

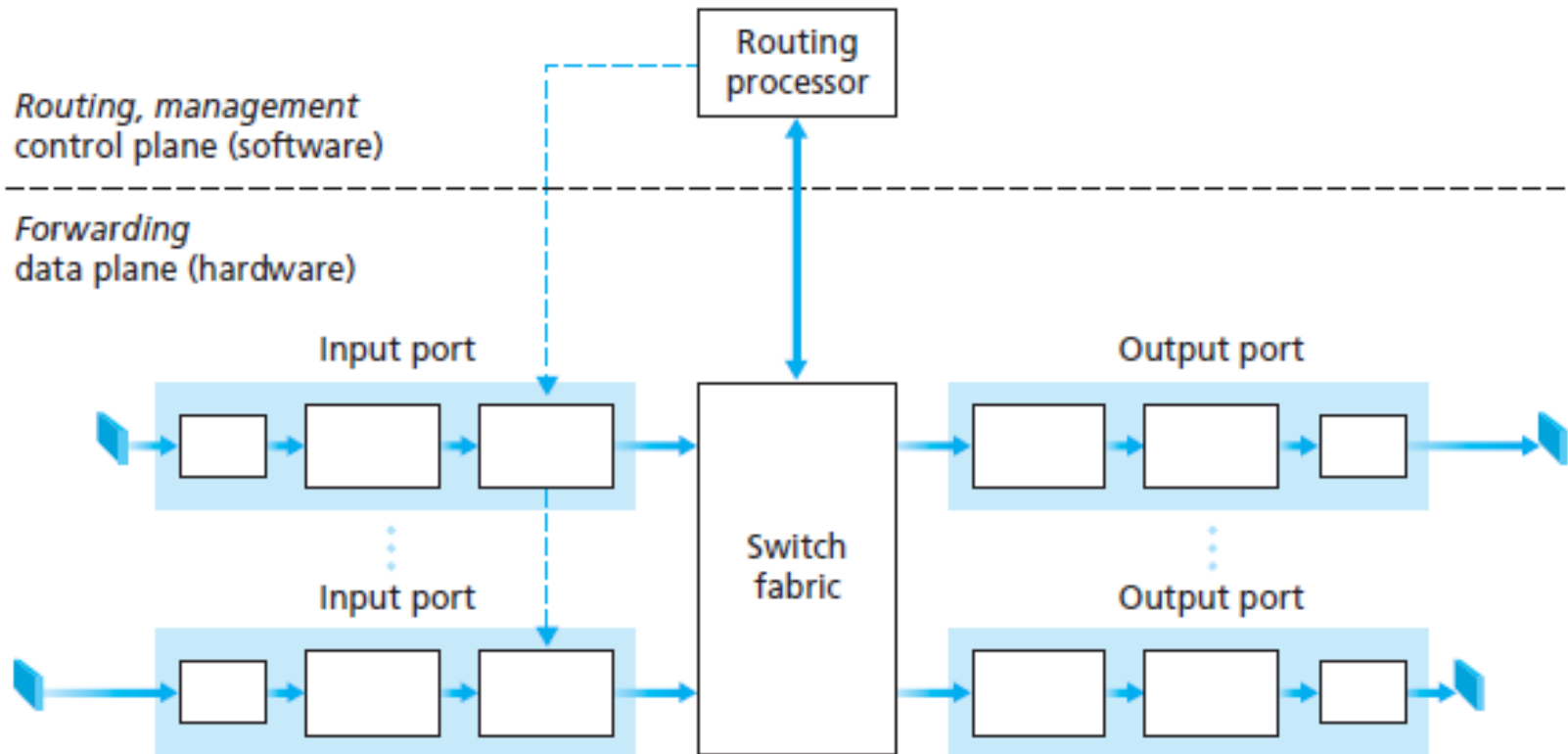
# 第六章 网络层

---

合肥工业大学  
计算机与信息学院

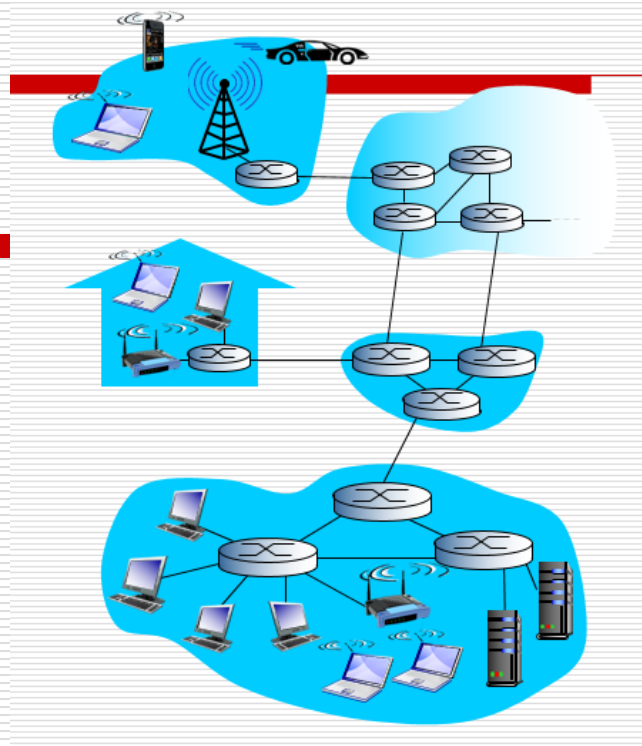
**数据平面：**从输入链路到输出链路的交换，根据转发表进行转发，硬件实现，纳秒级

**控制平面：**执行路由选择协议，生成路由表，软件实现，毫秒级



# 路由选择算法

---



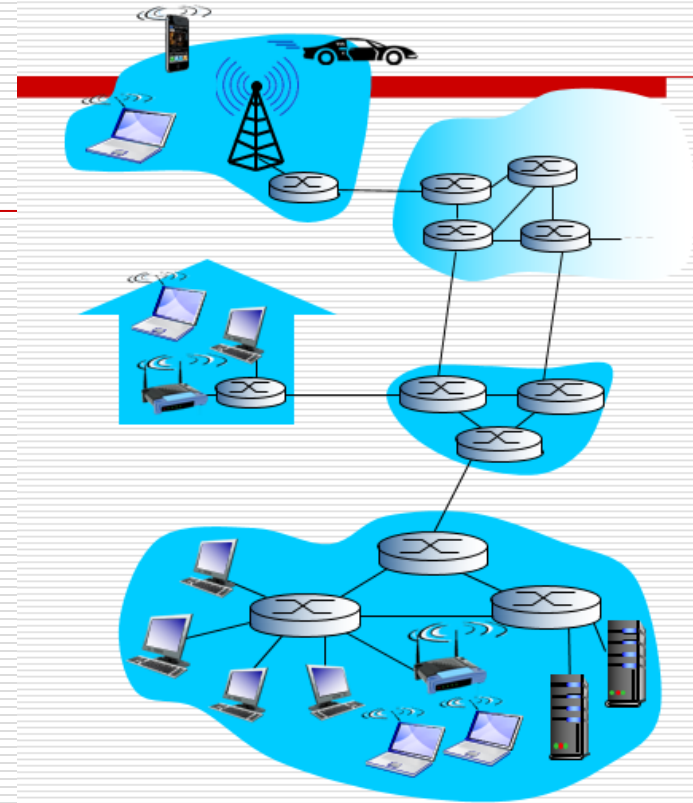
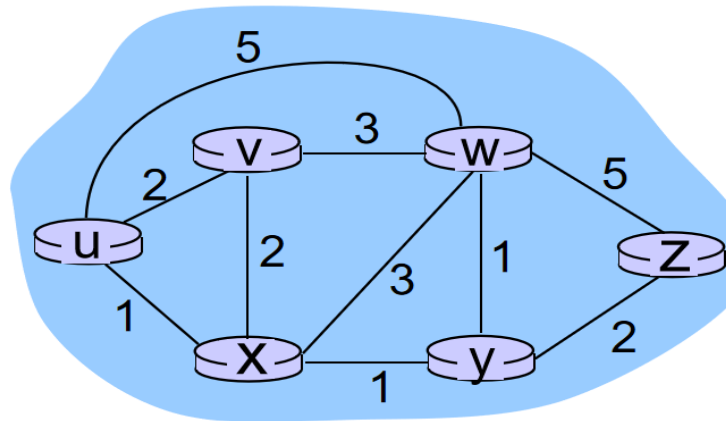
**目标：**从源节点到目的节点，选择一条“好”的路径

**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

---

图



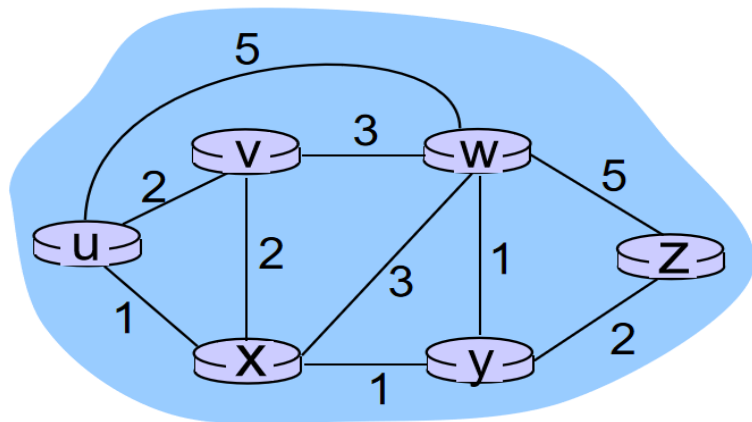
graph:  $G = (N, E)$

$N$  = set of routers =  $\{ u, v, w, x, y, z \}$

$E$  = set of links =  $\{ (u,v), (u,x), (v,x), (v,w), (x,w), (x,y), (w,y), (w,z), (y,z) \}$

cost of path  $(x_1, x_2, x_3, \dots, x_p) = c(x_1, x_2) + c(x_2, x_3) + \dots + c(x_{p-1}, x_p)$

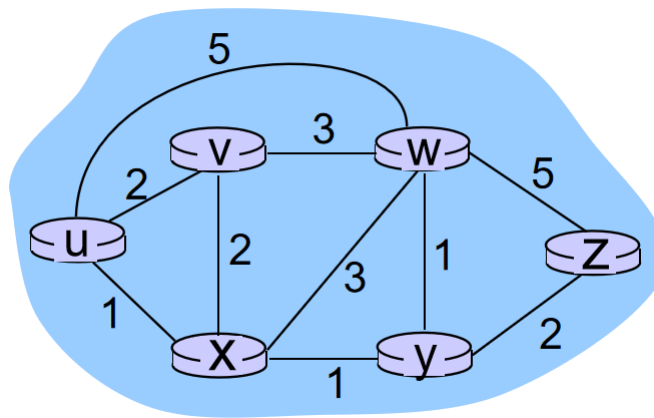
# 路由算法



- 全局：所有路由器有完整的网络拓扑及链路代价信息——**Link state**算法
- 局部：每个路由器只知道物理连接的邻居路由器，以及到邻居路由器的路径代价——**distance vector**算法

# 1) Link State (LS)

□ 网络拓扑和链路代价，对所有路由器是**可知**的

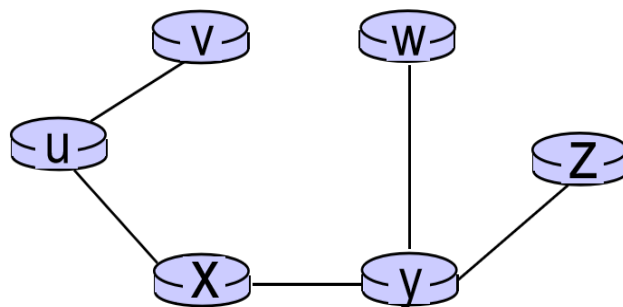
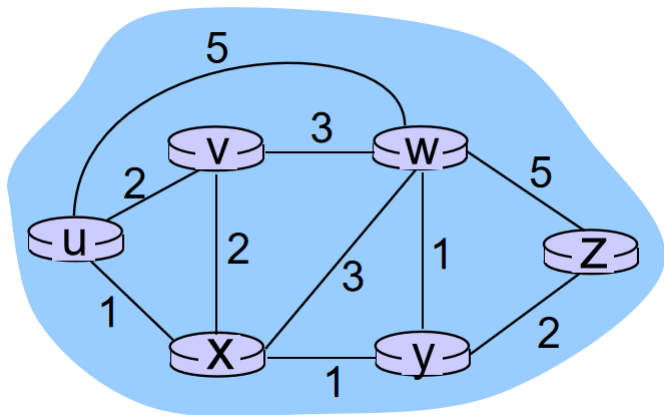


获得网络拓扑和  
链路代价信息

使用**最短路由算法**  
得到路由表

使用此路由表

# Dijkstra计算最短路径



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

# 获得拓扑和链路状态

---

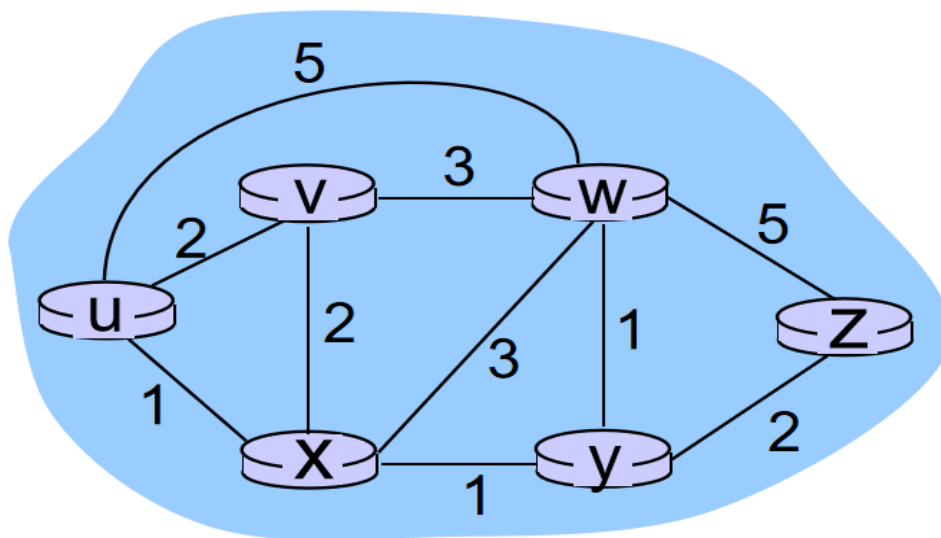
1. 发现相邻节点，获知其网络地址
  2. 测量到相邻节点的代价（延迟、开销）
  3. 组装一个分组，描述相邻节点的情况
  4. 将分组通过扩散的方法发到所有其它路由器
    - ✓ 如何区分新旧分组：顺序号
    - ✓ 序号回绕、路由器故障？
-



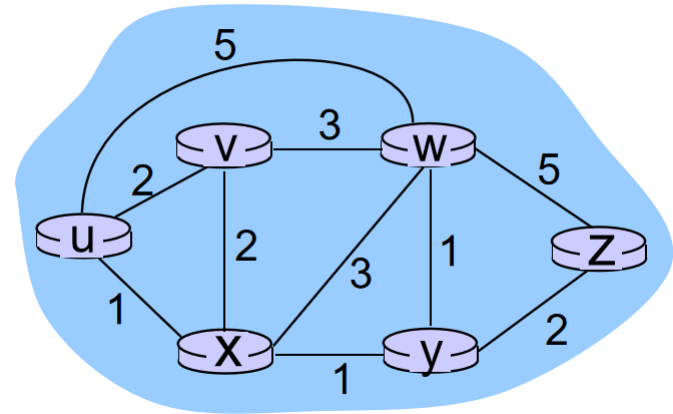
## 2) Distance Vector (DV)

---

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



# Bellman-Ford



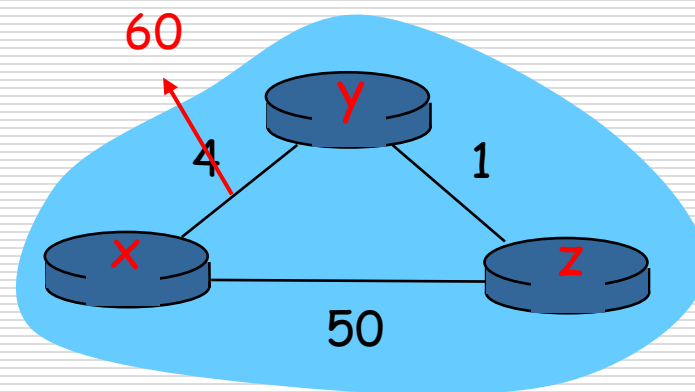
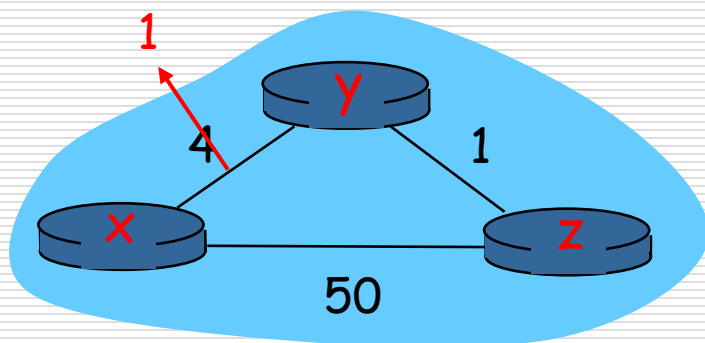
$$d_x(y) = \min_v \{ c(x,v) + d_v(y) \}$$

cost from neighbor v to destination y  
cost to neighbor v

min taken over all neighbors v of x

---

□ DV: 好消息传的快, 坏消息传的慢



无穷计数问题

# LS 和 DV 算法的比较

## 消息复杂度 (DV胜出)

- LS: 有  $n$  节点,  $E$  条链路, 发送报文  $O(nE)$  个
  - 局部的路由信息: 全局传播
- DV: 只和邻居交换信息
  - 全局的路由信息, 局部传播

## 收敛时间 (LS胜出)

- LS:  $O(n^2)$  算法
  - 有可能震荡
- DV: 收敛较慢
  - 可能存在路由环路
  - count-to-infinity 问题

## 健壮性: 路由器故障会发生什么 (LS胜出)

### LS:

- 节点会通告不正确的链路代价
- 每个节点只计算自己的路由表
- 错误信息影响较小, 局部, 路由较健壮

### DV:

- DV 节点可能通告对全网所有节点的不正确路径代价
  - 距离矢量
- 每一个节点的路由表可能被其它节点使用
  - 错误可以扩散到全网

2种路由选择算法都有其优缺点, 而且在互联网上都有应用

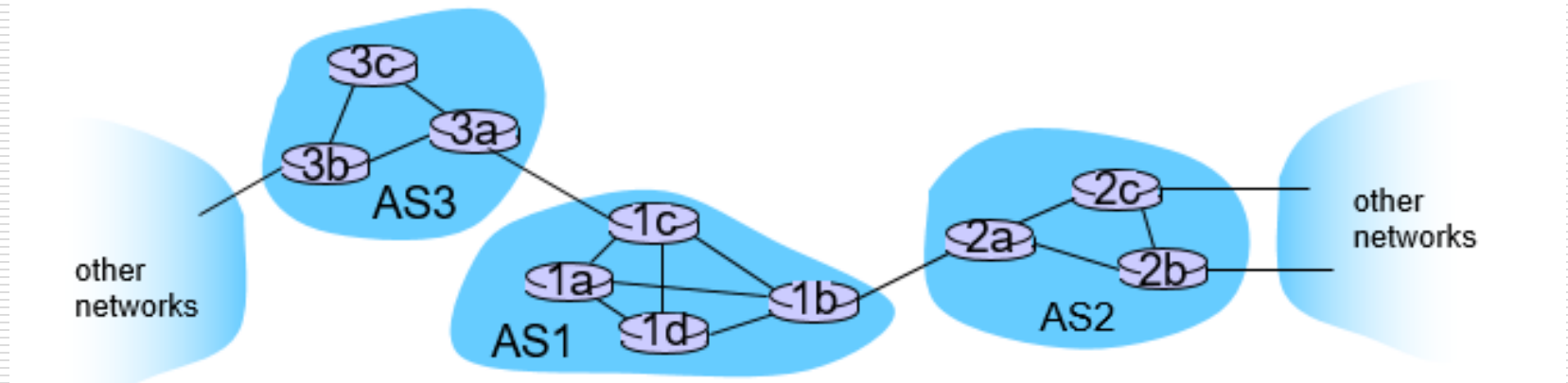
# 路由选择协议

---

## □ 设计协议需要考虑的现实问题

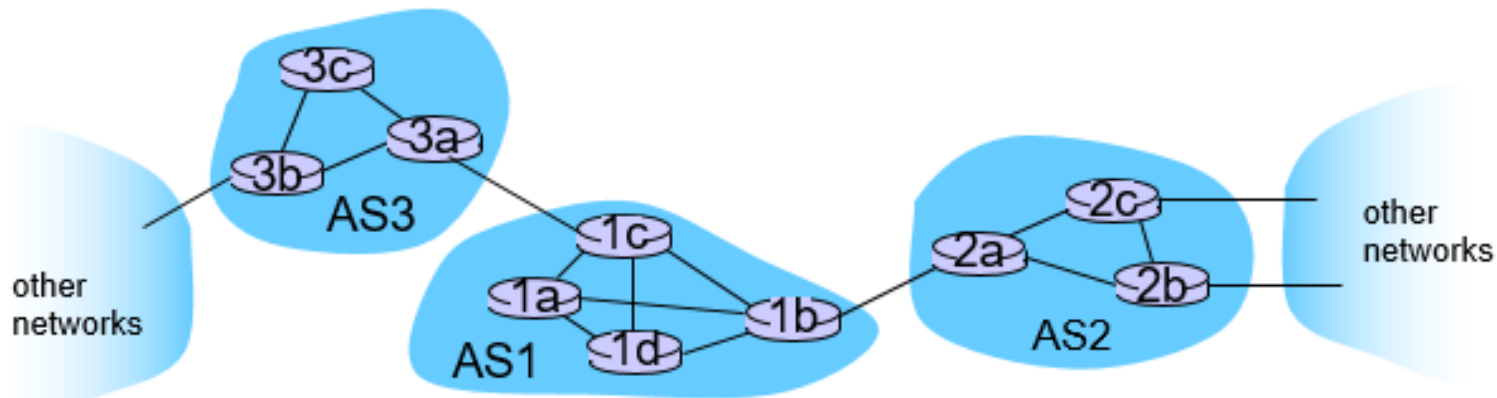
1. **规模：**以亿为数量级的目的网络，路由器的存储达不到这么大规模。路由表的交换，占用大量带宽
  2. **管理自治性：**互联网，网络的网络，每个网络管理员希望能自己控制网络内的路由
-

# AS



AS (Autonomous System)：由一组路由器构成的区域，通常处在相同的管理控制下

每个AS有全局的ASN（AS号），由ICANN区域注册机构分配



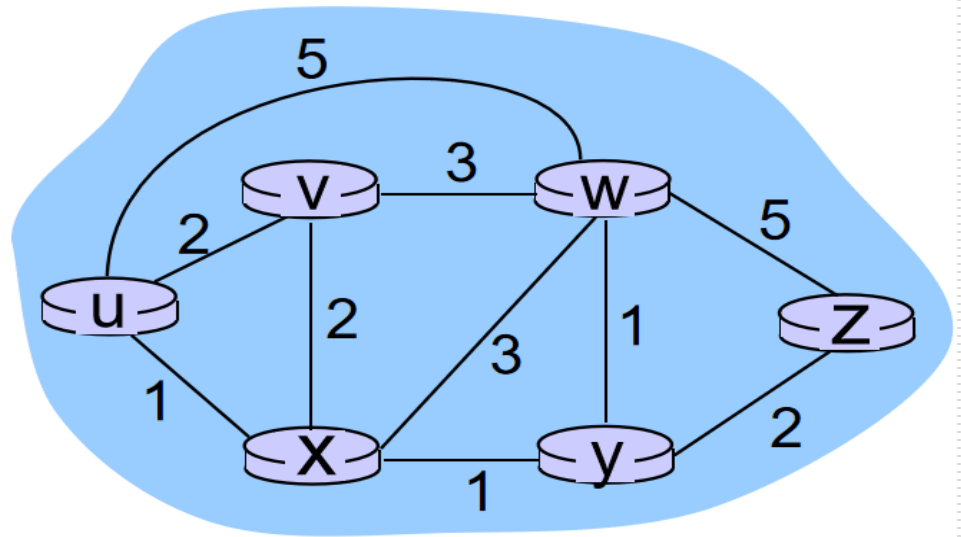
- **内部网关协议 (IGP) :** 在一个AS内进行路由计算, AS内所有路由器运行相同的路由协议: OSPF, RIP
- **外部网关协议 (EGP) :** AS之间的路由, BGP

# 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：DV算法
-





**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

# 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

---

# 例：更新路由表

---

路由器 R<sub>6</sub>

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

R<sub>4</sub> 发来的路由表

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

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

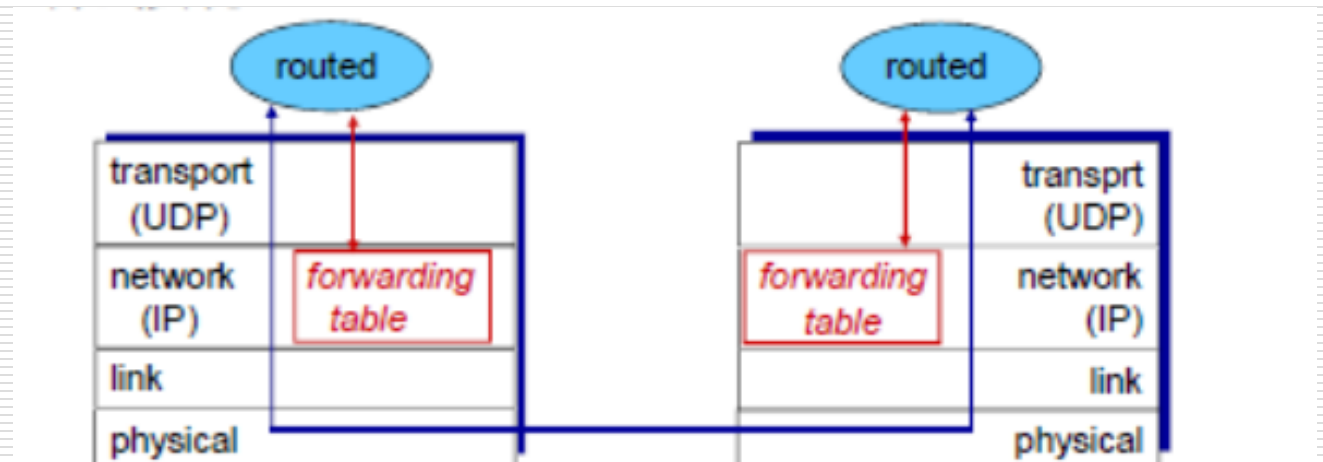
      : 不同的下一跳, 跳数相同, 不变

      : 不同的下一跳, 新跳数大, 不变

# RIP

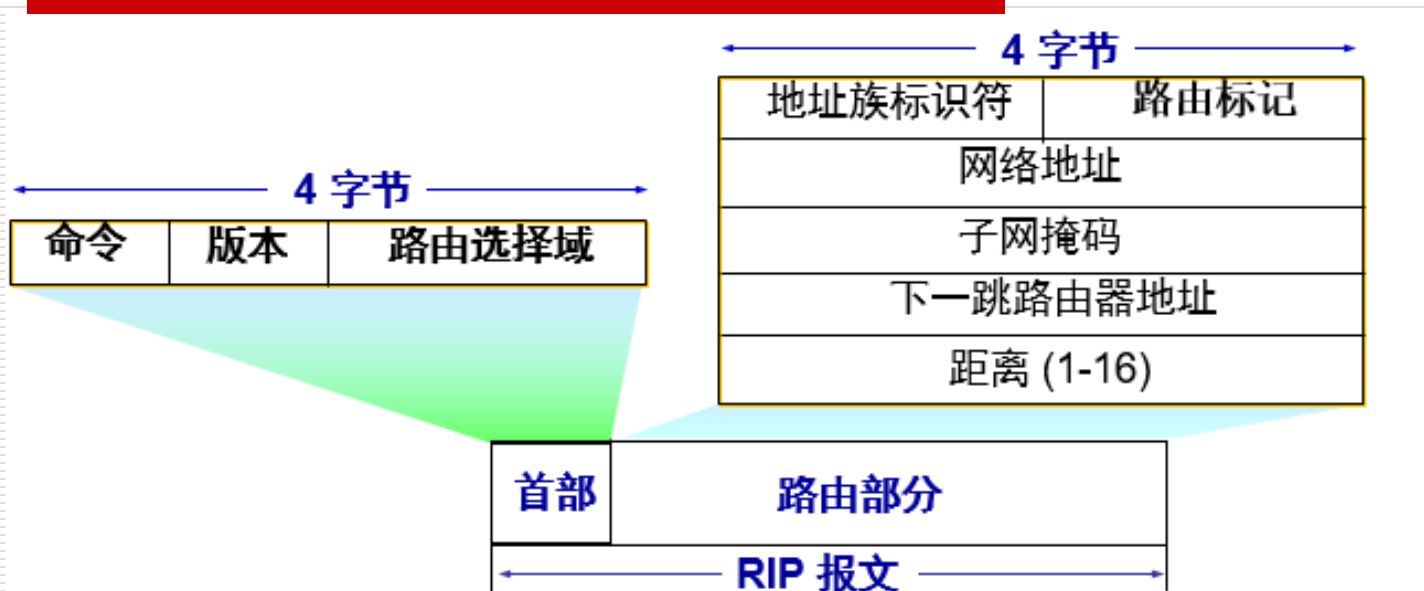
---

- 以应用进程的方式实现



- 通过UDP发送RIP报文
-

# RIP2报文格式



跳数为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）

## 2、OSPF

---

□ 开放最短路径优先（Open Shortest Path First）

✓ 采用LS算法

获得网络拓扑和  
链路代价信息

使用最短路由算法  
得到路由表

使用此路由表

---

# OSPF分组： 5种类型

---

类型1， 问候分组（Hello）： 发现和维持邻居路由器

类型2， 数据库描述分组（Database Description）： 向邻居发送本地的链路状态摘要信息

类型3， 链路状态请求分组（Link State Request）： 向邻居请求链路状态的详细信息

类型4， 链路状态更新分组（ Link State Update）： 广播发送链路更新状态

类型5， 链路状态确认分组（ Link State ACK）： 对更新分组/**hello**分组的确认

---

- 
- ✓ 通过IP直接发送，IP首部协议字段的值为89



OSPF2已成为互联网标准协议

---



---

## □ 路由协议（网络层的控制平面）

- 生成从源主机到目的主机的最佳路径
- 得到转发表

## ■ AS（自治系统）

### □ 内部网关协议

- RIP
- OSPF

### □ 外部网关协议

---

---

## □ RIP（路由信息协议）

- 距离向量路由算法（Distance Vector）
    - 相邻路由器交换信息——局部
    - Bellman-Ford计算最短路径
  - 距离：跳数（最大为16，表示网络不可达）
  - 应用进程，RIP报文通过传输层，到网络层.....
-

---

## □ OSPF(开放路径最短优先)

### ■ 链路状态路由算法（LS）

□ 每个路由器广播发送自身的链路状态——全局

□ Dijkstra计算最短路径

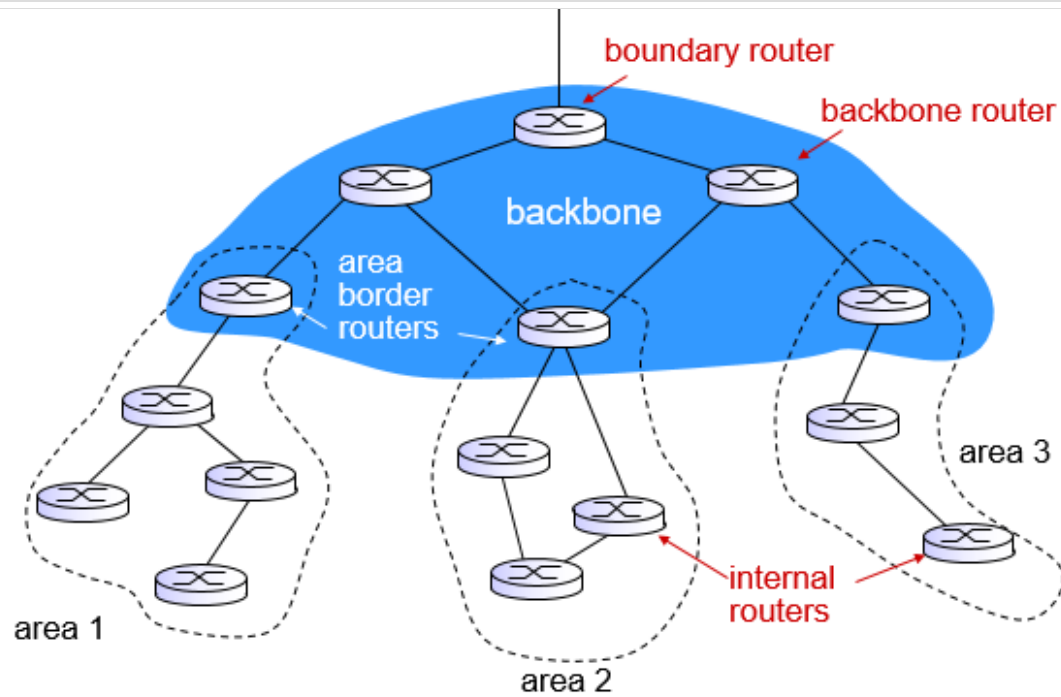
### ■ OSPF报文封装成IP分组（5种类型的分组）

---

# OSPF 特点

---

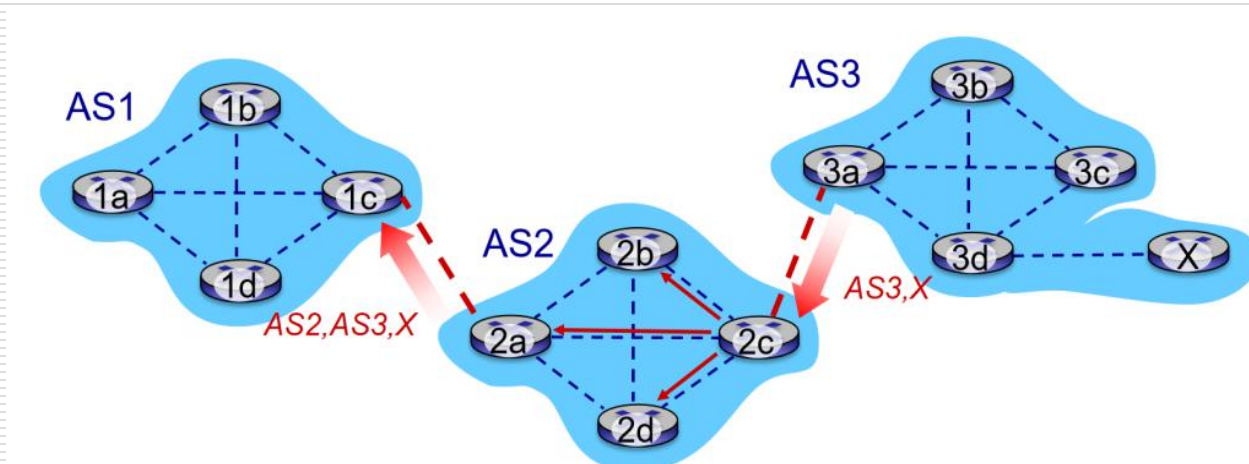
- ❑ 允许有多个代价相同的路径存在(在RIP协议中只有一个), 均衡负载;
  - ❑ 支持按照不同的代价（延时、带宽等）计算最优路径
  - ❑ 简单认证（安全）
  - ❑ 在大型网络中支持层次性OSPF
-

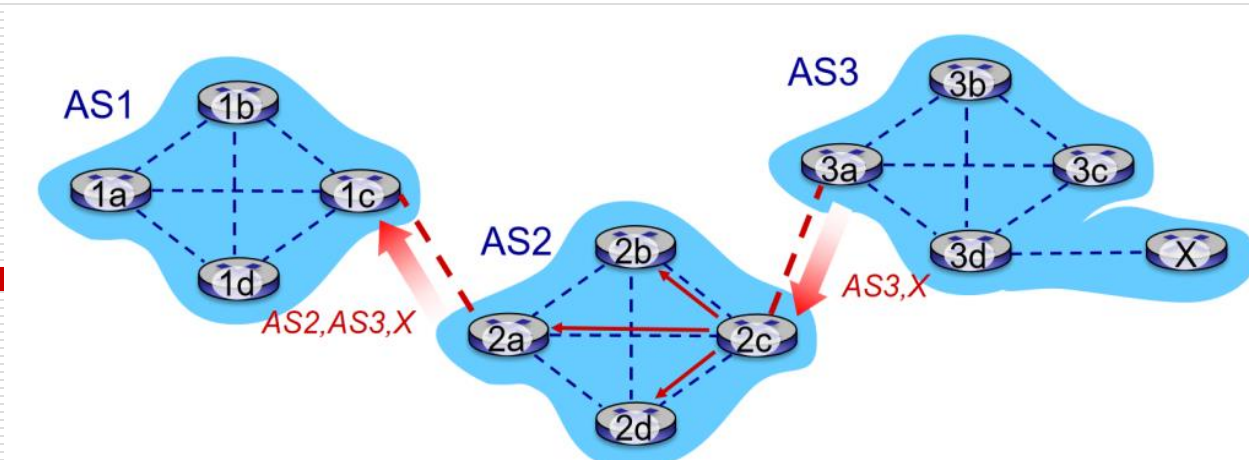


# 3、BGP

## □ 边界网关协议（Border Gateway Protocol）

- ✓ 不同AS之间通信的外部网关协议
- ✓ 目前使用的版本BGP-4

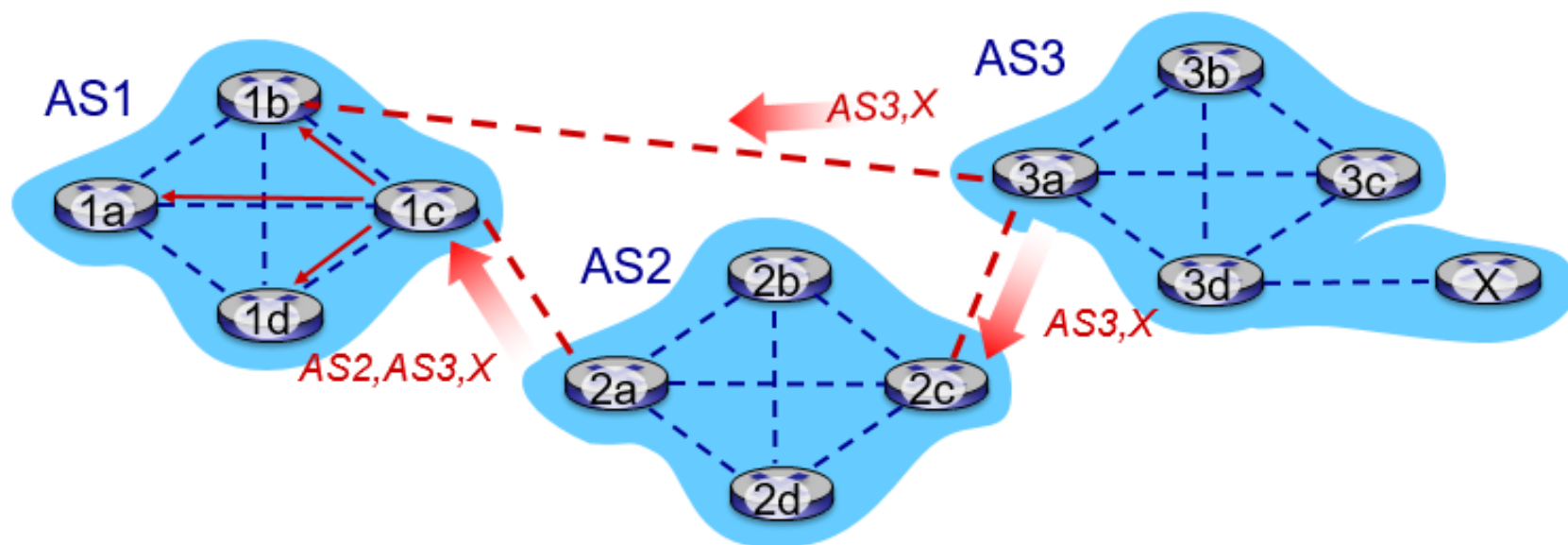




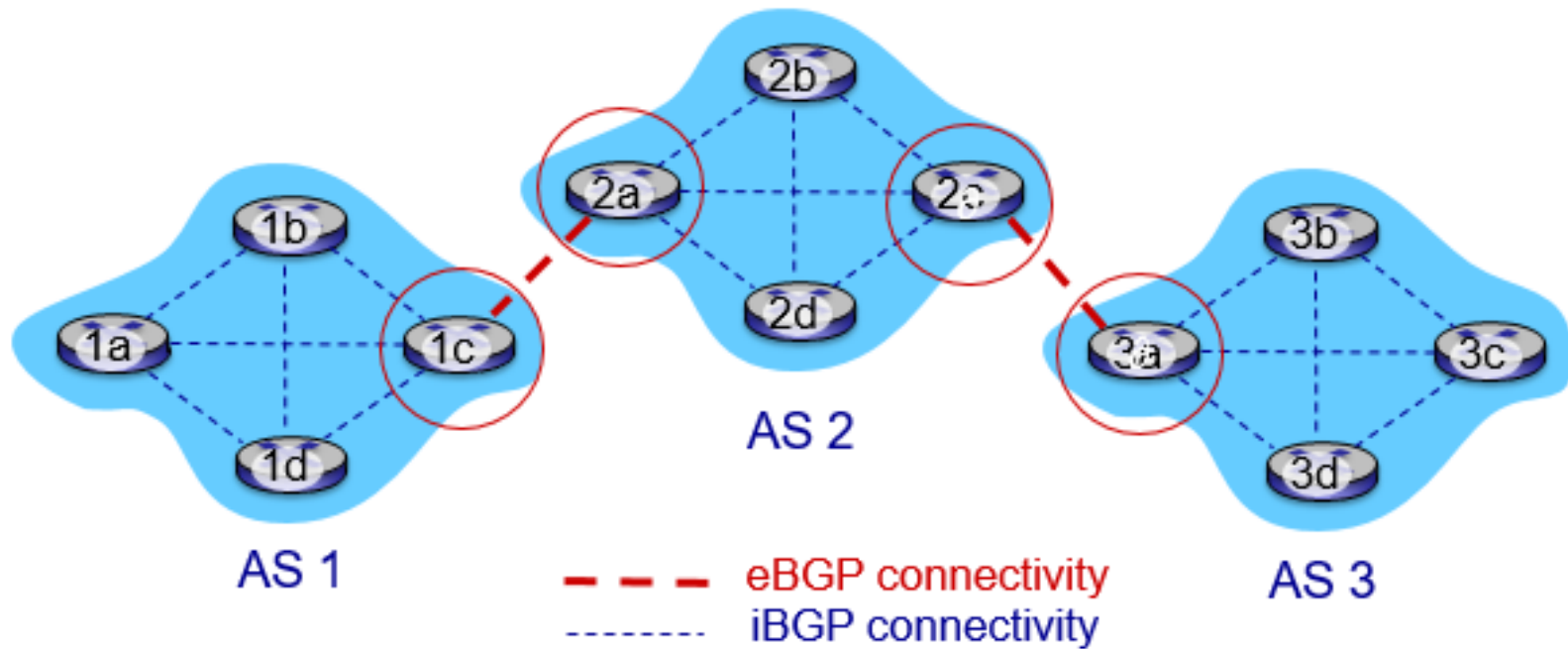
## BGP协议：

- 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）
- 路径向量路由算法（避免路由回路）

➤ 确定到达X的‘最好’路径（策略）







**eBGP:** 两个AS之间，从相邻的AS那里获得子网可达信息

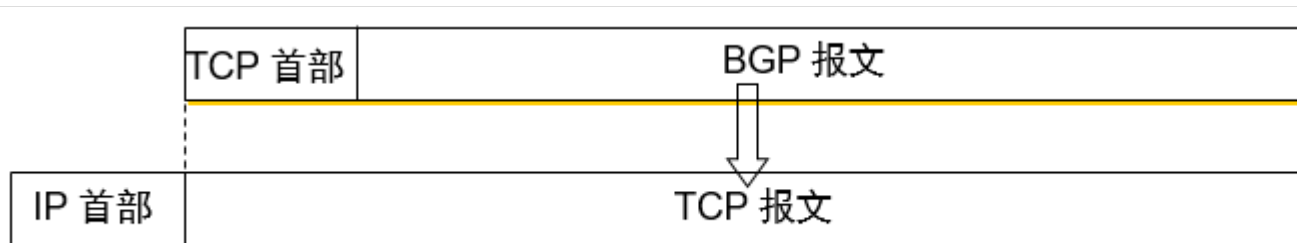
**iBGP:** 一个AS内部，将获得的子网可达信息传到AS内部的所有路由器

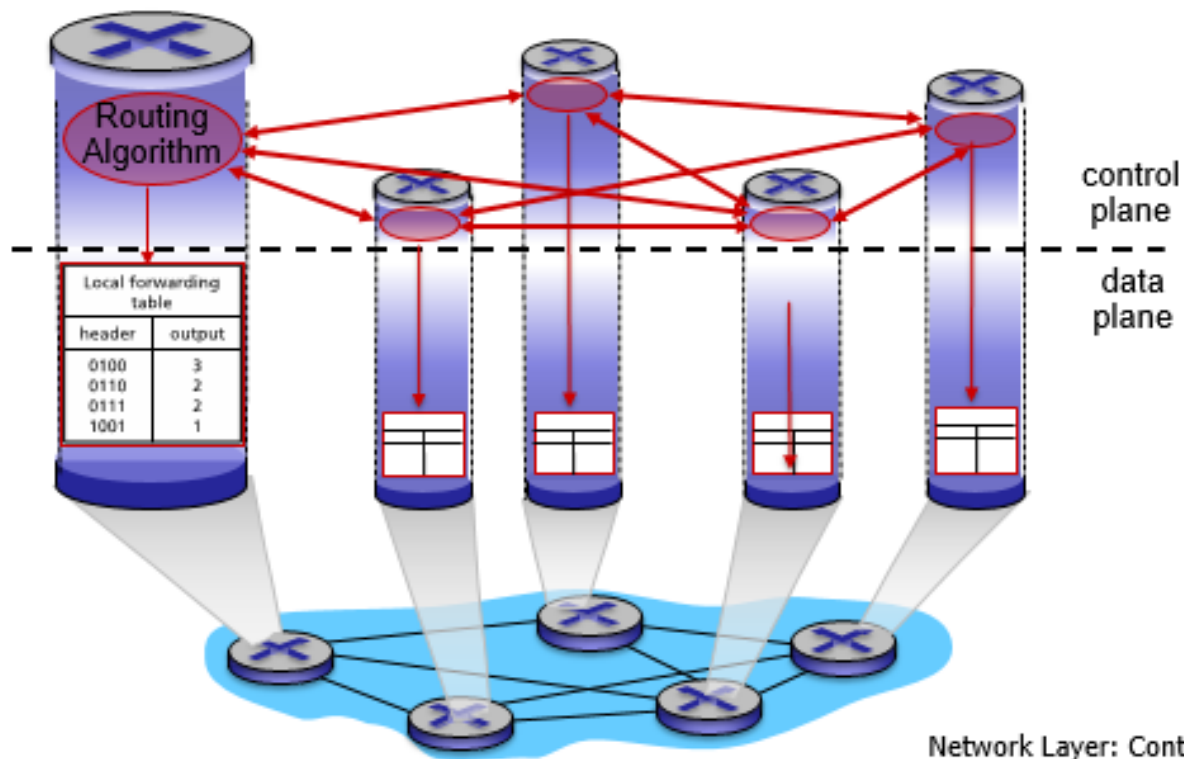
## □ BGP采用TCP传输

4种BGP报文：

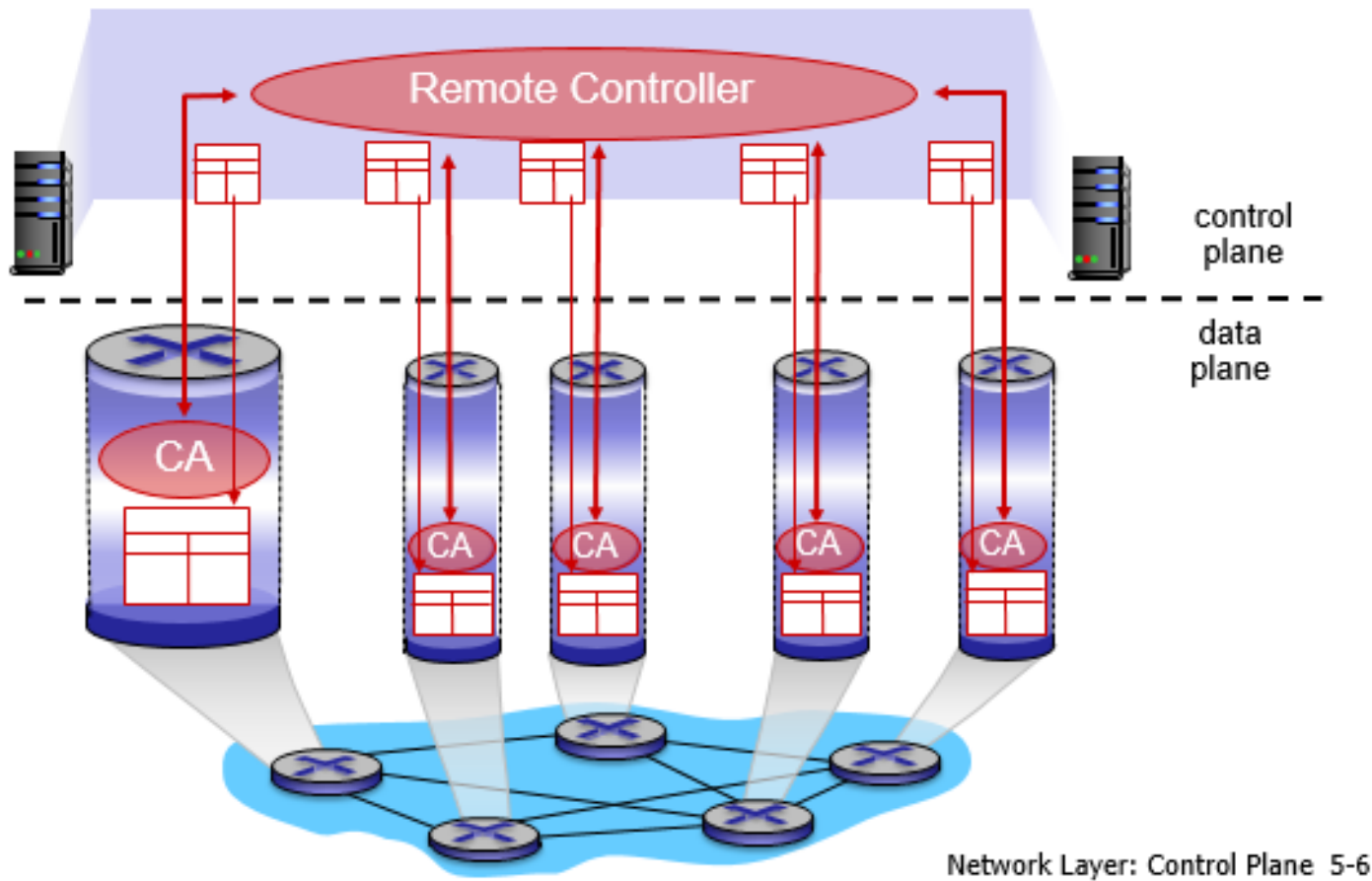
OPEN、UPDATE

KEEPALIVE、NOTIFICATION





传统方式：每个路由器独立运行路由算法，通过与其他路由器交换信息，构造转发表



**SDN（软件定义网络）：**一个或多个远程控制器通过与每个路由器上的**CA**交换信息，构造转发表

# 作业：

---

□ 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？

---

---

## □ 提交时间

- 11.12号（周日）

- 第四章（局域网）第五章（网络层）作业，合并在一个文档中