

### 3. ISP之间的路由选择 BGP

#### BGP 边界网关协议

##### 3.1 层次路由

###### 一个平面的路由

一个网络中所有路由器的地位一样

通过LS,DV算法,或者其他路由算法,所有路由器都要知道其他所有路由器(子网)如何走  
所有路由器在一个平面

问题:

规模巨大的网络中,路由信息的存储、传输和计算代价巨大

DV: 距离矢量很大, 而且不能够收敛

LS: 几百万个节点的LS分组的泛洪传输, 存储以及最短路径算法的计算

管理问题:

不同的网络所有者希望按照自己的方式管理网络

希望对外隐藏自己网络的细节

与其他网络互联

###### 层次路由

将互联网分成一个个AS(路由器区域)

某个区域内的路由器集合, 自治系统"autonomous system"

一个AS用AS Number(ASN)唯一标示

一个ISP可能包括1个或者多个AS

路由变成了: 2个层次路由

AS内部路由: 在同一个AS内路由器运行相同的路由协议

intra-AS routing protocol: 内部网关协议

不同的AS可能运行着不同的内部网关协议

能够解决规模和管理问题

如: RIP, OSPF, IGRP

网关路由器: AS边缘路由器, 可以连接到其他AS

AS间运行AS间路由协议:

inter-AS routing protocol: 外部网关协议

解决AS之间的路由问题, 完成AS之间的互相连通

优点:

解决了规模问题

内部网关协议解决: AS内部数量有限的路由器相互到达的问题, AS内部规模可控

如AS节点太多, 可分割AS, 使得AS内部的节点数量有限

AS之间的路由的规模问题

增加一个AS, 对于AS之间的路由总体上来说, 只是增加了一个节点=子网(每个AS可以用一个点来表示)

对于其他AS来说只是增加了一个表项, 就是这个新增的AS如何走的问题

扩展性强: 规模增大, 性能不会减太多

解决了管理问题

各个AS可以运行不同的内部网关协议

可以使自己的网络的细节不向外透露

##### 3.2 互联网AS间路由: BGP

BGP(Border Gateway Protocol): 边界网关协议--"事实上的"标准

将互联网各个AS粘在一起的胶水

BGP提供给每个AS以下方案:

eBGP 外部: 从相邻的ASes那里获得子网可达信息

**iBGP 内部：**将获得的子网可达信息传输到AS内部的所有路由器  
根据子网可达信息和决策来决定到达子网的"好"路径  
允许子网向互联网其他网络通告"我在这里"  
基于距离矢量算法(路径矢量)  
不仅仅是距离矢量，还包括到达各个目标网络的详细路径（AS序号的列表），能够避免简单DV算法的环路问题

**BGP会话：**

2个BGP路由器（"peers"）在一个半永久的TCP连接上交换BGP报文  
通告向不同目标的子网前缀的"路径"（BGP是一个"路径矢量"协议）

**路径的属性&BGP路由**

当通告一个子网前缀时，通告包括BGP属性

**prefix+attribute = route**

2个重要的属性： 路径+下一跳

**AS-PATH:**前缀的通告所经过的AS列表

监测环路；多路径选择

在向其他AS转发时，需要将自己的AS号加载路径上

**NEXT-HOP:**从当前AS到下一条AS有多个链路，在NEXT-HOP属性中，告诉对方通过那个I转发

其他属性：路由偏好指标，如何被插入的属性

**基于策略的路由：**

当一个网关路由器收到了一个路由通告，使用输入策略来接收或过滤（accept/decline）

过滤原因例1：

不想经过某个AS，转发某些前缀的分组

过滤原因例2：

已经有了一条往某前缀的偏好路径

策略也决定了是否向它别的邻居通告收到的这个路由信息

**BGP报文**

使用TCP协议交换BGP报文

**BGP报文：**

**OPEN:** 打开TCP连接，认证发送方

**UPDATE:** 通告新路径（或者撤销原路径）

**KEEPALIVE:** 在没有更新时保持连接，也用于对OPEN请求确认

**NOTIFICATION:** 报告以前消息的错误，也用来关闭连接

**路径选择**

路由器可能获得一个网络前缀的多个路径，路由器必须进行路径的选择，路由选择可以基于：

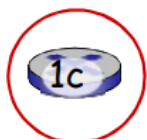
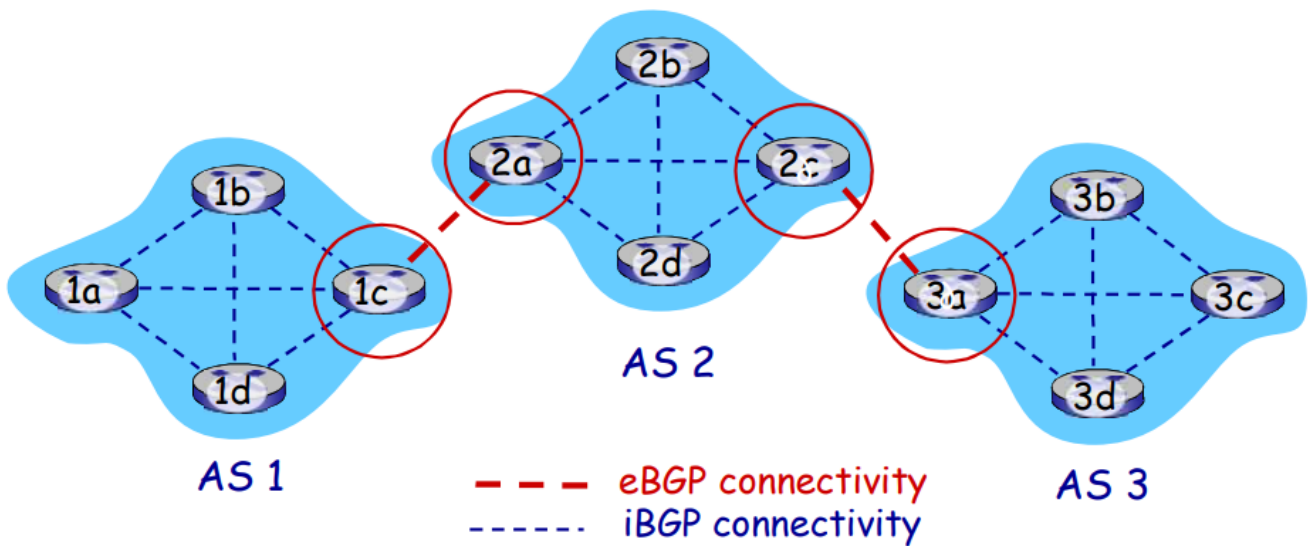
1. 本地偏好值属性：偏好策略决定
2. 最短AS-PATH：AS的跳数
3. 最近的NEXT-HOP路由器：热土豆路由
4. 附加的判据：使用BGP标示

一个前缀对应着多种路径，采用消除规则直到留下一条路径

如果区域内有多个网关，热土豆策略：

选择具备最小内部区域代价的网关作为往X的出口，不要操心域间的代价

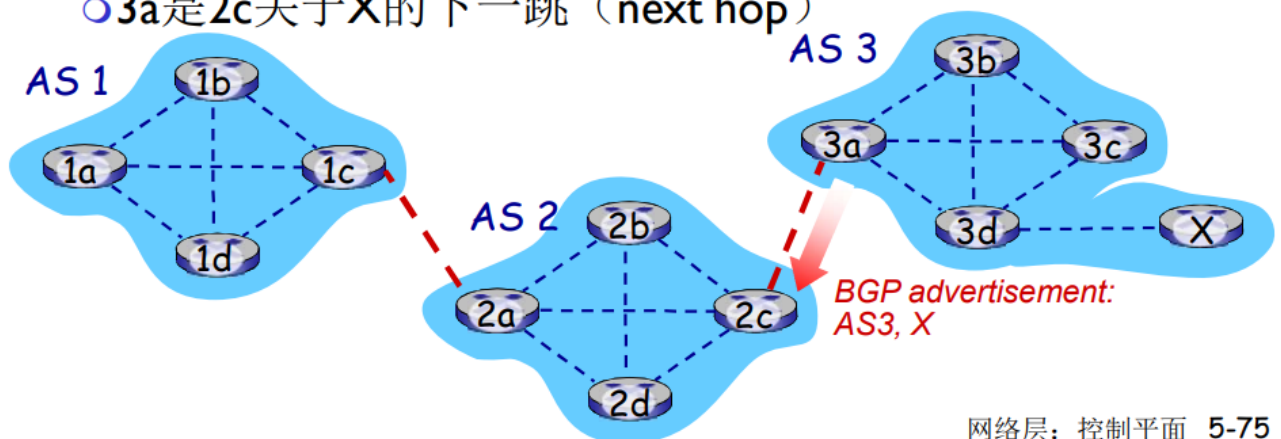
# eBGP, iBGP 连接



网关路由器同时运行eBGP和iBGP协议

□ 当AS3网关路由器3a向AS2的网关路由器2c通告路径: **AS3,X**

- 3a参与AS内路由运算, 知道本AS所有子网X信息
- 语义上: AS3向AS2**承诺**, 它可以向子网X转发数据报
- 3a是2c关于X的下一跳 (next hop)



网络层: 控制平面 5-75

## 3.3 为什么内部网关协议和外部网关协议如此不同

策略:

Inter-AS: 管理员需要控制通信路径, 谁在使用它的网络进行数据传输;

Intra-AS: 一个管理者, 所以无需策略; AS内部的各子网的主机尽可能地利用资源进行快速路由

规模:

AS间路由必须考虑规模问题, 以便支持全网的数据转发

AS内部路由规模不是一个大的问题

如果AS 太大, 可将此AS分成小的AS; 规模可控

AS之间只不过多了一个点而已

或者AS内部路由支持层次性，层次性路由节约了表空间，降低了更新的数据流量

性能：

Intra-AS：关注性能

Inter-AS：策略可能比性能更重要