

网络层：控制平面

概述

1. 每路由器控制(Per-router):每个路由器中的单独路由器算法原件，在控制平面进行交互，路由器都包含转发和路由选择功能
2. 逻辑集中式控制(SDN):一个不同的(通常是远程的)控制器于本地控制代理(CAs)交互

路由选择算法

路由的概念：按照某种指标找到一条从源节点到目标节点的较好路径

1. 较好路径:按照某种指标较小的路径，如延迟、费用，站数
2. 以网络为单位进行路由
 1. 网络为单位进行路由，路由信息传输、计算和匹配代价低

路由的输入：拓扑、边的代价、源节点 输出的输出：源节点和汇集树 最优化原则(optimality principle)：

1. 汇集树(sink tree)
 1. 此节点到所有其他节点的最优路径形成的树
 2. 路由选择算法就是为所有路由器找到并使用汇集树

路由选择算法的原则

1. 正确性(correctness)
2. 简单性(simplicity)
3. 健壮性(robustness)
4. 稳定性(stability):产生的路由不应该摇摆
5. 公平性(fairness)：对每一个站点都公平
6. 最优性(optimality):某一个指标的最优，时间上，费用上等指标，或综合指标。

路由算法分类

1. 全局(centralized routing algorithm)：
 1. 所有路由器拥有完整的拓扑和边的代价的信息
 2. link state算法
2. 分布式(decentralized routing algorithm):
 1. 路由器只知道与它有物理连接关系的邻居路由器，和到相应邻居路由器的代价值
 2. 迭代地与邻居交换路由信息，计算路由信息
 3. distance vector算法
3. 静态：路由随时间变化缓慢
 1. 非自适应算法(non-adaptive algorithm):不能适应网络拓扑和通信量的变化，路由表事先计算好
4. 动态：路由变化很快；周期性更新；根据链路代价的变化而变化
 1. 自适应路由选择(adaptive algorithm):能适应网络拓扑和通信量的变化
5. 负载敏感(load-sensitive):链路开销会动态地变化以反映出底层链路地当前拥塞水平
6. 负载迟钝(load-insensitive)：某条链路的开销并不明确地反映当前的拥塞水平

LS路由工作工程

1. 获得整个网络拓扑，所有链路代价等信息

2. 使用LS路由算法计算最优路径
3. 按照此路由表转发分组

链路状态路由选择(link state routing)

LS路由的基本工作过程

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
3. 组装一个LS分组，描述它到相邻节点的代价
4. 将分组通过扩散的方法发到所有其他的路由器
5. 通过Dijkstra算法找出最短路径

LS应用

1. OSPF协议是一种LS协议，被用于Internet上
2. IS-IS(intermediate system-intermediate system)

距离矢量路由选择(distance vector routing,DV)

1. 代价及相邻节点间代价的获得
 1. 跳数(hops)、延迟(delay)、队列长度
 2. 相邻节点间代价的获得：通过实测
2. 定期测量到相邻节点的代价
3. 定期与相邻节点交换路由表(DV)
4. 更新路由表

Bellman-Ford方程： $dx(y) = \min_v \{c(x,v) + v(y)\}$ 基本思想：每个节点x以 $Dx(y)$ 开始，对在N中的所有节点y，估计从x到y地最低开销路径地开销

距离矢量算法

1. 核心思路：
 1. 每个节点都将自己的距离矢量估计值传送给邻居，定时或者DV有变化时，让对方去算
 2. 当x从邻居收到DV时，自己运算，更新它自己的距离矢量
 1. 采用B-F equation： $Dx(y) < \min_v \{c(x,v) + Dv(y)\}$
 2. x往y的代价 < -x到邻居v代价 + v声称到y的代价
2. 异步式、迭代，有以下事件触发：
 1. 本地链路代价变化了
 2. 从邻居来了DV的更新信息
3. 分布式：每个节点只是在自己的DV改变后向邻居通告
4. 等待->重新计算->通告
5. 水平分裂(split horizon)算法：
 1. 路由选择环路(routing loop)
 2. 增加毒性逆转(poisoned reverse)：如果z通过y路由选择到目的地x，则z将通告y，它(即z)到x的距离是无穷大；用来解决环路

LS和DV算法的比较

1. 报文复杂性：
 1. LS：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

2. DV : 只和邻居交换信息 ; 全局的路由信息, 局部传播
2. 收敛时间
 1. LS: $O(n^2)$
 2. DV : 收敛缓慢
 3. 可能存在路由环路
 4. 无穷计算问题
3. 健壮性 : 在LS算法下, 路由计算在某种程度上是分离的, 提供了一定程度的健壮性。

因特网中自治系统内部的路由选择 : OSPF

RIP(Routing Information Protocol)

1. DV:在邻居之间每30秒交换通告报文
 1. 定期, 而且在改变路由的时候发送通告报文
 2. 在对方的请求下可以发送通告报文
2. 没有收到通告信息-邻居或者链路失效
 1. 发现路由失效
 2. 新的通告报文传递邻居(通告报文采用UDP报文, 以应用进程的方式实现)
 3. 邻居发出新通告
 4. 毒性逆转组织ping-pong回路

OSPF(Open Shortest Path First) : 开放最短路径优先

自治系统(Autonomous System,AS):有一组通常处在相同管理控制下的路由器组成 自治系统内部路由选择协议 (intra-autonomous system routing protocol):相同AS中的路由器都运行相同的路由并且有彼此的信息, 在一个自治系统内运行的路由选择算法

OSPF

1. 一种链路状态协议, 使用洪泛链路状态信息和Dijkstra最低开销路径算法
2. OSPF通告信息中携带 : 每一个邻居路由器一个表项
3. 通告信息会传遍AS全部(通过泛洪)
4. OSPF通告包含在OSPF报文中, 由IP直接承载
5. 具有报文传输、链路状态广播、检查链路正在运行、并允许OSPF路由器获得相邻路由器的网络范围链路状态的数据库等功能
6. 优点 :
 1. 安全 : 能够鉴别OSPF路由器之间的交换, 仅有受信任的路由器能参与一个AS内的OSPF协议。
MD5鉴别基于配置在所有路由器上的共享秘密密钥
 2. 多条相同开销路径 : 当存在多条相等开销的路径时无需仅选择单一的路径来承担所有的流量
 3. 对单播与多播路由选择的综合支持
 4. 支持在单个AS中的层次结构 : 每个区域都运行自己的OSPF链路状态路由选择算法, AS中只有一个OSPF区域配置成主干区域, 主干区域的主要作用是为该AS中其他区域之间的流量提供路由选择, 该主干总是包含本AS中的所有区域边界路由器

ISP之间的路由选择 : BGP

边界网关协议(Border Gateway Protocol,BGP)

1. 在Internet中, 所有的AS运行相同的AS间路由选择协议, 是一种分布式和异步的协议
2. BGP提供给每个AS以下方法 :

1. eBGP：从相邻的ASes获得子网可达信息
2. iBGP：将获得的子网可达信息传遍到AS内部的所有路由器
3. 根据子网可达信息和策略来决定到达子网的路径
3. 网关路由器(gateway router):位于AS边缘的路由器，直接连接到其他AS中的一台或多台路由器
4. 内部路由器(internal router):仅连接在它自己AS中的主机和路由器
5. 两个属性：
 1. AS-PATH:前缀的通告所经过的AS列表
 1. 检测环路：多路径选择
 2. 在向其他AS转发时，需要将自己的AS号加在路径上
 2. NEXT-HOP:AS-PATH起始的路由器接口的IP地址
 3. 目的前缀
6. 热土豆路由选择(hot potato routing)
 1. 选择的路由到开始该路由的NEXT-HOP路由器有最小开销，即选择具有最小最低开销的网关
 2. 当在转发表中增加AS向外前缀时，AS间路由选择协议(BGP)和AS内部路由选择协议(如OSPF)都要用到
 3. 思想：尽可能快地将分组送出其AS(用可能的最低开销)而不担心其AS外部到目的地余下部分的开销
7. 路由器选择算法
 1. 路由被指派一个本地偏好(local preference)值作为其属性之一，具有最高本地偏好值的路由将被选择
 2. 所有都具有相同的最高本地偏好值时，选择具有最短AS-PATH的路由，距离测度使用AS跳的跳数
 3. 都相同的情况下使用热土豆路由选择，即选择具有最靠近NEXT-HOP路由器的路由
 4. 还有，则用BGP标识符

SDN控制平面

SDN体系结构的4个关键特征

1. 基于流的转发：SDN控制的交换机的分组转发工作，能够基于运输层、网络层或链路层首部中任意数量的首部字段值进行
2. 数据平面与控制平面分离：数据平面由网络交换机组成，控制平面由服务器以及决定和管理交换机流表的软件组成
3. 网络控制功能：位于数据交换机外部，控制器维护准确的网络状态信息
4. 可编程的网络

SDN控制器功能的三个层次

1. 通信层：SDN控制器和受控网络设备之间的通信，称为控制器的“南向”接口
2. 网络范围状态管理层：由SDN控制平面所做出的最终控制决定，将要求控制器具有有关网络的主机、链路、交换机和其他SDN控制设备的最新状态信息。
3. 对于网络控制应用层的接口：控制器通过“北向”接口与网络控制应用程序交互

OpenFlow协议

1. 控制器和SDN交换机交互的协议
2. 采用TCP来交换报文,使用6653默认端口号
3. 3种OpenFlow报文类型
 1. 控制器>交换机
 2. 异步(交换机>控制器)

3. 对称(misc)
4. 一些关键的控制器到交换机的报文
 1. 配置：该报文允许控制器查询并设置交换机的配置参数
 2. 修改状态：控制器增加/删除或修改交换机流表中的表项并设置交换机端口特性
 3. 读状态：控制器用于从交换机的流表和端口收集统计数据和计数器值
 4. 发送分组：控制器用于在受控交换机从特定端口发送一个特定保温
5. 一些重要的受控交换机到控制器的报文
 1. 流删除：通知控制器已删除一个流表项
 2. 端口状态：向控制器通知端口状态的变化
 3. 分组入：将分组发送给控制器
6. 网络服务应用程序：用于决定数据平面转发和其他服务如何在受控交换机中完成
7. SAL：控制器的神经中枢，允许控制器组件和应用程序相互调用服务并订阅它们产生的事件
8. ODL的基本网络服务功能：OpenDaylight控制器的核心
9. ONOS控制器的三个层次：
 1. 北向抽象和协议：ONOS的意图框架允许应用程序请求高层服务而不必知道该服务的执行细节，或者同步/异步地经过北向API向网络控制的应用程序提供状态信息
 2. 分布式核：维护了网络的链路、主机和设备的状态
 3. 南向抽象和协议：屏蔽了底层主机、链路、交换机和协议的异构性，允许分布式核对设备和协议不可知

ICMP:因特网控制报文协议

1. 主要用途：差错报告
2. ICMP报文承载在IP分组中，即ICMP报文作为IP分组的有效载荷
3. ICMP报文有一个类型字段和一个编码字段，并且包含引起该ICMP报文首次生成的IP数据报的首部和前8个字节

网络管理

1. 定义：包括了硬件、软件和人类元素的设置、综合和协调，以监视、测试、轮询、配置、分析、评价和控制网络及网元资源，用合理的成本满足实时性、运营性能和服务质量的要求
2. 五大功能：
 1. 性能管理
 2. 故障管理
 3. 配置管理
 4. 账户管理
 5. 安全管理
3. 网络管理框架：
 1. 管理服务器(managing server):执行网络管理活动，控制网络管理信息的手机、处理、分析和/或显示
 2. 被管设备(managed device)
 3. 管理信息库(Management Information Base,MIB):一个被管设备中每个被管对象的关联信息收集在此。
 4. 网络管理代理(network management agent):在每个被管设备中，在管理服务器的命令和控制下在被管设备中采取本地动作
 5. 网络管理协议(network management protocol)：运行在管理服务器和被管设备之间，允许管理服务器查询被管设备的状态1，并经过其代理间接地在这些设备上采取行动。

简单网络管理协议(Simple Network Management Protocol,SNMP)

1. 类型：应用层协议
2. 作用：在管理服务器和代理管理服务器执行的代理之间传递网络管理控制和信息报文
3. 常用：
 1. 请求响应模式：SNMP管理服务器向代理发出请求，通常是改变MIB对象值，代理执行后回答
 2. 陷阱报文(trap message):代理向管理服务器发送的一种非请求报文，用于通知管理服务器一个异常情况已经导致了MIB对象值的改变