



4.1 网络层服务

4.2 虚电路vs数据报网络

4.3 路由器体系结构

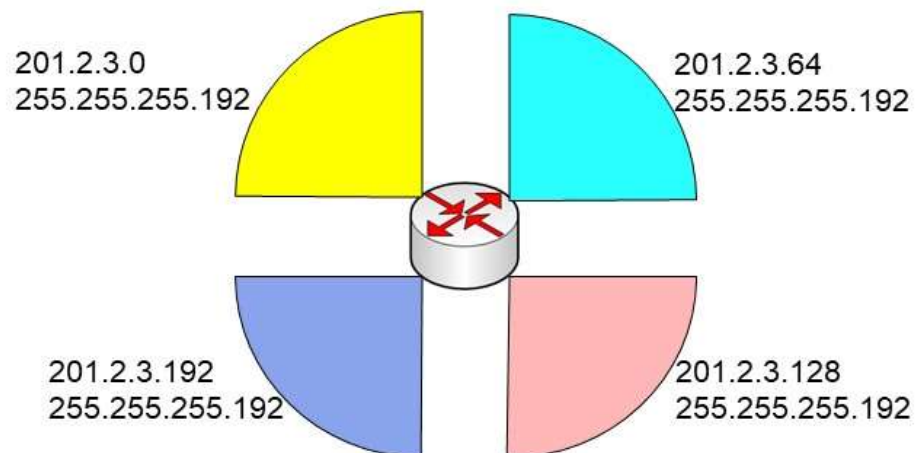
4.4 IP协议



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?



4.7 路由协议

刘亚维

340



4.1 网络层服务

4.2 虚电路vs数据报网络

4.3 路由器体系结构

4.4 IP协议

4.5 IP相关协议

4.6 路由算法

4.7 路由协议

RIP



AS内部路由

❖ Internet采用层次路由

❖ AS内部路由协议也称为内部网关协议**IGP** (interior gateway protocols)

❖ 最常见的AS内部路由协议:

- 路由信息协议: RIP(Routing Information Protocol)
- 开放最短路径优先: OSPF(Open Shortest Path First)
- 内部网关路由协议: IGRP(Interior Gateway Routing Protocol)
 - Cisco私有协议

341



4.1 网络层服务

4.2 虚电路vs数据报网络

4.3 路由器体系结构

4.4 IP协议

4.5 IP相关协议

4.6 路由算法

4.7 路由协议

RIP



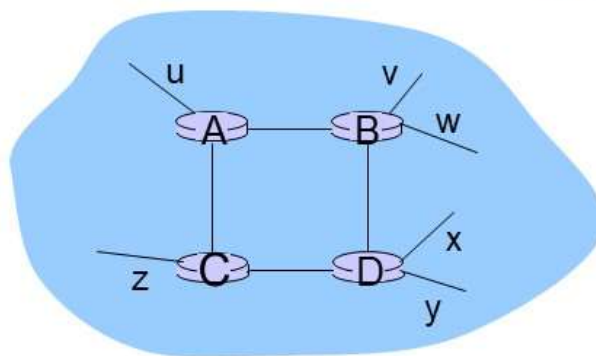
RIP

❖ 早于1982年随BSD-UNIX操作系统发布

❖ 距离向量路由算法

- 距离度量：跳步数 (max = 15 hops), 每条链路1个跳步
- 每隔30秒, 邻居之间交换一次DV, 称为通告(advertisement)
- 每次通告：最多25个目的子网(IP地址形式)

从路由器A到目的子网:



subnet	hops
u	1
v	2
w	2
x	3
y	3
z	2



4.1 网络层服务

4.2 虚电路vs数据报网络

4.3 路由器体系结构

4.4 IP协议

4.5 IP相关协议

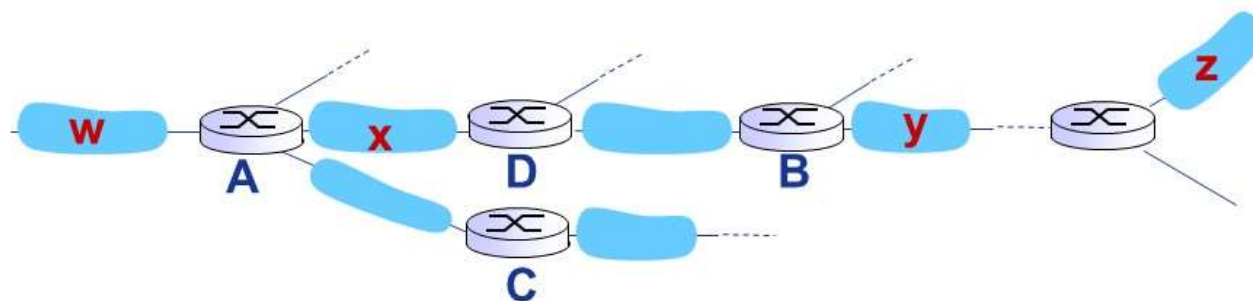
4.6 路由算法

4.7 路由协议

RIP



RIP: 举例



路由器D的路由表

destination subnet	next router	# hops to dest
W	A	2
Y	B	2
Z	B	7
X	--	1
....



4.1 网络层服务

4.2 虚电路vs数据报网络

4.3 路由器体系结构

4.4 IP协议

4.5 IP相关协议

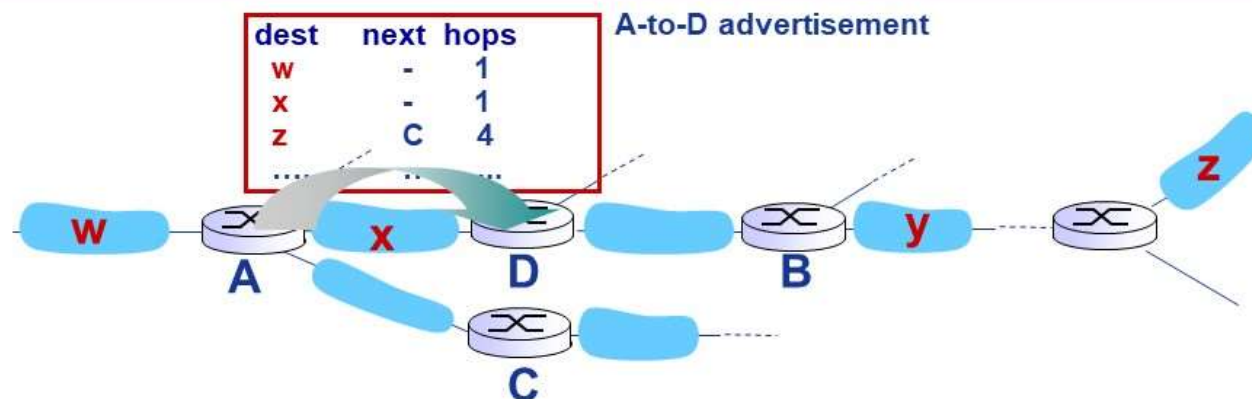
4.6 路由算法

4.7 路由协议

RIP



RIP: 举例



路由器D的路由表

destination subnet	next router	# hops to dest
W	A	2
Y	B	2
Z	B	7
X	--	1
....

Note: Red arrows in the original image point from 'A' in the next router column to 'B' for destination Z, and from '2' in the # hops column to '5' for destination Z.



4.1 网络层服务

4.2 虚电路vs数据报网络

4.3 路由器体系结构

4.4 IP协议

4.5 IP相关协议

4.6 路由算法

4.7 路由协议

RIP



RIP: 链路失效、恢复

如果180秒没有收到通告→邻居/链路失效

- 经过该邻居的路由不可用
 - 重新计算路由
- 向邻居发送新的通告
- 邻居再依次向外发送通告（如果转发表改变）
- 链路失效信息能否快速传播到全网？
 - 可能发生无穷计数问题
- **毒性逆转技术**用于预防乒乓(ping-pong)环路
(另外：无穷大距离 = 16 hops)

346



4.1 网络层服务

4.2 虚电路vs数据报网络

4.3 路由器体系结构

4.4 IP协议

4.5 IP相关协议

4.6 路由算法

4.7 路由协议

RIP

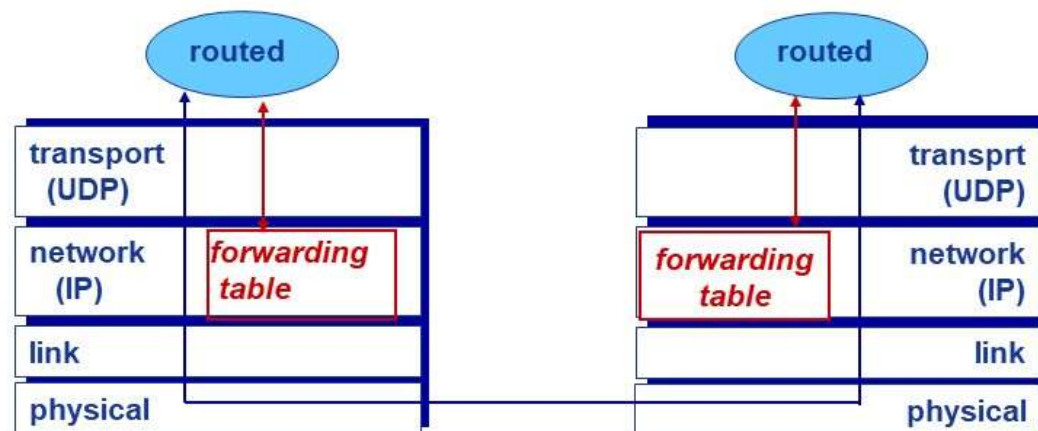


RIP路由表的处理

❖ RIP路由表是利用一个称作route-d (daemon)的**应用层**进程进行管理

❖ 应用进程实现

❖ 通告报文周期性地通过UDP数据报发送



347

单选题 1分

4.1 网络层服务

4.2 虚电路vs数据报网络

4.3 路由器体系结构

4.4 IP协议

4.5 IP相关协议

4.6 路由算法

4.7 路由协议

RIP



某自治系统内采用RIP协议，若该自治系统内的路由器R1收到其邻居路由器R2的距离矢量，距离矢量中包含信息<net1, 16>，则能得出的结论是

- ☐ A R2可以经过R1到达net1，跳数为17
- ☐ B R2可以到达net1，跳数为16
- ☐ C R1可以经过R2到达net1，跳数为17
- ☒ D R1不能经过R2到达net1

348



4.1 网络层服务

4.2 虚电路vs数据报网络

4.3 路由器体系结构

4.4 IP协议

4.5 IP相关协议

4.6 路由算法

4.7 路由协议

OSPF



OSPF (Open Shortest Path First)

- ❖ “开放”：公共可用
- ❖ 采用链路状态路由算法
 - LS分组扩散（通告）
 - 每个路由器构造完整的网络(AS)拓扑图
 - 利用Dijkstra算法计算路由
- ❖ OSPF通告中每个入口对应一个邻居
- ❖ OSPF通告在**整个AS**范围泛洪
 - OSPF报文直接封装到**IP**数据报中
- ❖ 与OSPF极其相似的一个路由协议：**IS-IS路由协议**

349



4.1 网络层服务

4.2 虚电路vs数据报网络

4.3 路由器体系结构

4.4 IP协议

4.5 IP相关协议

4.6 路由算法

4.7 路由协议

OSPF



OSPF优点(RIP不具备)

- ❖ **安全(security):** 所有OSPF报文可以被认证(预防恶意入侵)
- ❖ 允许使用**多条**相同费用的**路径** (RIP只能选一条)
- ❖ 对于每条链路, 可以针对不同的**TOS**设置多个不同的费用度量
 - e.g., 卫星链路可以针对“尽力” (best effort) ToS设置“低”费用; 针对实时ToS设置“高”费用
- ❖ **集成单播路由与多播路由:**
 - 多播OSPF协议(MOSPF) 与OSPF利用相同的网络拓扑数据
- ❖ **OSPF支持对大规模AS**分层(hierarchical)****

350



4.1 网络层服务

4.2 虚电路vs数据报网络

4.3 路由器体系结构

4.4 IP协议

4.5 IP相关协议

4.6 路由算法

4.7 路由协议

BGP

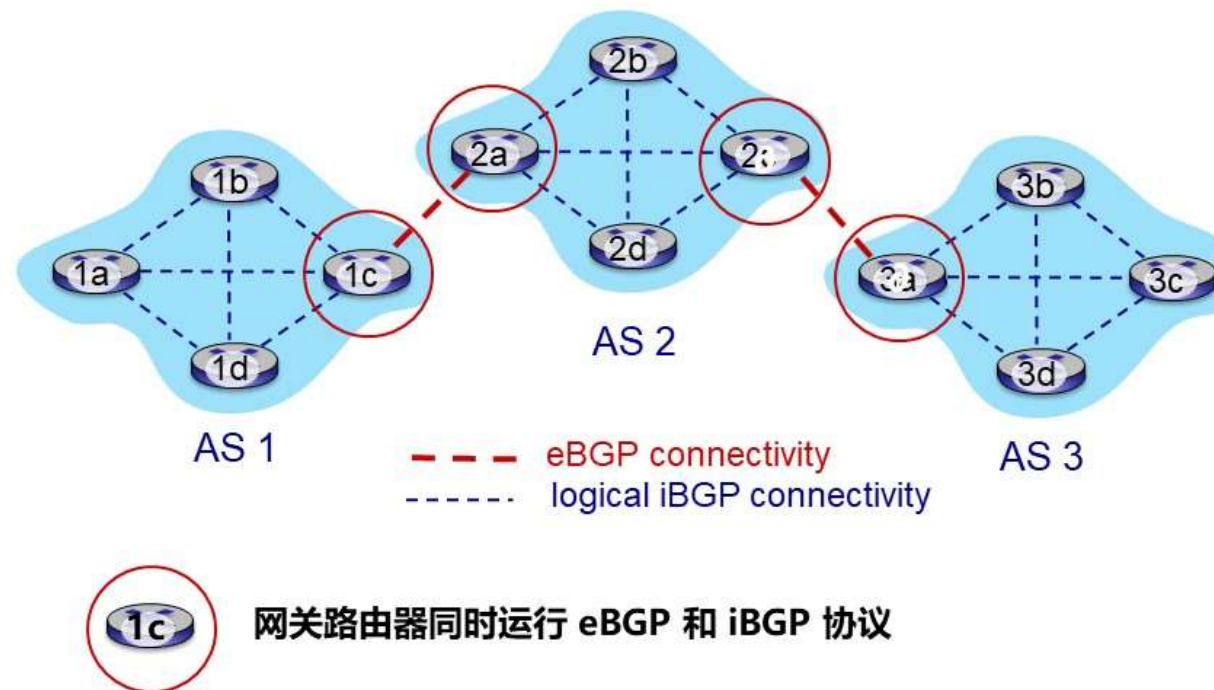


Internet AS间路由协议: BGP

- ❖ **边界网关协议BGP (Border Gateway Protocol): 事实上的标准域间路由协议**
 - 将Internet “粘合” 为一个整体的关键
- ❖ **BGP为每个AS提供了一种手段:**
 - **eBGP:** 从邻居AS获取子网可达性信息.
 - **iBGP:** 向所有AS内部路由器传播子网可达性信息.
 - 基于可达性信息与策略, 确定到达其他网络的 **“好”** 路径.
- ❖ **容许子网向Internet其余部分通告它的存在: 我在这儿!**

357

eBGP, iBGP连接



Network Layer: 5-359



4.1 网络层服务

4.2 虚电路vs数据报网络

4.3 路由器体系结构

4.4 IP协议

4.5 IP相关协议

4.6 路由算法

4.7 路由协议

BGP



BGP基础

❖ **BGP会话(session):** 两个BGP路由器 (“Peers”) 交换BGP报文:

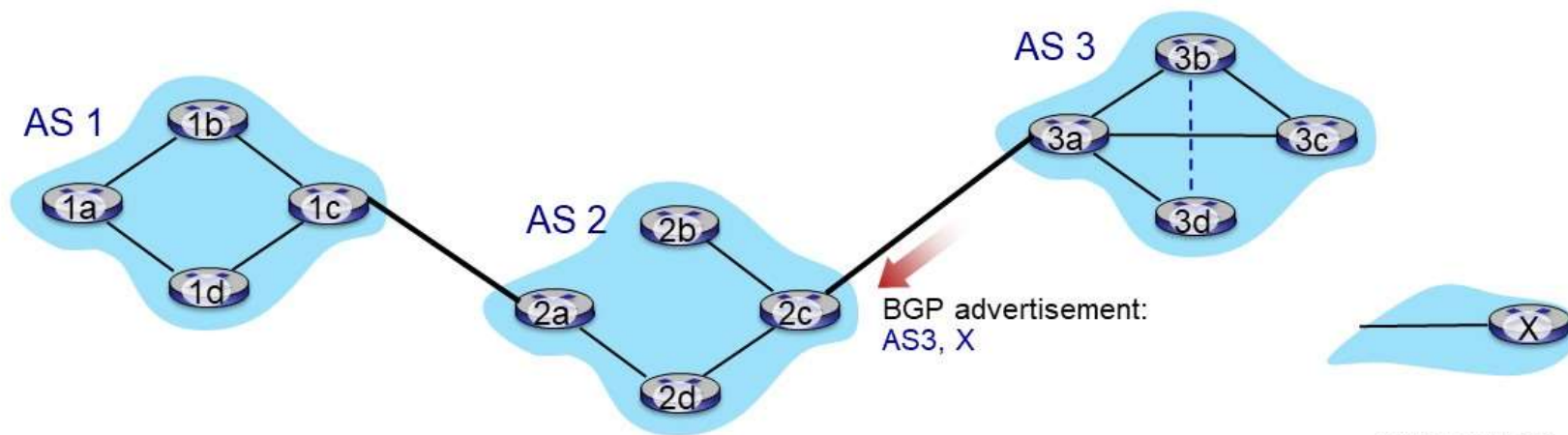
- 通告去往不同目的**前缀** (prefix) 的**路径** (“路径向量(path vector)” 协议)
- 报文交换基于半永久的**TCP**连接

❖ **BGP报文:**

- **OPEN:** 与peer建立TCP连接, 并认证发送方
- **UPDATE:** 通告新路径 (或撤销原路径)
- **KEEPALIVE:** 在无UPDATE时, 保活连接; 也用于对OPEN请求的确认
- **NOTIFICATION:** 报告先前报文的差错; 也被用于关闭连接

BGP 基础

- 当 AS3 网关 3a 通告**路径 AS3,X**到 AS2 网关 2c 时:
 - AS3 向 AS2 **承诺**会将数据报转发到 X



Network Layer: 5-363

BGP基础: 分发路径信息

4.1 网络层服务

4.2 虚电路vs数据报网络

4.3 路由器体系结构

4.4 IP协议

4.5 IP相关协议

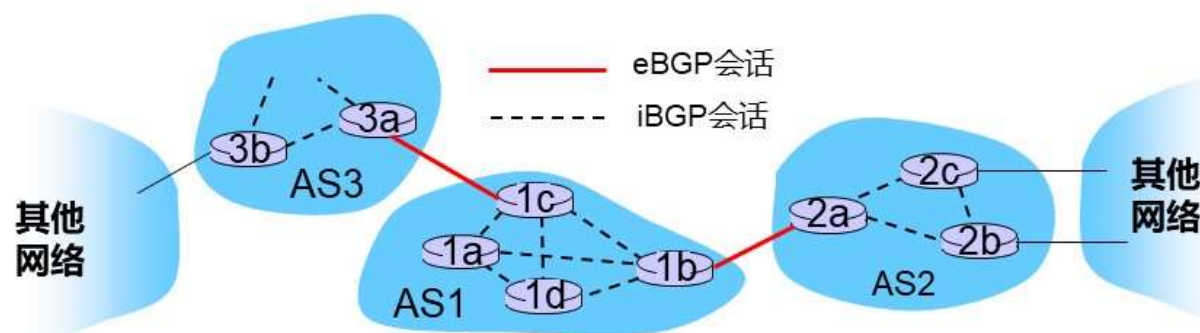
4.6 路由算法

4.7 路由协议

BGP



- ❖ 在3a与1c之间, AS3利用eBGP会话向AS1发送前缀可达性信息.
 - 1c则可以利用iBGP向AS1内的所有路由器分发新的前缀可达性信息
 - 1b可以(也可能不)进一步通过1b-到-2a的eBGP会话, 向AS2通告新的可达性信息
- ❖ 当路由器获得新的前缀可达性时, 即在其转发表中增加关于该前缀的入口(路由项)。



路径属性与BGP路由 (route)

4.1 网络层服务

4.2 虚电路vs数据报网络

4.3 路由器体系结构

4.4 IP协议

4.5 IP相关协议

4.6 路由算法

4.7 路由协议

BGP

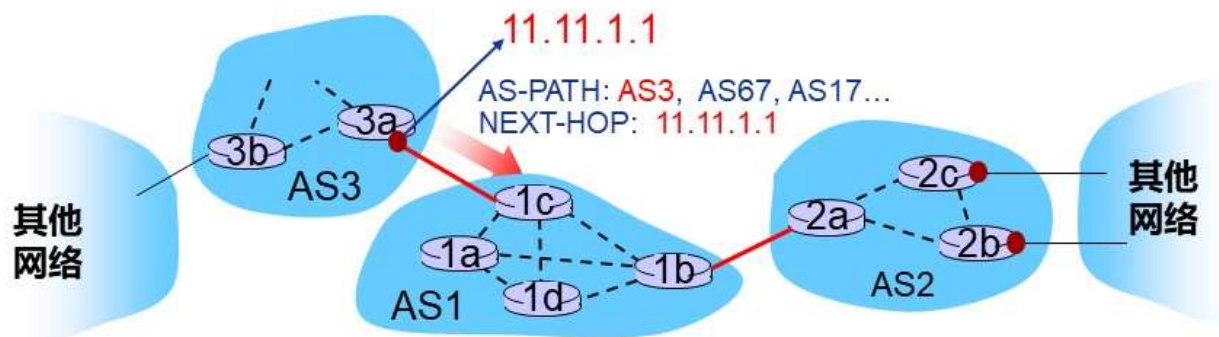


❖ 通告的前缀信息包括BGP属性

- 前缀+属性= “路由”

❖ 两个重要属性:

- **AS-PATH(AS路径):** 包含前缀通告所经过的AS序列: e.g., AS 67, AS 17
- **NEXT-HOP(下一跳):** 一个AS-PATH的起始路由器的接口, 指向下一跳AS.
 - 可能从当前AS到下一跳AS存在**多条链路**



366



4.1 网络层服务

4.2 虚电路vs数据报网络

4.3 路由器体系结构

4.4 IP协议

4.5 IP相关协议

4.6 路由算法

4.7 路由协议

BGP



BGP路由选择

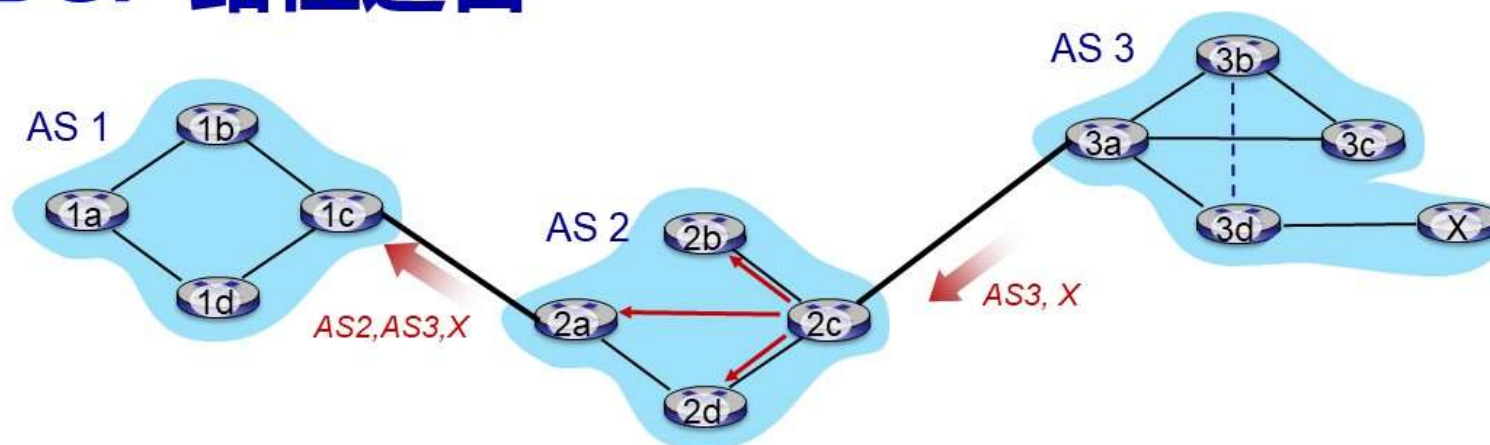
❖ 基于策略(policy-based) 路由

- 网关路由器收到路由通告后，利用其输入策略(import policy) 决策接受/拒绝该路由
- AS 策略还确定是否把路径通告到其他相邻 AS
- e.g., 从不将流量路由到AS x

❖ 路由器可能获知到达某目的AS的多条路由，基于以下准则选择：

1. 本地偏好(preference)值属性: 策略决策(policy decision)
2. 最短AS-PATH
3. 最近NEXT-HOP路由器: 热土豆路由(hot potato routing)
4. 附加准则

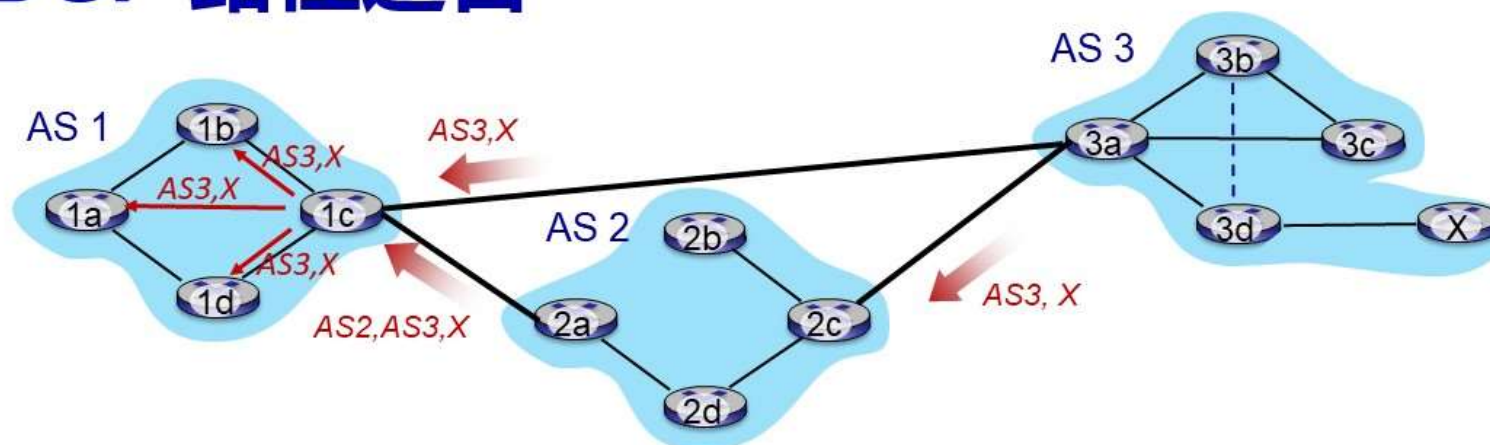
BGP 路径通告



- AS2 路由器 2c 从 AS3 路由器 3a 收到路径通告 **AS3, X** (通过 eBGP)
- 根据 AS2 策略, AS2 路由器 2c 接受路径 AS3, X, 传播 (通过 iBGP) 到所有 AS2 路由器
- 根据 AS2 策略, AS2 路由器 2a 通告 (通过 eBGP) 路径 **AS2、AS3、X** 到 AS1 路由器 1c

Network Layer: 5-369

BGP 路径通告



网关路由器可能知道通往目的地的多条路径：

- AS1 网关路由器 1c 从 2a 获知路径 **AS2, AS3, X**
- AS1 网关路由器 1c 从 3a 学习路径 **AS3, X**
- 根据**策略**，AS1网关路由器1c选择路径**AS3, X**，并通过iBGP在AS1内通告

Network Layer: 5-371

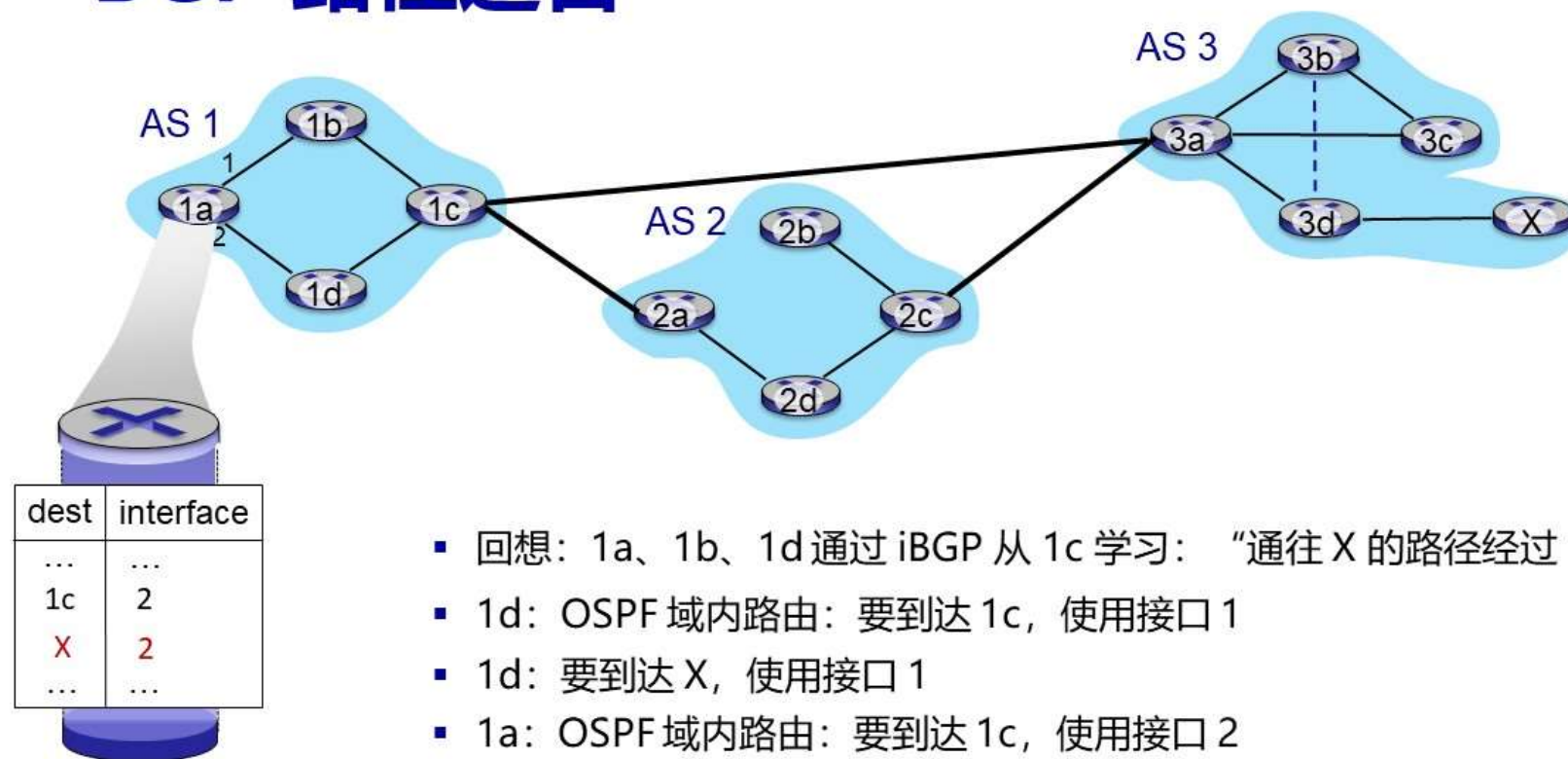
BGP 路径通告



- 回想：1a、1b、1d通过 iBGP 从 1c 学习：“通往 X 的路径经过 1c”
- 1d：OSPF 域内路由：要到达 1c，使用接口 1
- 1d：要到达 X，使用接口 1

Network Layer: 5-373

BGP 路径通告



- 回想：1a、1b、1d通过 iBGP 从 1c 学习：“通往 X 的路径经过 1c”
- 1d：OSPF 域内路由：要到达 1c，使用接口 1
- 1d：要到达 X，使用接口 1
- 1a：OSPF 域内路由：要到达 1c，使用接口 2
- 1a：要到达 X，使用接口 2

Network Layer: 5-374



4.1 网络层服务

4.2 虚电路vs数据报网络

4.3 路由器体系结构

4.4 IP协议

4.5 IP相关协议

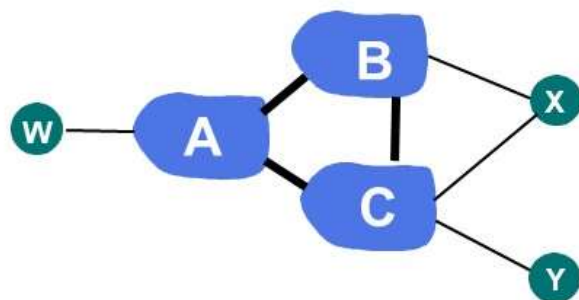
4.6 路由算法

4.7 路由协议

BGP



BGP路由选择策略



图例:  服务提供商网络
 客户网络

- ❖ A,B,C是**提供商网络**/AS(provider network/AS)
- ❖ X,W,Y是**客户网络**(customer network/AS)
- ❖ W,Y是**桩网络**(stub network/AS): 只与一个其他AS相连
- ❖ X是**双宿网络**(dual-homed network/AS): 连接两个其他AS
 - X不希望经过他路由B到C的流量
 - ... 因此, X不会向B通告任何一条到达C的路由



4.1 网络层服务

4.2 虚电路vs数据报网络

4.3 路由器体系结构

4.4 IP协议

4.5 IP相关协议

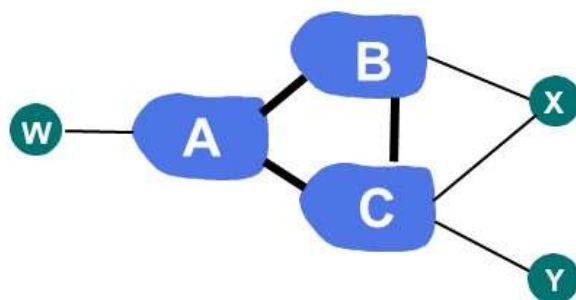
4.6 路由算法

4.7 路由协议

BGP



BGP路由选择策略



图例:  服务提供商网络

 客户网络

- ❖ A向B通告一条路径: AW
- ❖ B向X通告路径: BAW
- ❖ B是否应该向C通告路径BAW呢?
 - **绝不!** B路由CBAW的流量没有任何“收益”, 因为W和C均不是B的客户。
 - B期望强制C通过A向W路由流量
 - B期望只路由去往/来自**其客户**的流量!



4.1 网络层服务

4.2 虚电路vs数据报网络

4.3 路由器体系结构

4.4 IP协议

4.5 IP相关协议

4.6 路由算法

4.7 路由协议

为什么采用不同的AS内与AS间路由协议？

策略(policy):

- ❖ inter-AS: 期望能够管理控制流量如何被路由，谁路由经过其网络等.
- ❖ intra-AS: 单一管理，无需策略决策

规模(scale):

- ❖ 层次路由节省路由表大小，减少路由更新流量
- ❖ 适应大规模互联网

性能(performance):

- ❖ intra-AS: 侧重性能
- ❖ inter-AS: 策略主导



练一练

- ❖ 1. 请将192.168.1.0/24剩余IP地址分配给局域网1~3，其中局域网1需要IP地址数不少于60个，局域网2、3需要IP地址数不少于30个。说明局域网1~3的子网地址、广播地址、子网掩码、可分配IP地址数及范围。

