

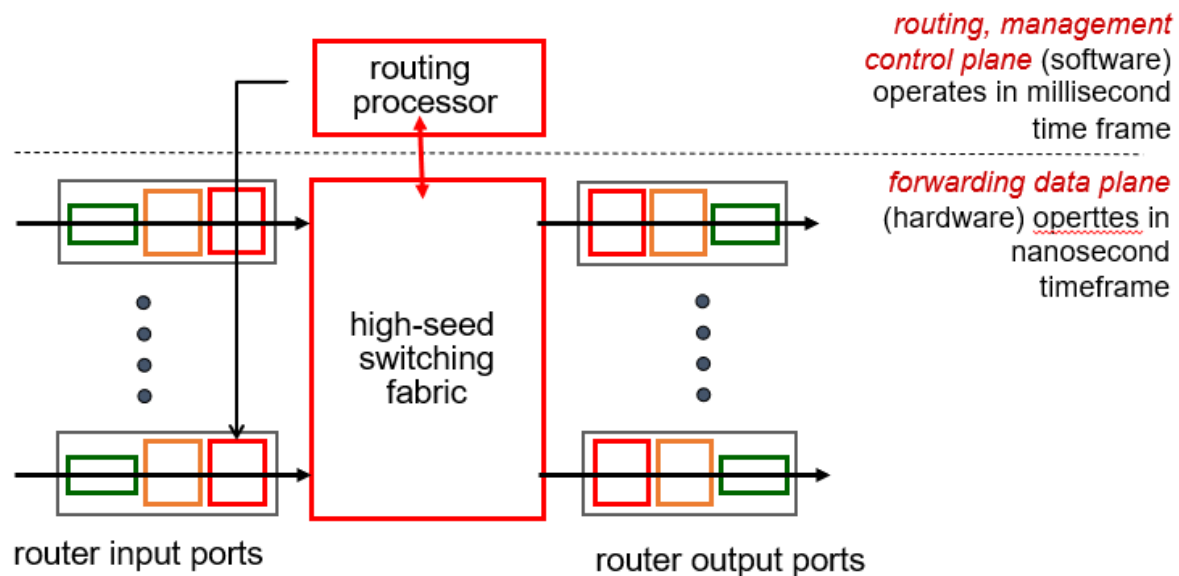
计算机网络



计算机与信息学院
人工智能学院

本章内容

- 网络层提供的服务类型
- 数据平面（转发：IP、ARP、ICMP、NAT）
- 控制平面（路由：RIP、OSPF、BGP）

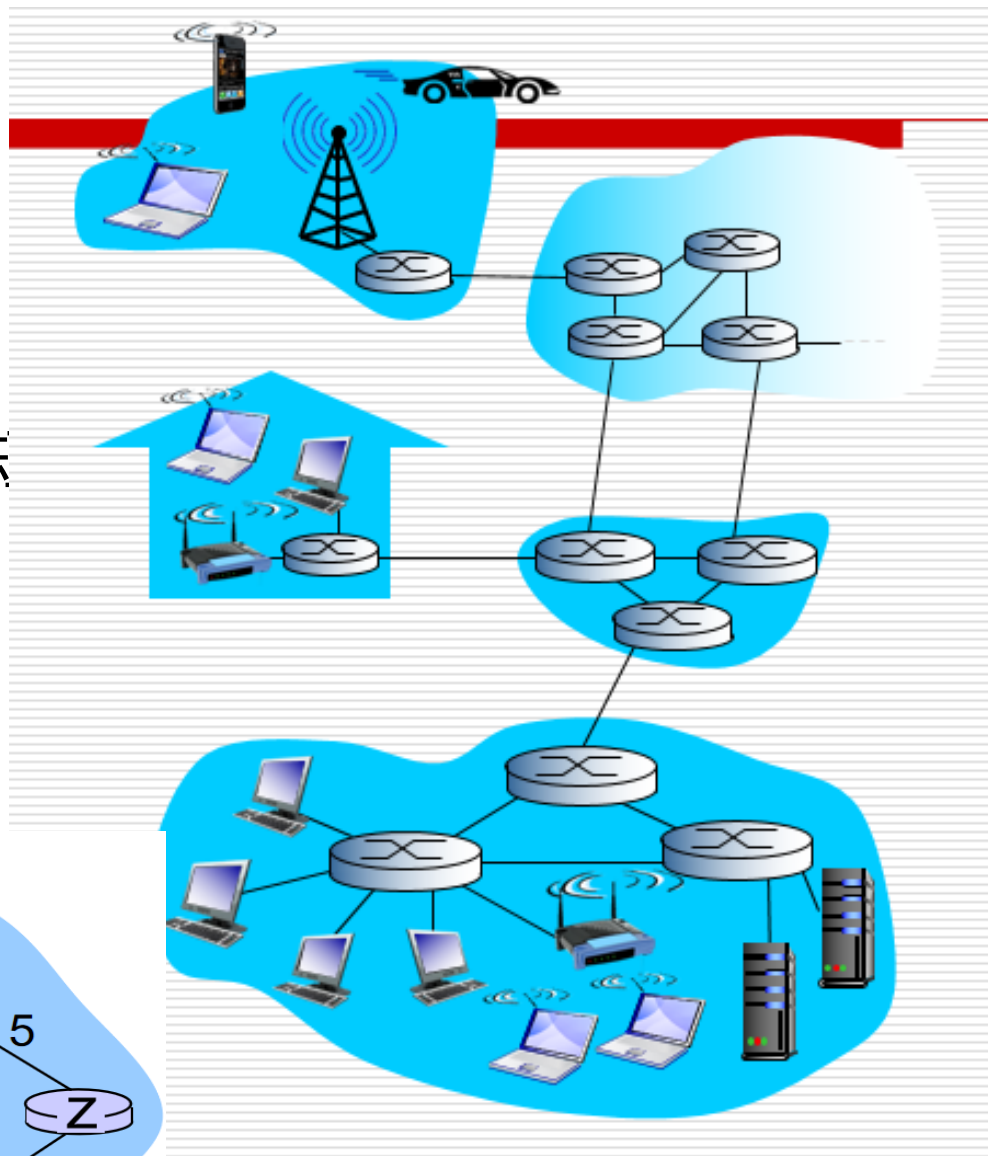
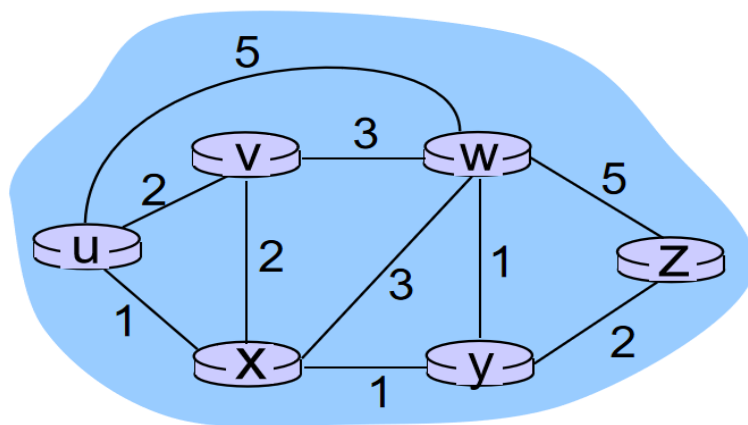


路由选择算法

从源节点到目的节点，选择一条“好”的路径

好？通信代价最小：时间最短、没有拥塞、路径最短

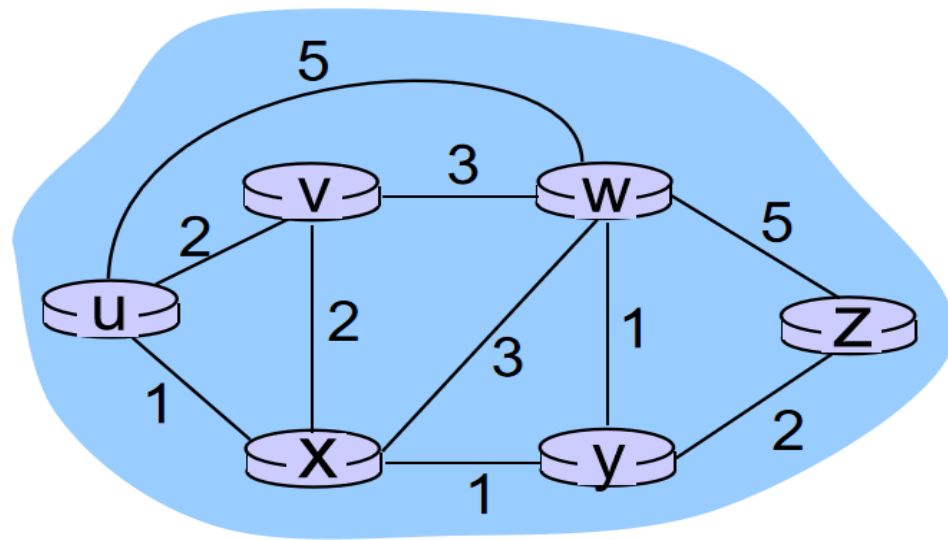
路径：经过的路由器队列



- 距离向量路由算法 (DV: Distance Vector)
- 链路状态路由算法 (LS: Link state)

距离向量路由算法 (DV: Distance Vector)

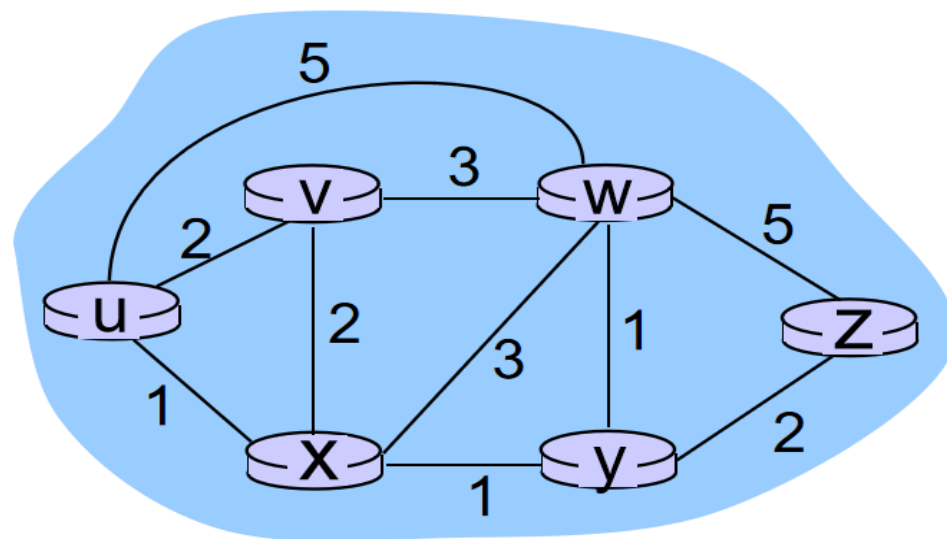
- ✓ 相邻路由器之间互相发送信息
- ✓ 每个路由器仅知道**相邻路由器**的信息



$$d_x(y) = \min_v \{ c(x,v) + d_v(y) \}$$

cost from neighbor v to destination y
cost to neighbor v

\min taken over all neighbors v of x

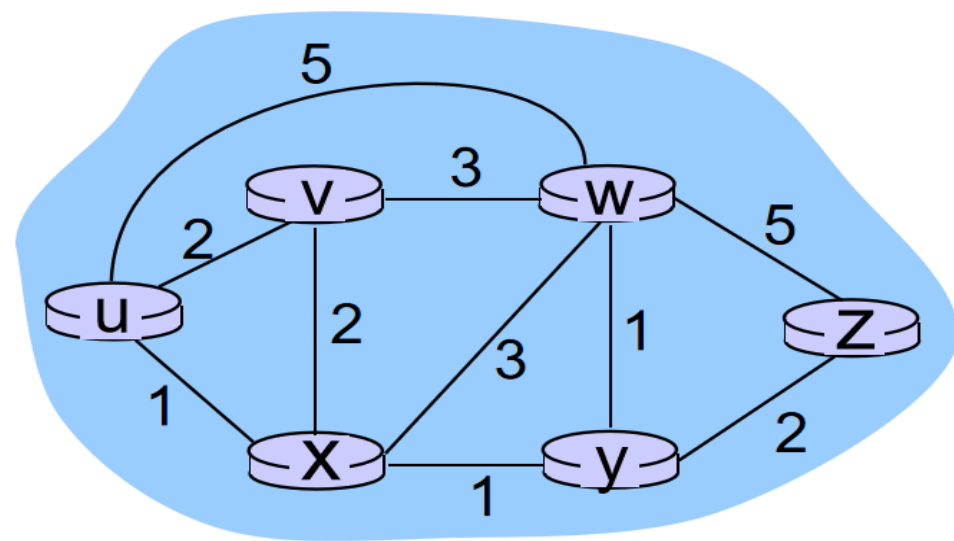


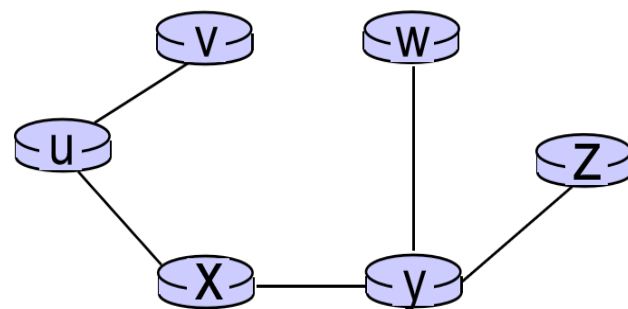
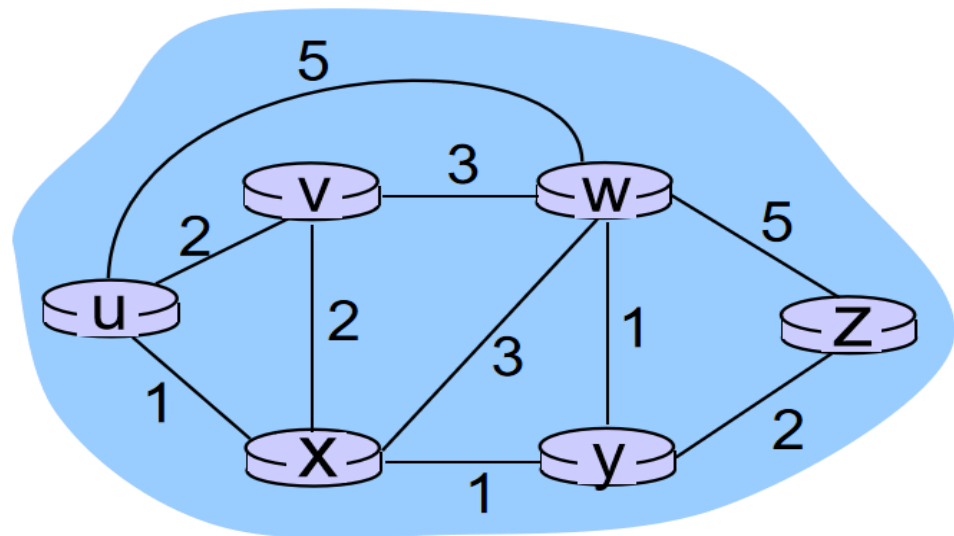
Bellman-Ford算法：求最短路径

链路状态路由算法 (LS: Link state)

- ✓ 发现相邻节点，测量到相邻节点的代价（延迟、开销）
- ✓ 组装分组，将分组通过**扩散**的方法发到所有其它路由器

——网络拓扑和链路代价，对所有路由器是**可知**的





Dijkstra算法：求最短路径

destination	link
v	(u,v)
x	(u,x)
y	(u,x)
w	(u,x)
z	(u,x)

- **距离向量路由算法 (DV: Distance Vector)**

- ✓ Bellman-Ford 求最短路径

- ✓ 相邻路由器交换信息

- **链路状态路由算法 (LS: Link state)**

- ✓ Dijkstra求最短路径

- ✓ 每个路由器广播发送链路状态

路由选择协议

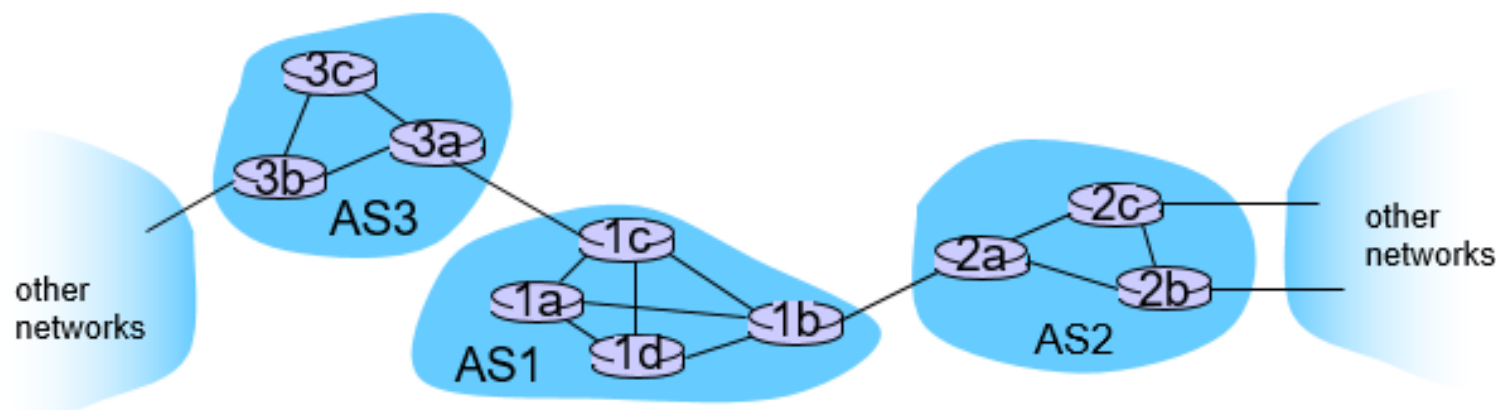
设计协议需要考虑的现实问题

- ✓ **规模**：以亿为数量级的目的网络，路由器的存储达不到这么大规模。路由表的交换，占用大量带宽
- ✓ **管理自治性**：互联网，网络的网络，每个网络管理员希望能自己控制网络内的路由

AS（Autonomous System）：一组路由器构成的区域，通常处在相同的管理控制下（相同的路由协议及代价度量）

每个AS有全局的ASN（AS号），由国际互联网管理机构统一分配

ASN	名称	IP 数量
AS4134	CHINANET-BACKBONE No.31	115485184
AS9808	CHINAMOBILE-CN China Mobile Communications Group Co.	65057536
AS4837	CHINA169-BACKBONE CHINA UNICOM China169 Backbone	61357056
AS4538	ERX-CERNET-BKB China Education and Research Network Center	42737408
AS9394	CTTNET China TieTong Telecommunications Corporation	38098944



- 内部网关协议（IGP）：在一个AS内进行路由计算，AS内所有路由器运行相同的路由协议：OSPF, RIP
- 外部网关协议（EGP）：AS之间的路由，BGP-4

1) RIP

路由信息协议 (Routing Information Protocol)

- 在1982年发布的BSD-Unix中实现，两个版本RIP1和RIP2

——路由选择算法：距离向量 (DV)

RIP

距离： 路由器与直接相连的网络距离为1， 每经过1个路由器， 跳数+1

相邻路由器 交换信息

——路由表（目的网络， 距离， 下一跳）

RIP

1. 相邻路由器交换信息

——路由表（目的网络，距离，下一跳）

2. 更新本地路由表 $d_x(y) = \min_v \{c(x,v) + d_v(y)\}$

3. 定期更新

——如每隔30秒交换信息，网络拓扑改变时交换信息

路由器 R₆

目的网络	距离	下一跳路由器
Net2	3	R ₄
Net3	4	R ₅
...

R₄ 发来的路由表

目的网络	距离	下一跳路由器
Net1	3	R ₁
Net2	4	R ₂
Net3	1	直接交付

R6更新路由表

Net1: 一条新路由, 增加

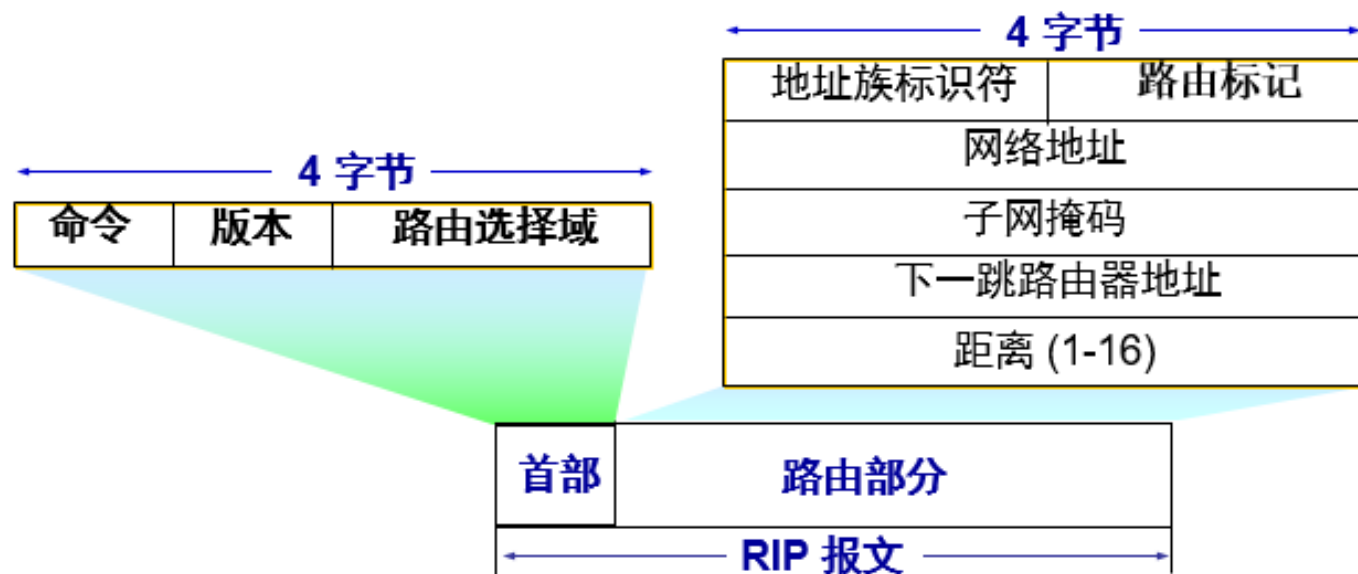
Net2: 相同的下一跳, 替换

Net3: 不同的下一跳, 新跳数小, 替换

: 不同的下一跳, 跳数相同, 不变

: 不同的下一跳, 新跳数大, 不变

RIP报文格式



20个字节表示一条路由选项，最多可支持25条路由

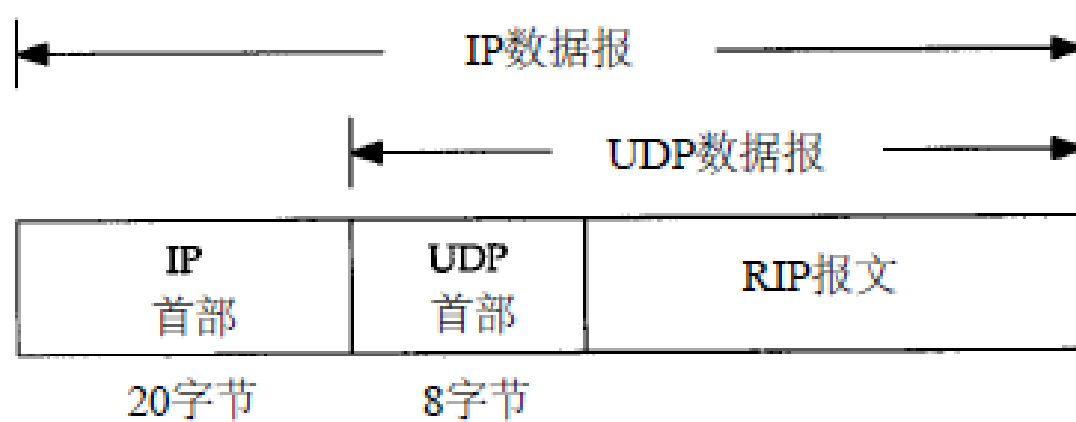
RIP: 规定跳数最大为16, 16代表该网络不可达

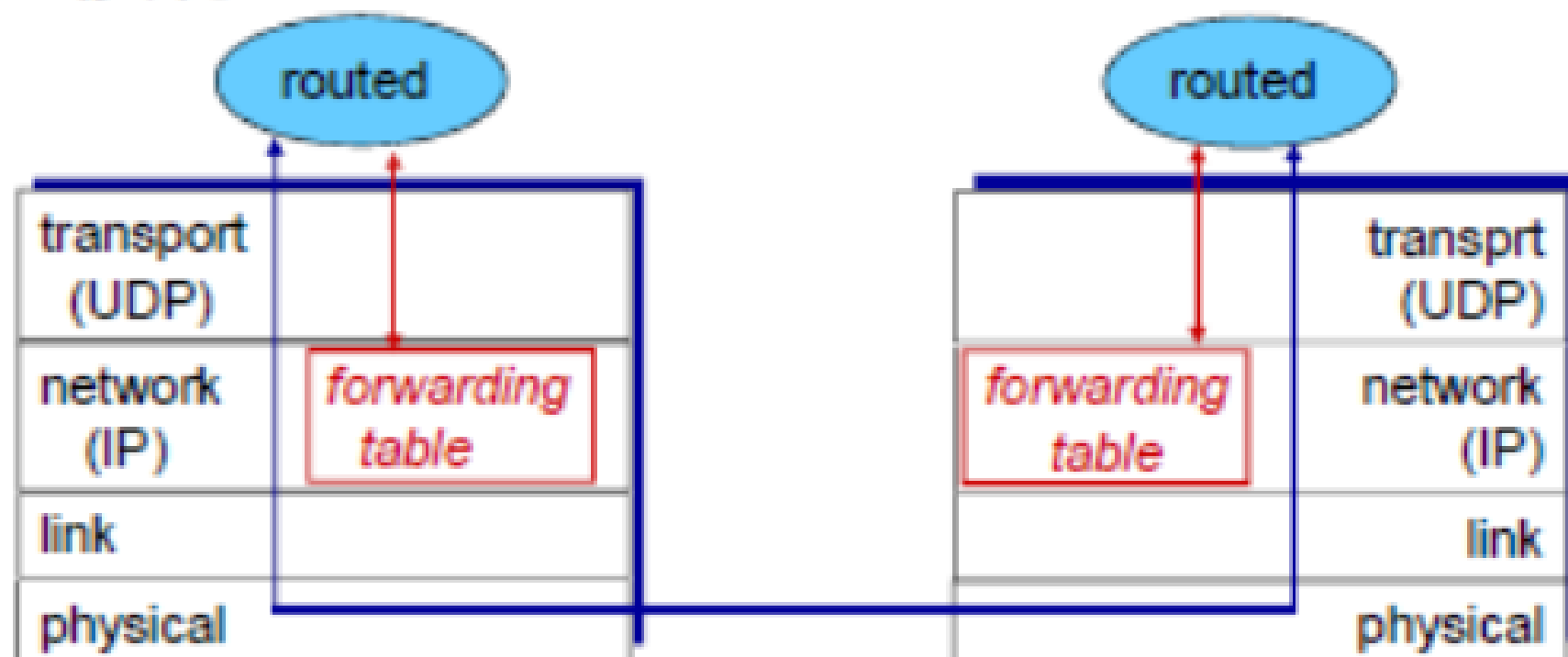
——因此一个路径至多包含15个路由器, (网络规模较小)

DV算法：无穷计数问题

链路故障/回路，链路开销增大

RIP以应用进程的方式实现，UNIX的守护进程





RIP的优缺点

- ✓实现简单、交换报文的复杂度较低
- ✓DV算法的无穷计数问题，限制了网络的规模
- ✓健壮性较差（一个节点路由表，可能会影响到很多节点）

——1989，开发了OSPF (Open Shortest Path First)

2) OSPF

开放最短路径优先 (Open Shortest Path First)

原理简单，实现复杂

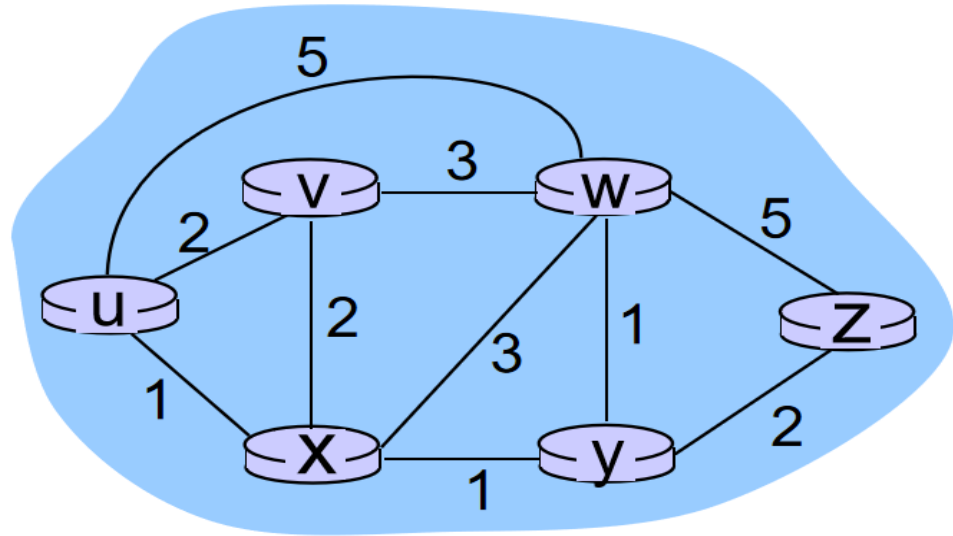
- 路由选择算法：链路状态 (LS)

- 路由选择算法：链路状态（LS）

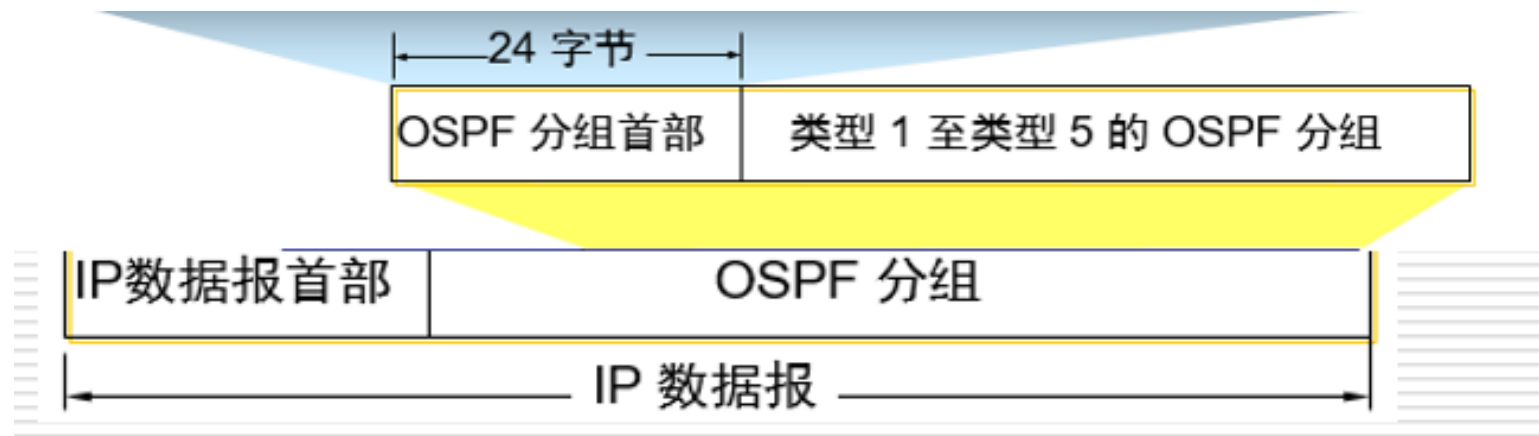
1. 发现相邻节点,获知对方网络地址
2. 测量到相邻节点的代价(延迟,开销)
3. 组装一个LS分组,描述它到相邻节点的代价情况
4. 将分组通过扩散的方法发到所有其它路由器

—————以上4步让每个路由器获得拓扑和边代价

5. 通过Dijkstra算法找出最短路径



OSPF的五种报文

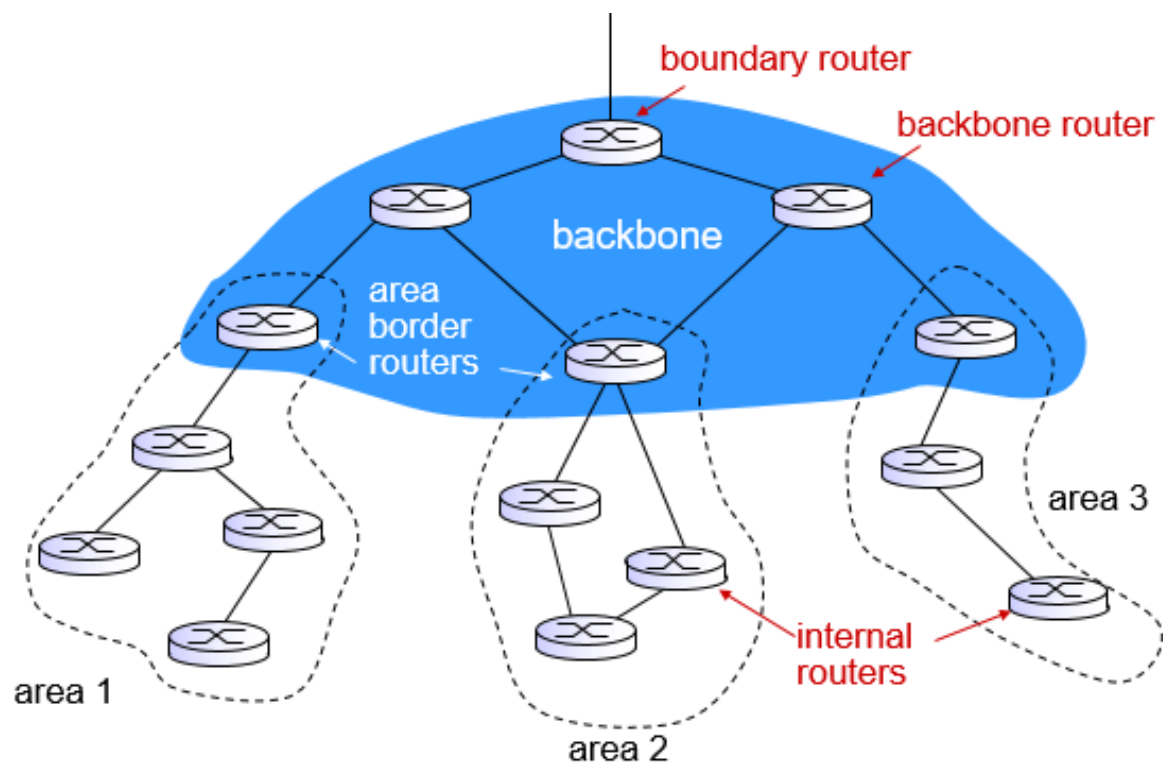


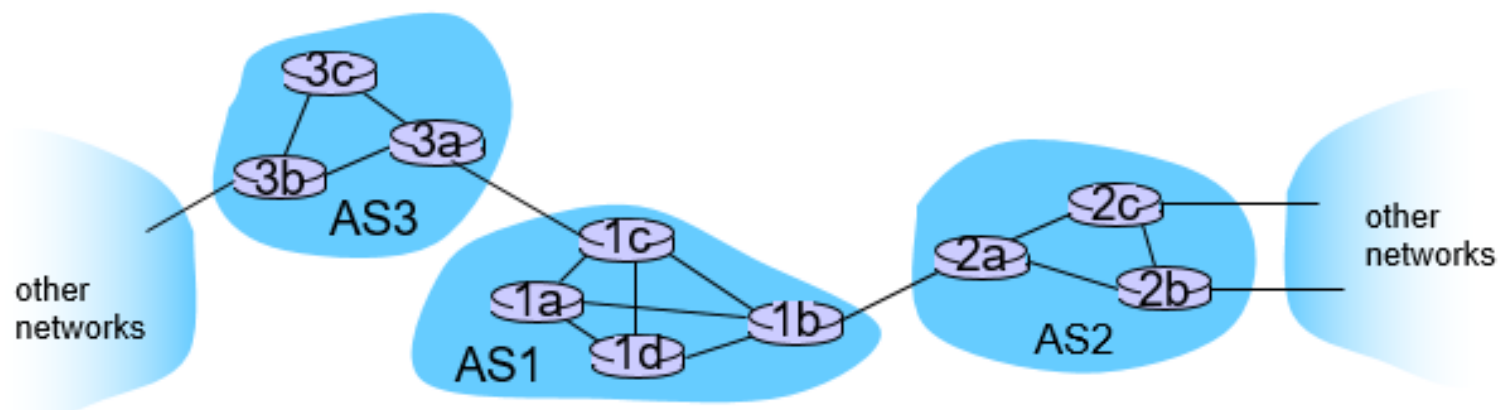
- 1) Hello: 问候分组, 发现和维持邻居路由器 (组播: 224.0.0.5)
- 2) Link State Update (LSU) : 通过 “扩散” 机制实现报文在网络中的广播



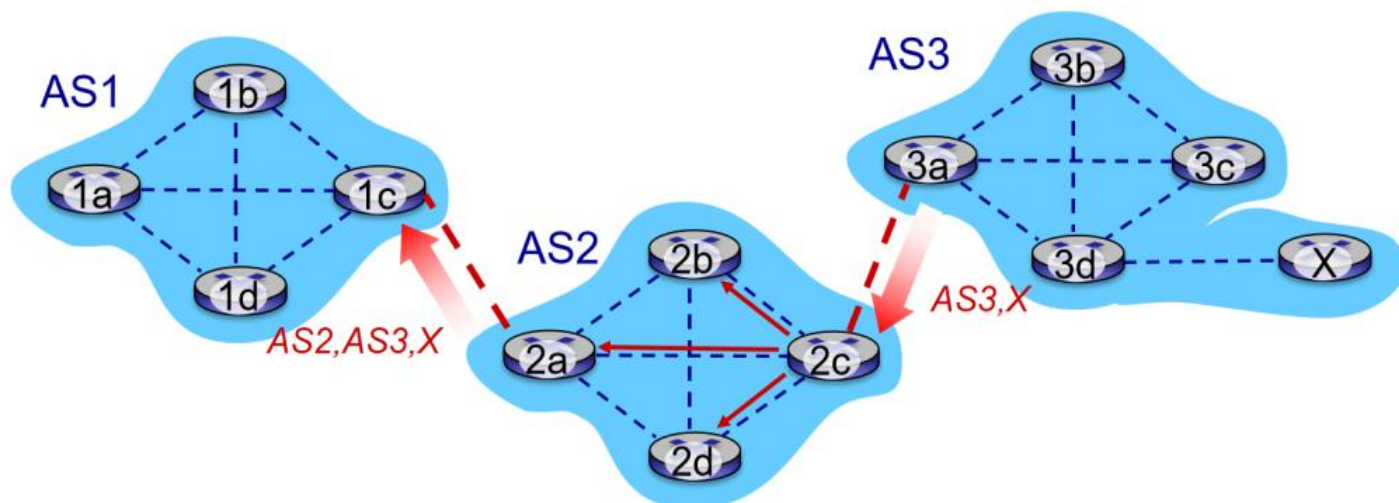
- 3) **Link State ACK**: 对LSU/hello分组的确认 (可靠传输)
- 4) **Database Description**: 数据库描述分组
- 5) **Link State Request**: 链路状态请求分组

✓ OSPF适用更大规模的网络，支持分层





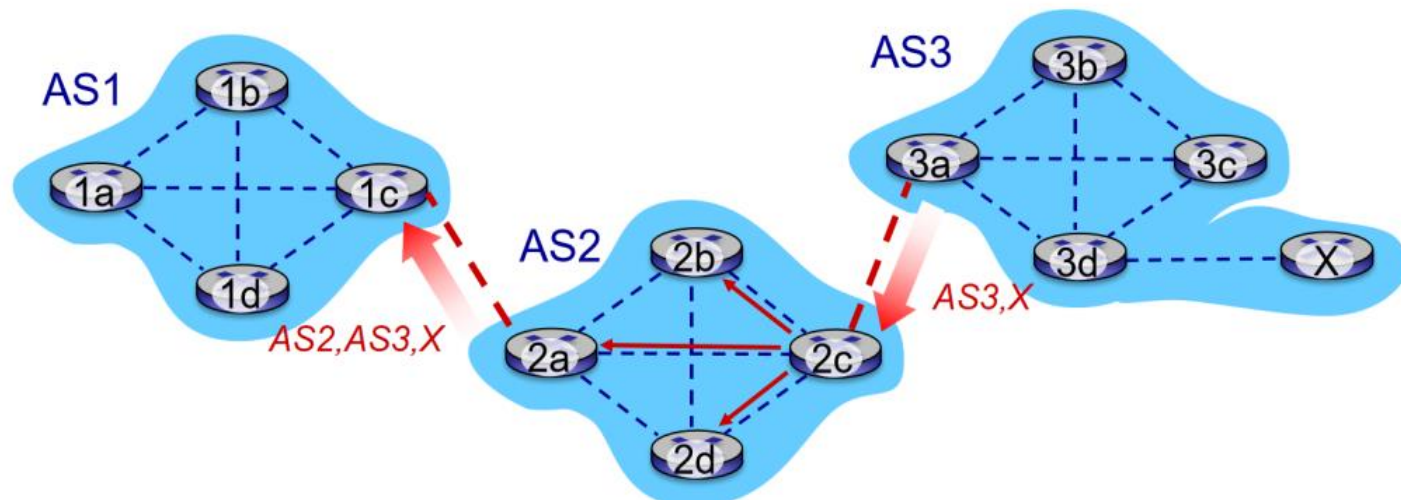
- 内部网关协议（IGP）： OSPF, RIP
- 外部网关协议（EGP）： AS之间的路由, **BGP-4**



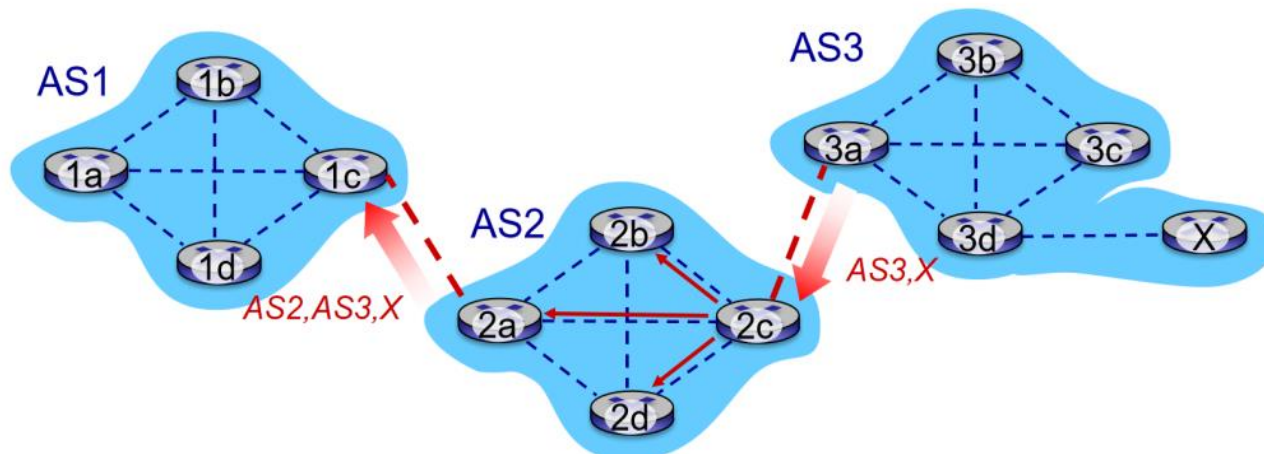
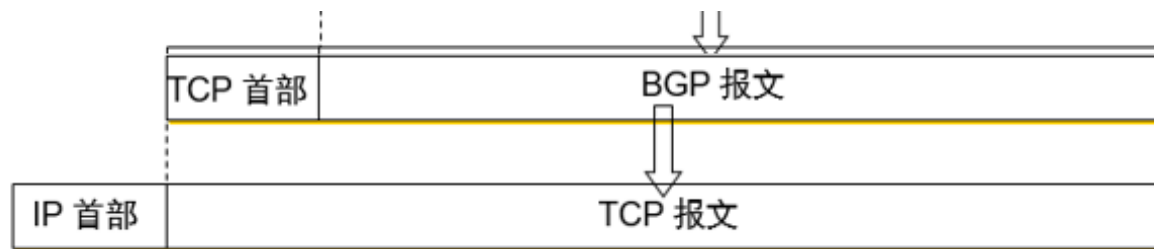
边界网关路由器： 连接到其他AS中1台或多台路由器

内部路由器： 仅连接AS内的主机和路由器

BGP (边界网关协议: Border Gateway Protocol)



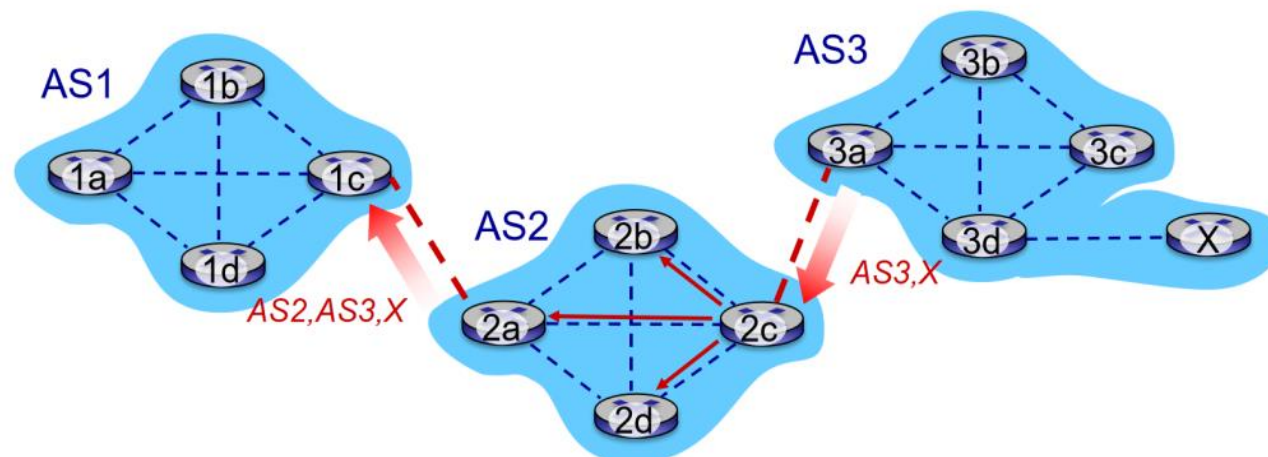
- ✓ 通过交换BGP报文，获得到达X（网络前缀）的路径
- ✓ 路径：AS序列（X: AS2, AS3）



TCP端口号: 179

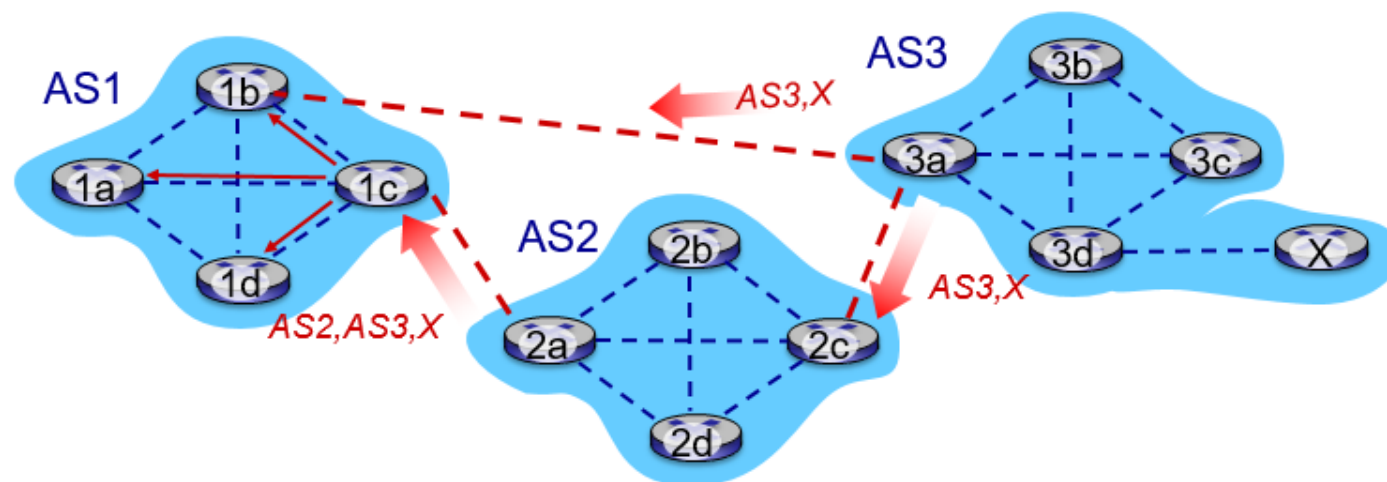
eBGP: 两个AS之间的边界路由器建立BGP连接，交换BGP报文

iBGP: AS内部两天路由器建立BGP连接，交换BGP报文



- 1) **OPEN:** 与相邻路由器建立连接
- 2) **UPDATE:** 发送新路径/取消旧路径 (到达X的AS Path、Next hop)
- 3) **KEEPALIVE:** 对OPEN的确认/在没有UPDATE消息时, 保持连接
- 4) **NOTIFICATION:** 报错及关闭连接

存在多条路径时，确定到达X的‘最好’路径



- 1) 策略制定
- 2) 最短AS—Path
- 3) 选择最近（内部路由代价最小）的下一跳：热土豆路由
- 4) 其他

总结

- 网络层提供的服务类型：虚电路和数据报
- 数据平面（转发：IP、ARP、ICMP、NAT）
- 控制平面（路由：RIP、OSPF、BGP）

IP：封装上层数据，转发，实现主机之间的通信

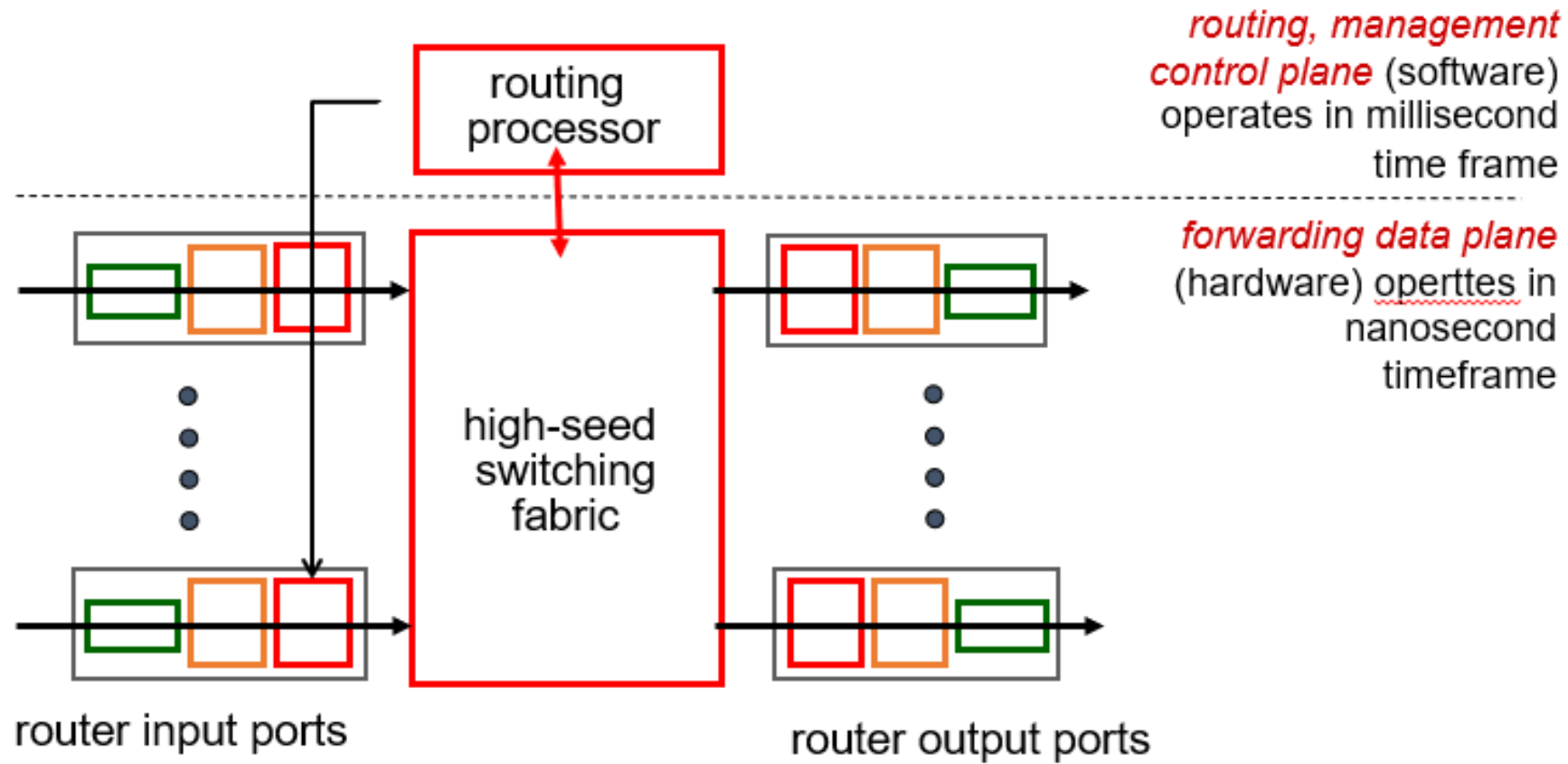
ARP：封装成帧，实现IP地址和MAC地址的解析

ICMP：封装成IP数据报（协议类型：1），差错报告和网络层请求应答

RIP：由传输层UDP封装（端口号：520），应用进程，交换路由表信息

OSPF：封装成IP数据报（协议类型：89），AS内广播链路状态信息

BGP：由传输层TCP封装（端口号：179），AS之间，以及AS内部交换路径



作业

P249: 5.1

P250: 5.8, 5.11, 5.12, 5.13

补充:

对比分析路由协议RIP、OSPF、BGP-4, (运行区域、路由算法、报文传输)

为什么要采用域内和域间两种路由?