

网络层控制平面

路由选择算法

路由的输入：拓扑、边的代价，源节点

输出的输出：源节点的汇集树

汇集树：

- 此节点到所有节点的最优路径形成的树
- 路由选择算法就是为了所有路由器找到并使用汇集树

路由选择算法的原则

正确性：算法必须是正确的和完整的

简单性：算法要求简单

健壮性：算法能适应通信量和网络拓扑的变化

稳定性：产生的路由不应该摇摆

公平性：对每一个站点都公平

最优性：某一个指标的最优

路由选择算法分类

link state 算法（全局）

过程

- 各点通过各个渠道获得整个网络拓扑结构，网络中所有链路代价等信息
- 使用最短路算法得到路由表
- 使用此路由表

链路状态路由选择

- 发现节点，获得网络地址
 - 一个路由器上电之后，向所有线路发送HELLO分组
 - 其他路由器收到HELLO分组，回应自己的名字
 - 通过HELLO广播获得其他路由器的信息
- 测量相邻节点的代价
 - 实测法，发送一个分组让对方立刻响应
 - 回送一个ECHO分组

- 组装一个分组，描述相邻节点的情况
 - 给出相邻节点和它相邻节点的延迟
- 将分组通过扩散的方法发到所有其他路由器
 - 顺序号
 - 年龄字段
- 通过Dijkstra算法找出最短路径

符号标记:

$c(i,j)$: 从节点 i 到 j 链路代价(初始状态下非相邻节点之间的链路代价为 ∞)

$D(v)$: 从源节点到节点 V 的当前路径代价(节点的代价)

$p(v)$: 从源到节点 V 的路径前序节点

N' : 当前已经知道最优路径的节点集合(永久节点的集合)

两类节点:

- 临时节点: 没有找到从源节点到此节点的最优路径节点
- 永久节点: 已经找到了从源节点到此节点的最优路径节点

distance vector 算法 (分布式)

距离矢量算法

使用Bellman-ford算法

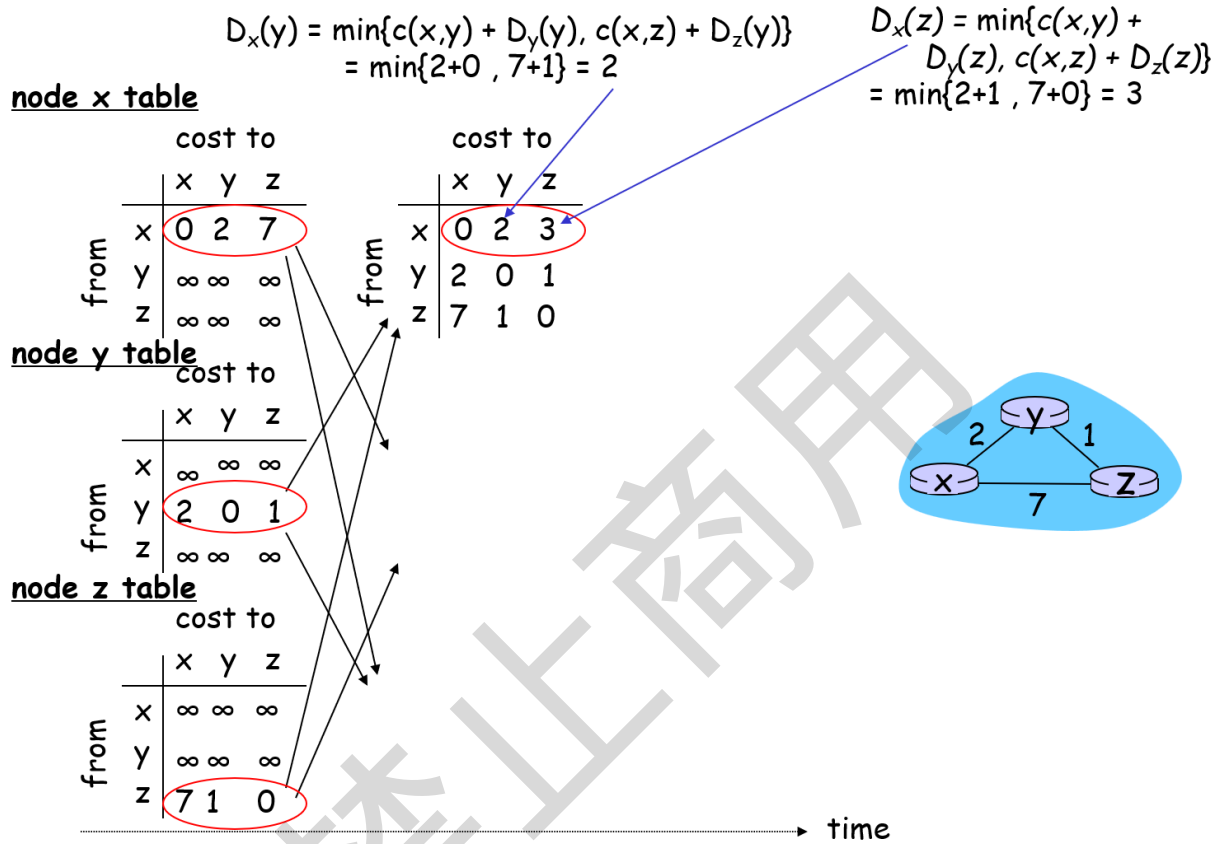
- $D_x(y)$ = 从 x 到 y 最小开销的估计
- $c(x,v)$ 节点 x 到相邻节点 v 的开销
- $D_x = [D_x(y): y \in N]$ 节点 x 维护的距离矢量
- 节点 x 维护它相邻节点的距离矢量
 - 对于每个相邻节点 v , x 维护
 $D_v = [D_v(y): y \in N]$

特点

- 每个节点向邻居发送距离矢量拷贝
- 异步的
- 当节点从邻居获得距离矢量之后更新自己的距离矢量

- “好消息”传得快：出现近距离路，路由更新快
- “坏消息”传得慢：路由稳定时间长
- 毒性逆转：
 - 定义最大度量：定义一个最大的有效的费用值
 - 当一条链路失效时将到达目标节点的费用算作无穷大，但有利于清除环路

算法图例



LS和DV算法的比较

消息复杂度：DV交换的信息少

- 假设n个节点，E条边，LS的报文O(nE)个
- DV只需要和邻居交换

收敛时间：LS更短

- LS使用O(n*n)
- DV收敛较慢，可能存在环路，并且有毒性逆转问题

健壮性：LS更好

- LS每个节点计算自己的链路代价
- DV可能会导致通告全网所有路由器不正确的链路代价

互联网中自治系统内的路由选择

RIP（路由信息协议）

使用DV算法，每30秒通告一次，每个通告最多25个子网，最大跳数为16

进程处理

报文通过UDP报文传送，周期性重复

OSPF（开放最短路径优先）

使用LS算法，通告信息中携带每一个邻居路由器的一个表项，通告信息会泛洪

IS-IS路由协议：几乎和OSPF相同

特点

- 安全：所有OSPF报文都是经过认证的
- 允许多个代价相同的路径存在（RIP中只允许有一个）
- 在大型网络中支持层次性OSPF

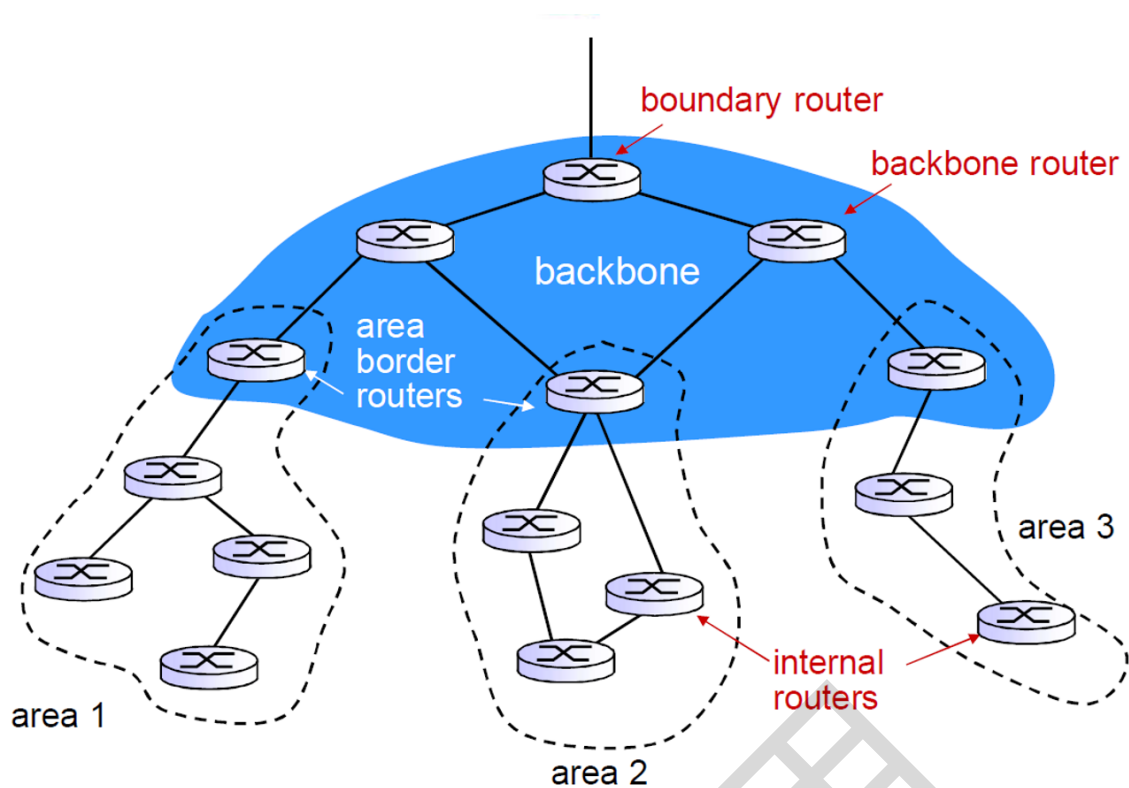
层次性OSPF

2个级别的层次性：本地，骨干

区域边界路由器：汇总到自己区域内网络的距离，向其他边界路由器通告

骨干路由器：仅在骨干区域内实行OSPF

边界路由器：连接AS



ISP之间的路由选择 (BGP)

BGP: 边界网关协议

层次路由的优点:

- 解决规模问题
- 解决管理问题

BGP给每个AS提供以下方法:

- eBGP: 从相邻的AS那里获得子网可达信息
- iBGP: 将获得的子网信息传给AS内部的所有路由器

基于DV算法

BGP会话: 在一个半永久的TCP连接上

路径的属性

□ 2个重要的属性:

- **AS-PATH:** 前缀的通告所经过的**AS**列表: **AS 67 AS 17**
 - ⑩ 检测环路; 多路径选择
 - ⑩ 在向其它**AS**转发时, 需要将自己的**AS**号加在路径上
- **NEXT-HOP:** 从当前**AS**到下一跳**AS**有多个链路, 在**NETX-HOP**属性中, 告诉对方通过那个 **I** 转发.
- 其它属性: 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**来接受或过滤 (**accept/decline.**)
 - ⑩ 过滤原因例1: 不想经过某个**AS**, 转发某些前缀的分组
 - ⑩ 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

OPEN: 打开TCP连接, 认证发送方

UPDATE: 通告新路径

KEEPALIVE: 在没有更新时保持连接

NOTIFICATION: 报告以前消息错误或者关闭连接

BGP路径选择

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

热土豆策略: 选择具备最小内部区域代价的网关作为X的出口

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