

假设某网络操作员具有在每个入口点进入和发向每个出口点的进入网络的流量估计。该操作员接下来可能要设置特定入口到出口的流路由选择，以最小化经所有网络链路的最大利用率。但使用如 OSPF 这样的路由选择算法，操作员调节通过网络的流的路由选择，主要手段是链路权重。因此，为了取得最小化最大链路利用率的目标，该操作员必须找出取得该目标的链路权重集合。这是一种相反的因果关系，即所希望的流路由选择已知，必须找到 OSPF 链路权重，使得该 OSPF 路由选择算法导致这种希望的流路由选择。

4.6.3 自治系统间的路由选择：BGP

我们刚才学习了 ISP 如何使用 RIP 和 OSPF 来决定位于相同 AS 内部的源和目的对之间的优化路径。我们现在研究一下，跨越多个 AS 的源和目的对之间是如何确定路径的。由 RFC 4271（亦可参见 [RFC 4274]）定义的边界网关协议（Broder Gateway Protocol, BGP）版本 4 是当今因特网中域间路由选择协议事实上的标准。它通常被称为 BGP4 或简称为 BGP。作为一个自治系统间路由选择协议（参见 4.5.3 节），BGP 为每个 AS 提供了进行以下工作的手段：

- 1) 从相邻 AS 处获得子网可达性信息。
- 2) 向本 AS 内部的所有路由器传播这些可达性信息。
- 3) 基于可达性信息和 AS 策略，决定到达子网的“好”路由。

更为重要的是，BGP 使得每个子网向因特网的其余部分通告它的存在。一个子网高声宣布“我存在，我在这里”，并且 BGP 确保在因特网中的所有 AS 知道该子网以及如何到达那里。如果没有 BGP 的话，每个子网将是隔离的，即它们孤独并且不为因特网其余部分所知。

1. BGP 基础

BGP 极其复杂，许多专著致力于研究该主题，而且许多问题仍没有得到很好理解 [Yannuzzi 2005]。此外，作为较高层 ISP 的设计者或管理员，即使在阅读了这些专著和这些 RFC 以后，如果不花数月（即使不是数年）实际实践 BGP 的话，可能发现难以全面掌握 BGP。无论如何，因为 BGP 是因特网中绝对至关重要的协议，即从本质上讲，正是这个协议将所有的东西粘合在一起了，因此我们至少需要获得该协议工作原理的初步知识。我们从描述 BGP 在简单的例子网络环境下如何工作开始，我们曾在图 4-32 中学习过该例子网络。在下面的描述中，我们的讨论将基于 4.5.3 节中的层次路由选择，希望读者能复习一下该材料。

在 BGP 中，路由器对通过使用 179 端口的半永久 TCP 连接来交换路由选择信息。对于图 4-32 中的网络而言，该半永久 TCP 显示在图 4-40 中。对于每条直接连接位于两个不同的 AS 中的路由器的链路而言，通常有一条这样的 BGP TCP 连接。因此，在图 4-40 中，在网关路由器 3a 和 1c 之间有一条 TCP 连接，在网关路由器 1b 和 2a 之间有另一条 TCP 连接。在一个 AS 中的路由器之间还有许多半永久 BGP TCP 连接。特别是，图 4-40 显示了一个 AS 内部的每对路由器之间的一条 TCP 连接的通常配置，在每个 AS 内部产生了网状的 TCP 连接。对于每条 TCP 连接，位于该连接端点的两台路由器称为 BGP 对等方（BGP peers），沿着该连接发送所有 BGP 报文的 TCP 连接称为 BGP 会话（BGP session）。此外，跨越两个 AS 的 BGP 会话称为外部 BGP（eBGP）会话（external BGP session），在同一个