自治系统

开始之前，我们先介绍下自治系统的概念。

互联网连接着世界各地的组织结构，不仅包括语言不相通的，甚至包括宗教信仰完全不同的组织，这里面没有管理者，也没有被管理者，每个组织之间保持着平等的关系。制定自己的路由策略，并以此为准在一个或多个网络群体中采用的较小单位叫做自治系统（AS：Autonomous System）。

说到自治系统，区域网络、ISP（网络服务提供商，如电信、联通、移动） 都是典型的例子。在区域网络或及 ISP 内部，由构造、管理和运维网络的管理员、运营者制定出路由控制相关方针，然后根据此方针进行具体路由控制的设定。而接入到区域网络或 ISP 的组织机构，则必须根据管理员的指示进行路由控制设定，否则会给自己或其他使用者带来负面影响，甚至无法通信。

每个自治系统都有边界路由器，通过它和外面的世界建立联系。

自治系统内部动态路由采用的协议是内部网关协议（Interior Gateway Protocol），简称 IGP。而自治系统之间的路由控制采用的是外部网关协议（Exterior Gateway Protocol），简称 EGP。

IP 地址分为网络地址和主机地址，它们有各自的分工，EGP 和 IGP 的关系和 IP 地址的两部分类似。就像 IP 地址的网络部分用于在网络之间进行路由选择、主机部分用于链路内主机识别一样，可以根据 EGP 在自治系统之间进行路由选择，根据 IGP 在自治系统内部进行主机识别。

由此，路由协议分为 EGP 和 IGP 两个层次，没有 EGP 就不可能有世界上不同组织机构之间的通信，没有 IGP 机构内部也就不能通信。

IGP 中可以使用 RIP、OSPF 等众多协议。与之相对的，EGP 使用的是 BGP 协议。

下面我们就来详细介绍上面三个常见的路由协议。

RIP

RIP（Routing Information Protocol，路由信息协议）基于距离向量算法，是距离向量型的一种路由协议，广泛用于局域网。

RIP 将路由控制信息定期（30秒）向全网广播。如果没有收到路由控制信息，连接就会被断开，不过这有可能是丢包导致的，所以 RIP 规定等待 5 次。

RIP 基于距离向量算法决定路径。距离的单位为跳数（Metrics），跳数是指所经过路由器的个数。RIP希望尽可能少经过路由器将数据包转发到目标 IP 地址。

OSPF

OSPF（Open Shortest Path First，开放式最短路径优先）基于链路状态算法，是一种链路状态型路由协议，主要用在自治系统内部进行路由决策。由于采用链路状态类型，所以即使网络中有环路，也能够进行稳定的路由控制。

OSPF 支持子网掩码，为了减少网络流量，还引入了「区域」这一概念，将一个自治网络划分为若干个更小的范围。

在 OSPF 中，路由器之间交换链路状态生成网络拓扑信息，然后再根据这个拓扑信息生成路由控制表。

RIP 的路由选择，要求途中所经过的路由器个数越少越好，但路由器个数最少不一定是最优路径，与之相比，OSPF 可以给每条链路赋予一个权重，并始终选择一个权重最小的路径作为最终路由。

BGP

BGP（Border Gateway Protocol，边界网关协议）是连接不同自治系统的协议，属于外部网关协议。主要用于 ISP（网络服务提供商） 之间相连接的部分。

只有 BGP、RIP 和 OSPF 共同进行路由控制，才能进行整个互联网的路由控制。

在 RIP 和 OSPF 中根据 IP 地址的网络地址部分进行路由控制，然而 BGP 则需要放眼整个互联网进行路由控制。BGP 的路由控制表最终由网络地址和下一站的路由器组来表示，不过它会根据所要经过的 AS 个数进行路由控制。

ISP、区域网络会将每个网络区域编配成一个个自治系统进行管理，它们为每个自治系统分配一个 16 比特的 AS 编号。BGP 就是根据这个编号进行相应的路由控制。

BGP 协议使用的算法是路径向量算法，它是距离向量算法的升级版，解决了坏消息传递慢（某个路由器故障系统感知慢）的问题，将单个自治系统看作一个整体，不区分自治系统内部的路由器，因此规模有限。