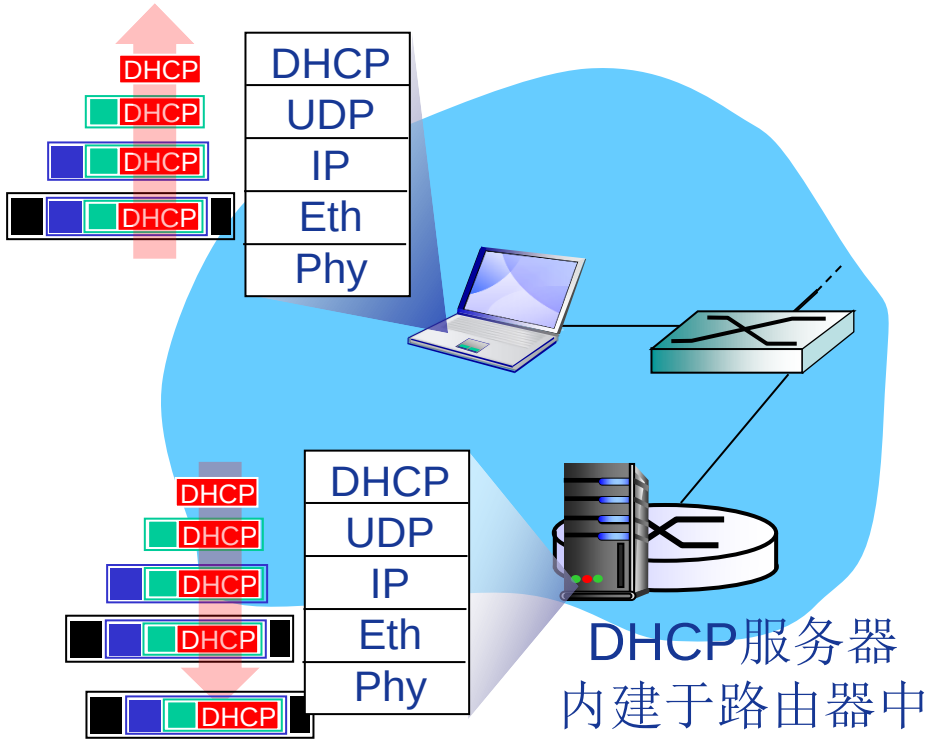
DHCP工作过程示例

- 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

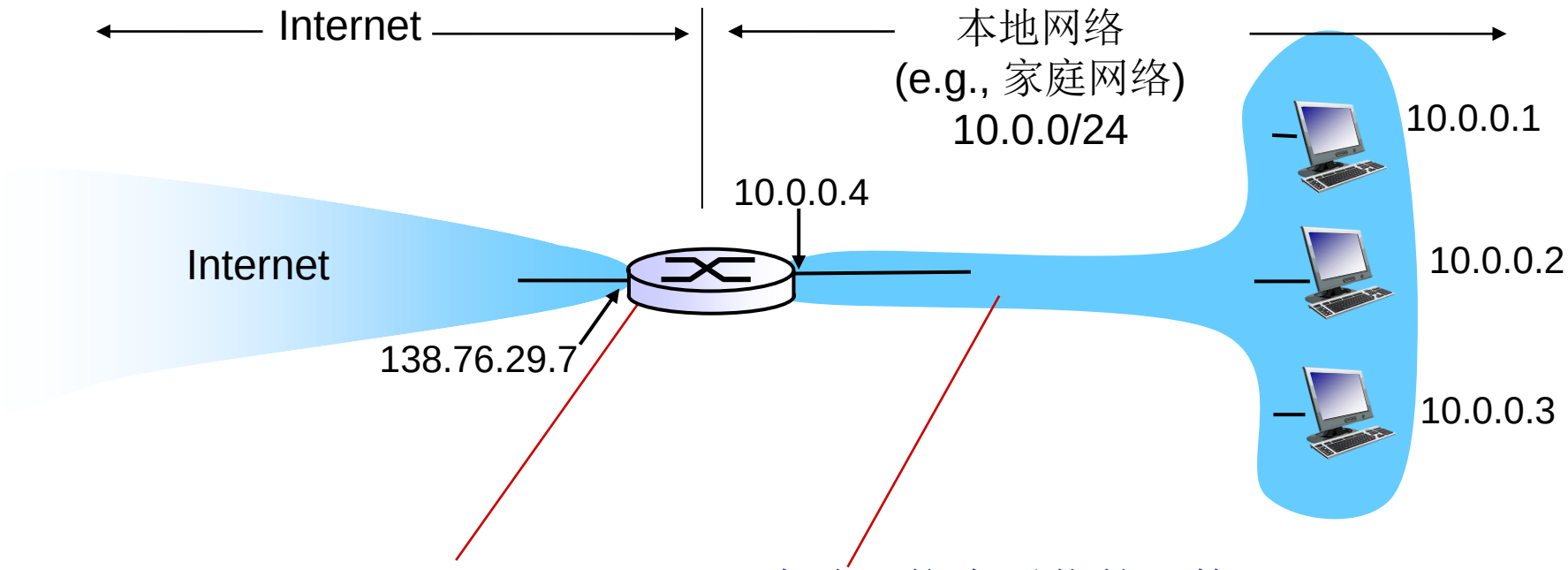
- DHCP服务器构造ACK报文
  - 包括分配给客户的IP地址、子网掩码、默认网关、DNS服务器地址



本讲主题

网络地址转换(NAT)

网络地址转换(NAT)



所有离开本地网络去往Internet的数据报的源IP地址需替换为相同的NAT IP地址：138.76.29.7以及不同的端口号

本地网络内通信的IP数据报的源与目的IP地址均在子网10.0.0/24内

网络地址转换(NAT)

动机:

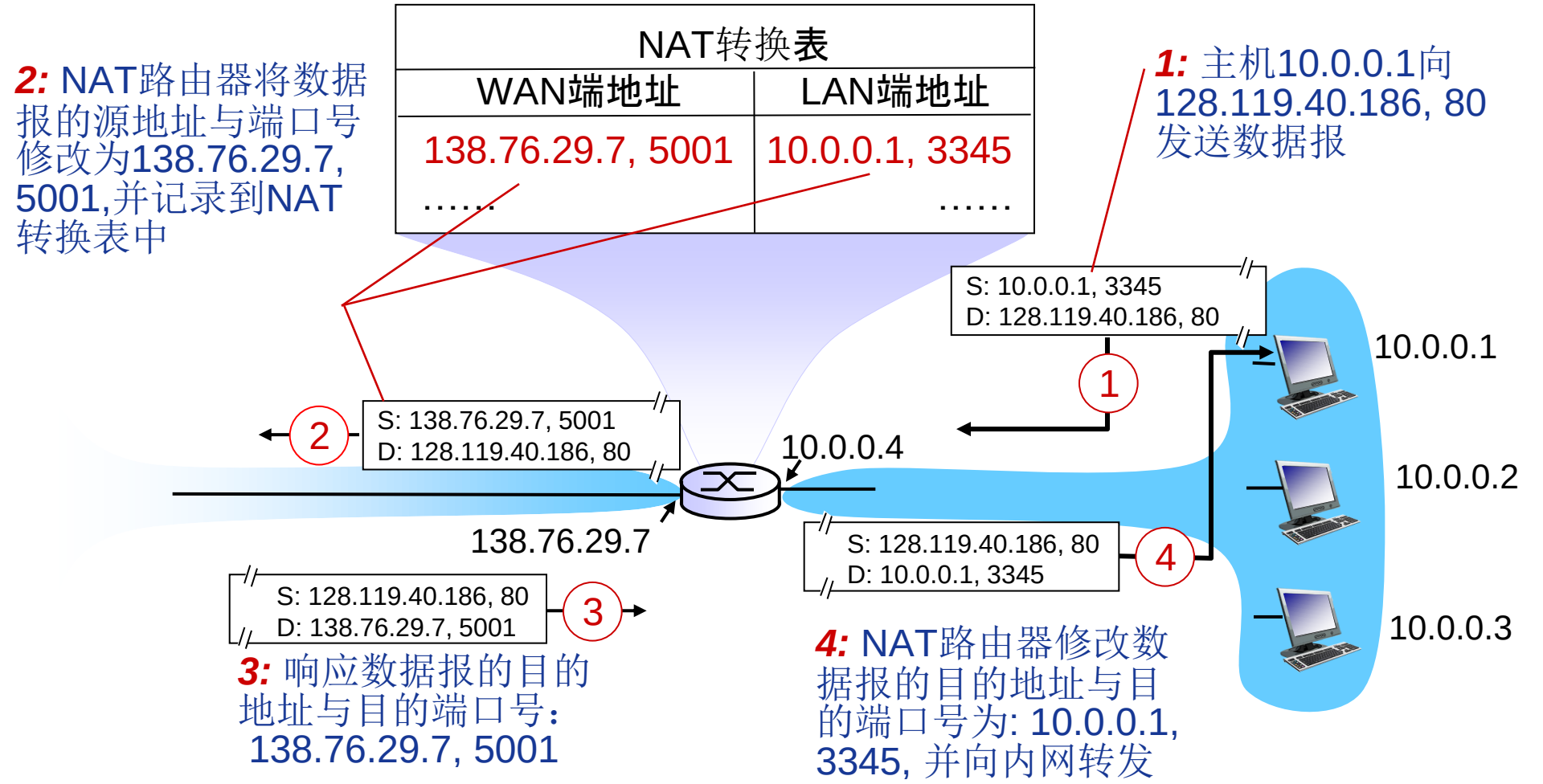
- 只需/能从ISP申请一个IP地址
  - IPv4地址耗尽
- 本地网络设备IP地址的变更，无需通告外界网络
- 变更ISP时，无需修改内部网络设备IP地址
- 内部网络设备对外界网络不可见，即不可直接寻址(安全)

网络地址转换(NAT)

实现:

- 替换
  - 利用(**NAT IP地址,新端口号**)替换每个外出IP数据报的(**源IP地址,源端口号**)
- 记录
  - 将每对(**NAT IP地址, 新端口号**) 与(**源IP地址, 源端口号**)的替换信息存储到NAT转换表中
- 替换
  - 根据NAT转换表，利用(**源IP地址, 源端口号**)替换每个进入内网IP数据报的(**目的IP地址,目的端口号**)，即(NAT IP地址, 新端口号)

网络地址转换(NAT)

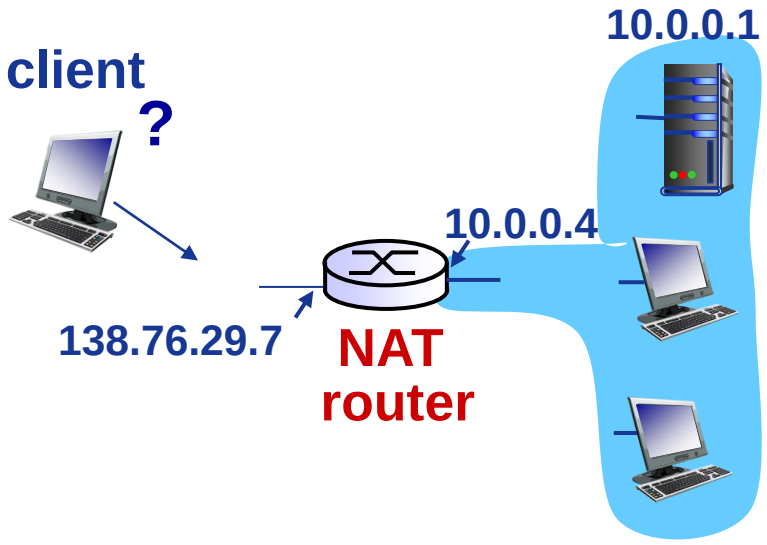

NAT转换表
WAN端地址
LAN端地址
138.76.29.7, 5001
10.0.0.1, 3345
……
……
2: NAT路由器将数据报的源地址与端口号修改为138.76.29.7, 5001,并记录到NAT转换表中
1: 主机10.0.0.1向128.119.40.186, 80发送数据报
S: 10.0.0.1, 3345
D: 128.119.40.186, 80
1
2
S: 138.76.29.7, 5001
D: 128.119.40.186, 80
10.0.0.4
138.76.29.7
S: 128.119.40.186, 80
D: 10.0.0.1, 3345
4
S: 128.119.40.186, 80
D: 138.76.29.7, 5001
3
3: 响应数据报的目的地址与目的端口号: 138.76.29.7, 5001
4: NAT路由器修改数据报的目的地址与目的端口号为: 10.0.0.1, 3345, 并向内网转发
10.0.0.1
10.0.0.2
10.0.0.3

网络地址转换(NAT)

- 16-bit端口号字段:
  - 可以同时支持60,000多并行连接!
- NAT主要争议:
  - 路由器应该只处理第3层功能
  - 违背端到端通信原则
    - 应用开发者必须考虑到NAT的存在，e.g., P2P应用
  - 地址短缺问题应该由IPv6来解决

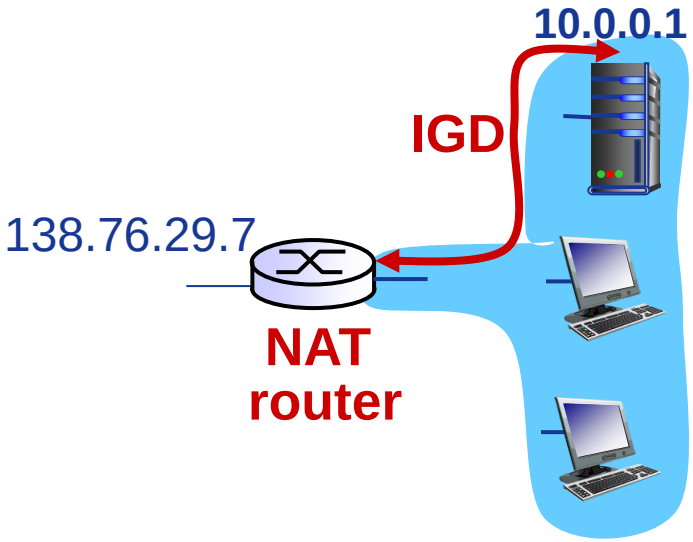
NAT穿透问题

- 客户期望连接内网地址为10.0.0.1的服务器
  - 客户不能直接利用地址10.0.0.1直接访问服务器
  - 对外唯一可见的地址是NAT地址: 138.76.29.7
- 解决方案1: 静态配置NAT，将特定端口的连接请求转发给服务器
  - e.g., (138.76.29.7, 2500) 总是转发给(10.0.0.1, 25000)



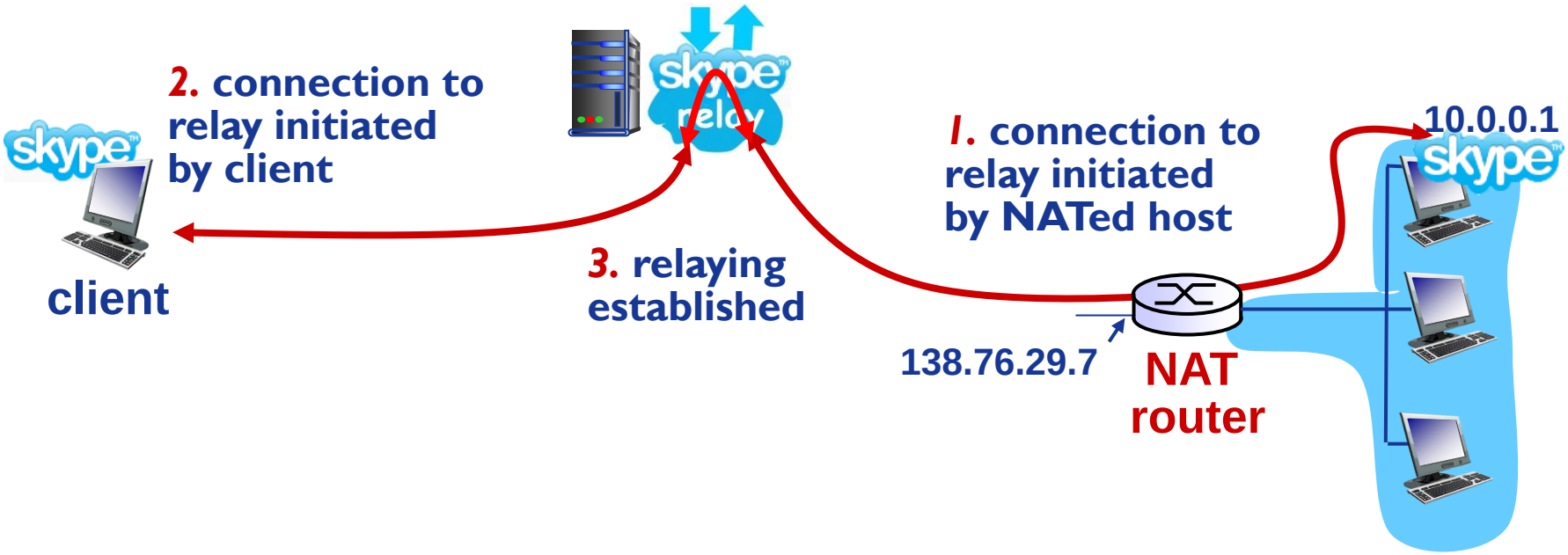
NAT穿透问题

- 解决方案2: 利用UPnP (Universal Plug and Play) 互联网网关设备协议 (IGD-Internet Gateway Device) 自动配置:
  - 学习到NAT公共IP地址 (138.76.29.7)
  - 在NAT转换表中，增删端口映射



NAT穿透问题

- 解决方案3: 中继(如Skype)
  - NAT内部的客户与中继服务器建立连接
  - 外部客户也与中继服务器建立连接
  - 中继服务器桥接两个连接的分组



本讲主题

互联网控制报文协议(ICMP)

互联网控制报文协议(ICMP)

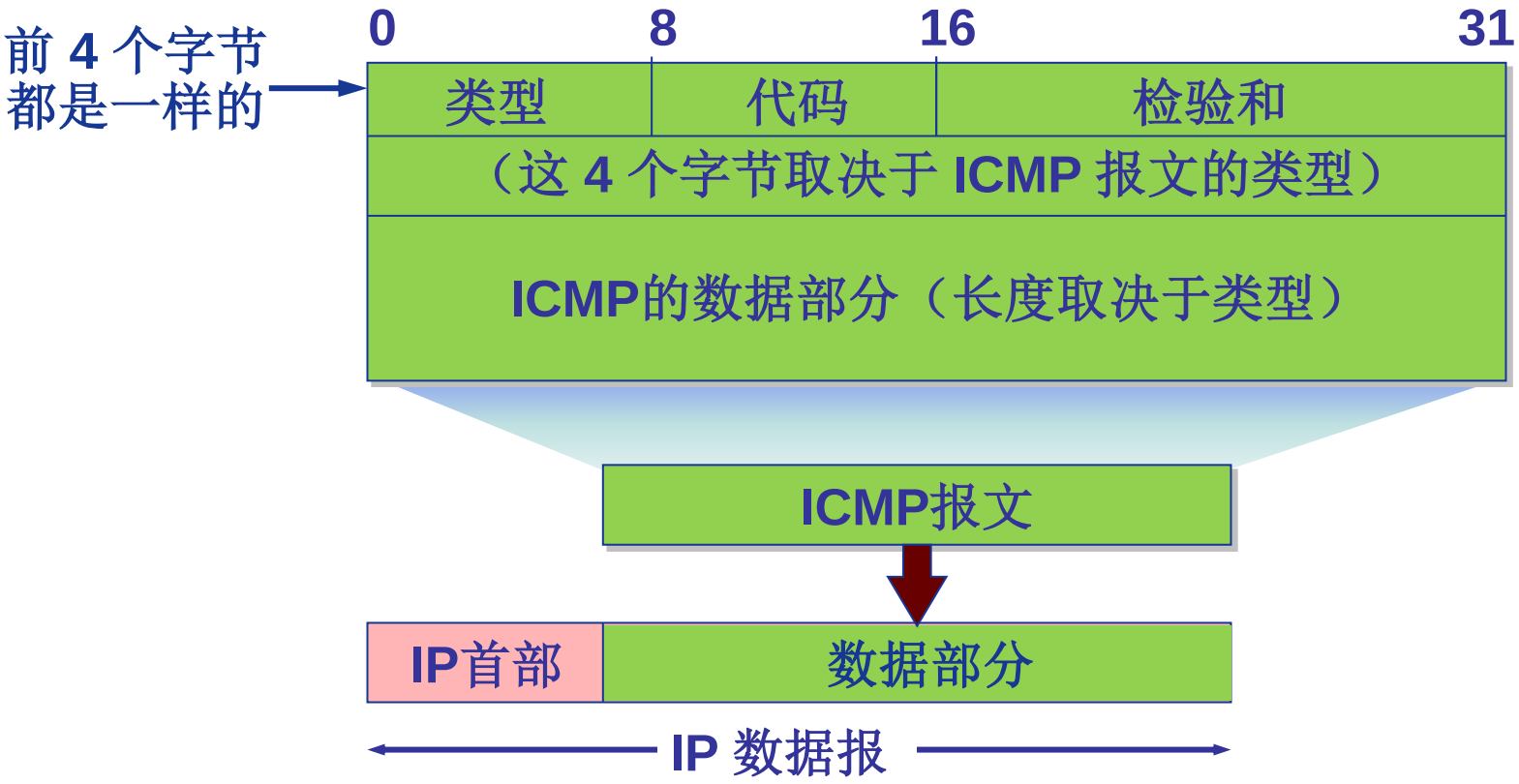
- 互联网控制报文协议 ICMP (Internet Control Message Protocol)支持主机或路由器：
  - 差错(或异常)报告
  - 网络探询
- 两类ICMP 报文:
  - 差错报告报文(5种)
    - 目的不可达
    - 源抑制(Source Quench)
    - 超时/超期
    - 参数问题
    - 重定向 (Redirect)
  - 网络探询报文(2组)
    - 回声(Echo)请求与应答报文(Reply)
    - 时间戳请求与应答报文

ICMP报文

类型(Type)	编码(Code)	description
0	0	回声应答 (ping)
3	0	目的网络不可达
3	1	目的主机不可达
3	2	目的协议不可达
3	3	目的端口不可达
3	6	目的网络未知
3	7	目的主机未知
4	0	源抑制(拥塞控制-未用)
8	0	回声请求(ping)
9	0	路由通告
10	0	路由发现
11	0	TTL超期
12	0	IP首部错误

ICMP报文的格式

- ICMP报文封装到IP数据报中传输

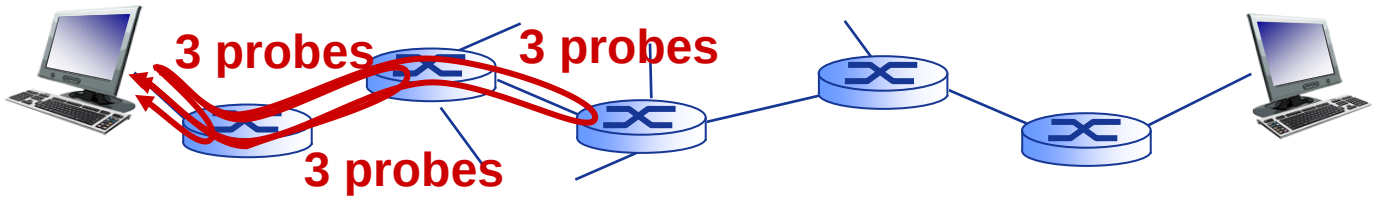


ICMP的应用举例：Traceroute

- 源主机向目的主机发送一系列UDP数据报
  - 第1组IP数据报TTL =1
  - 第2组IP数据报TTL=2, etc.
  - 目的端口号为不可能使用的端口号
- 当第*n*组数据报(TTL=*n*)到达第*n*个路由器时：
  - 路由器丢弃数据报
  - 向源主机发送ICMP报文(type=11, code=0)
  - ICMP报文携带路由器名称和IP地址信息
- 当ICMP报文返回到源主机时，记录RTT

停止准则：

- **UDP数据报最终到达目的主机**
- **目的主机返回“目的端口不可达”ICMP报文 (type=3, code=3)**
- **源主机停止**



本讲主题

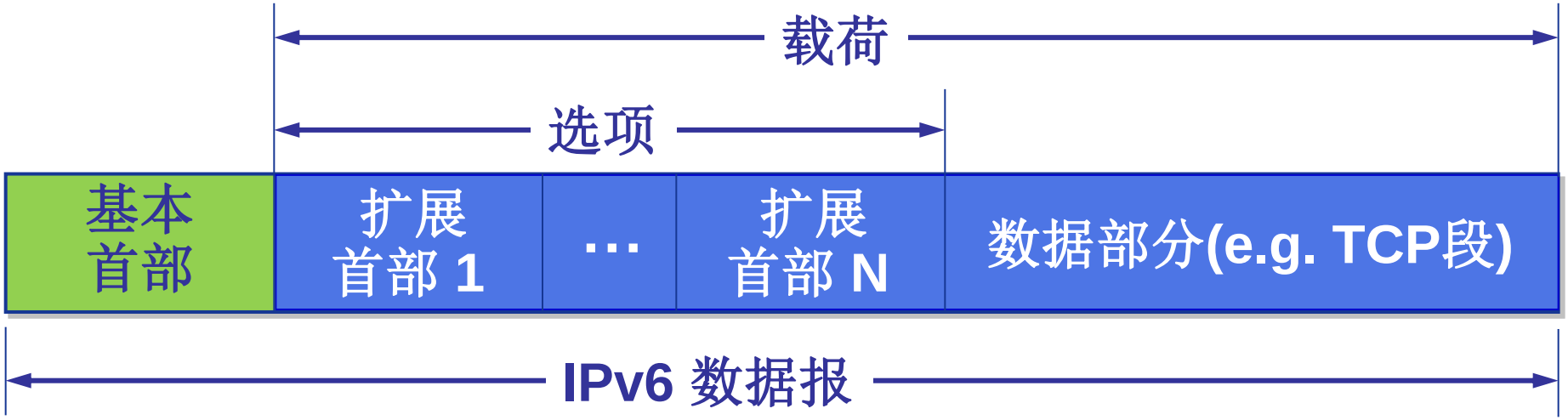
IPv6简介

IPv6：动机

- 最初动机: 32位IPv4地址空间已分配殆尽
- 其他动机: 改进首部格式
  - 快速处理/转发数据报
  - 支持QoS

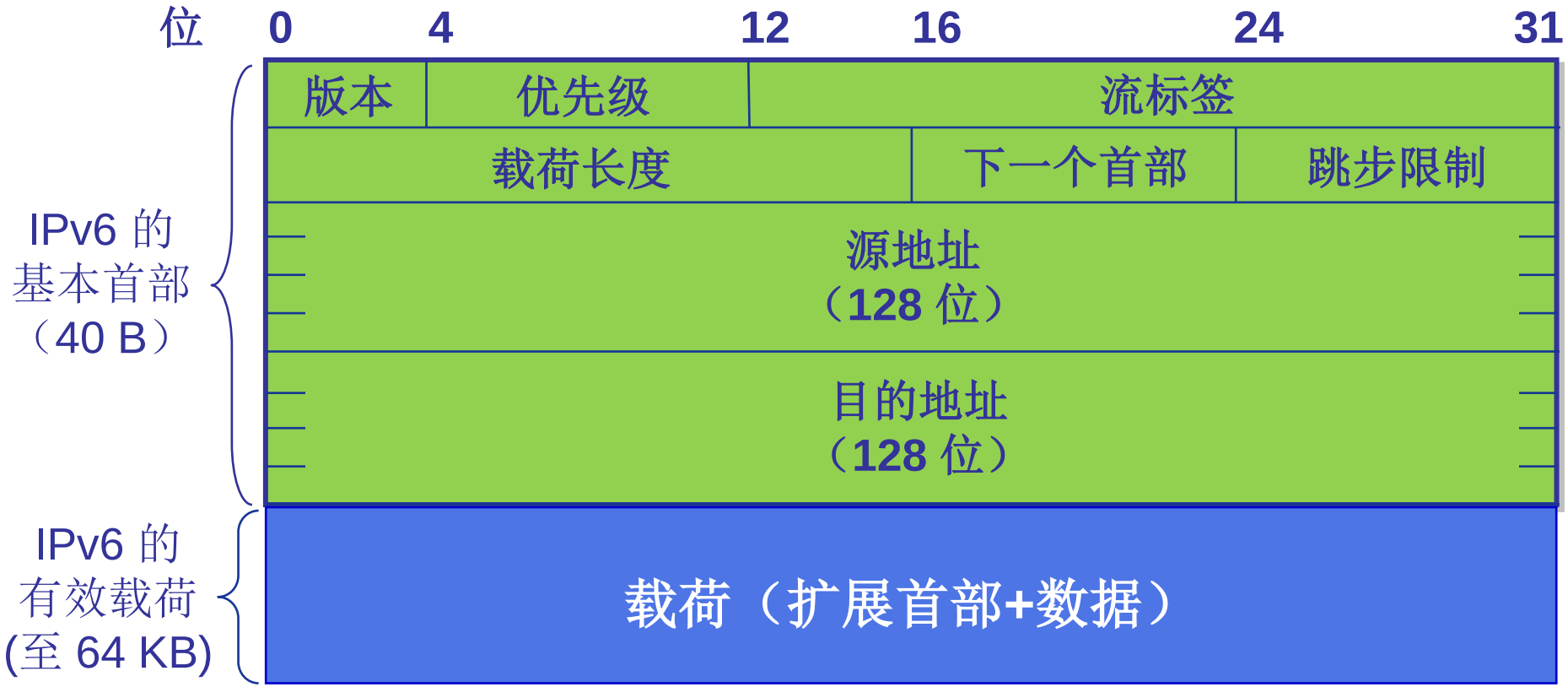
IPv6数据报格式:

- 固定长度的40字节基本首部
- 不允许分片



IPv6数据报格式

优先级(priority): 标识数据报的优先级
流标签(flow Label): 标识同一“流”中的数据报
下一个首部(next header): 标识下一个选项首部或上层协议首部(如TCP首部)



其他改变 vs IPv4

- 校验和(checksum): 彻底移除，以减少每跳处理时间
- 选项(options): 允许，但是从基本首部移出，定义多个选项首部，通过“下一个首部”字段指示
- ICMPv6: 新版ICMP
  - 附加报文类型，e.g. “Packet Too Big”
  - 多播组管理功能

IPv6基本地址类型

单播(unicast):

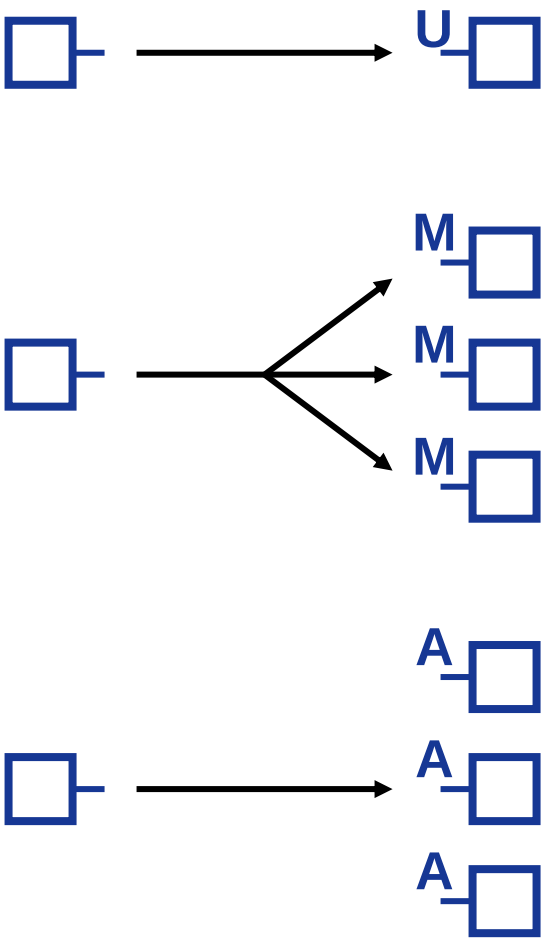
一对一通信

多播(multicast):

一对多通信

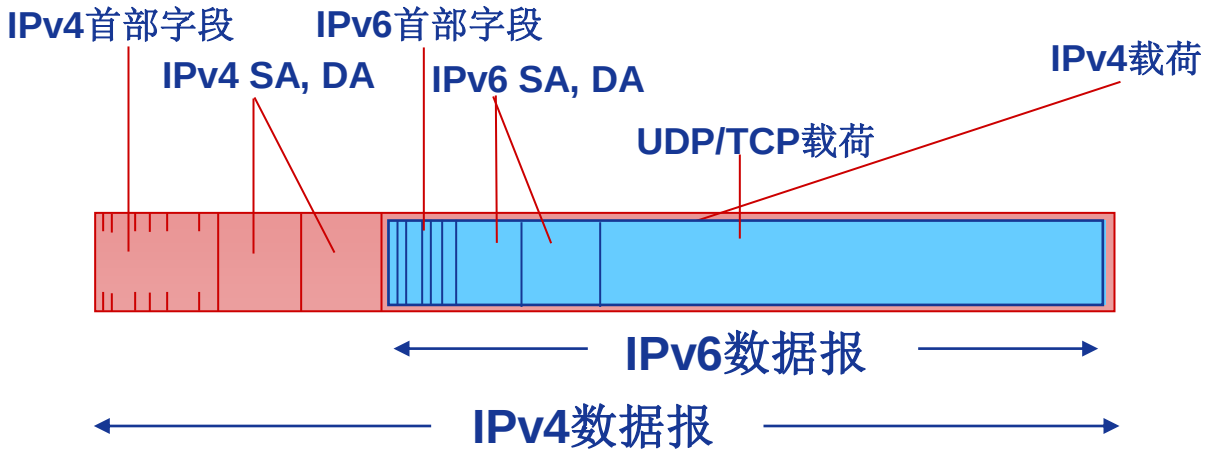
任意播(anycast):

一对一组之一（最近一个）通信



IPv4向IPv6过渡

- 不可能在某个时刻所有路由器同时被更新为IPv6
  - 不会有 “标志性的日期”
  - IPv4和IPv6路由器共存的网络如何运行？
- 隧道(tunneling): IPv6数据报作为IPv4数据报的载荷进行封装，穿越IPv4网络



隧道（tunneling）