

### 3. BGP 路由选择

如本节前面所述, BGP 使用 eBGP 和 iBGP 向在 AS 中的所有路由器发布路由。根据这种发布, 路由器可能知道到达任何一条前缀的多条路由, 在这种情况下路由器必须在可能的路由中选择一条。进入这个路由选择进程的输入是被路由器知道并接受的所有路由的集合。如果对相同前缀存在两条或多条路由, 则 BGP 顺序地调用下列消除规则, 直到留下一条路由。

- 路由被指派一个本地偏好值作为它们的属性之一。一条路由的本地偏好可能由该路由器设置或可能由在相同 AS 中的另一台路由器学习到。这是一条由 AS 的网络管理员决定的决策。(我们随后将更为详细地讨论 BGP 策略问题。) 具有最高本地偏好值的路由将被选择。
- 在余下的路由中(所有都具有相同的本地偏好值), 具有最短 AS-PATH 的路由将被选择。如果该规则是路由选择的唯一规则的话, 则 BGP 将使用一种距离向量算法来决定路径, 其中距离测度使用 AS 跳的数目而不是路由器跳的数目。
- 在余下的路由中(所有都具有相同的本地偏好和相同的 AS-PATH 长度), 将选择具有最靠近 NEXT-HOP 路由器的路由。这里, 最靠近是指具有最低费用路径的费用的路由器, 它由 AS 内部算法来决定。如在 4.5.3 节中所讨论的那样, 该进程经常被称为热土豆路由选择。
- 如果仍留下多条路由, 该路由器使用 BGP 标识符来选择路由; 参见 [Stewart 1999]。

消除规则甚至比上面描述的更为复杂。为了避免 BGP 出现严重问题, 少量地学习一些 BGP 选择规则方面的知识是有帮助的。

### 实践原则

#### 将它们拼装起来: 表项是怎样进入路由器的转发表的?

前面讲过路由器转发表中的一个表项是由前缀(如 138.16.64/22)和一个对应的路由器输出端口(如端口 7)组成的。当一个分组到达路由器时, 该分组的目的 IP 地址与转发表中的前缀进行比较, 找到与最长前缀匹配的表项。该分组则(在本路由器中)转发到与该前缀相关的端口。我们现在总结一下一个路由选择表项(前缀和相关端口)是如何进入转发表中的。这种简单的操作将我们已经学习的路由选择和转发的知识有机地结合了起来。为了使其有趣, 我们假设该前缀是一种“外部前缀”, 即它不属于该路由器的 AS 而属于某些其他 AS。

为使一个前缀进入路由器的转发表, 路由器必须首先知晓该前缀(对应于一个子网或子网的一个聚合)。如我们已经学习的那样, 经过 BGP 路由通告, 该路由器知道了前缀。这种通告可能(从位于另一个 AS 中的一台路由器)经一个 eBGP 会话, 或者(从位于相同 AS 中的一台路由器)经一个 iBGP 会话发送给它。

在路由器知道了前缀之后, 在它能够在其转发表中加入该前缀之前, 需要决定适当的输出端口——目的地指向该前缀的数据报将转发到该端口。如果路由器接收到对该前缀的多个路由通告, 该路由器使用 BGP 路由选择进程, 如本小节前面描述的那样, 找到对于该前缀的“最好的”路由。假定这样一条最好的路由已选定。如前所述, 所选择