

# 路由算法

jask

09/16/2024

## OSPF 开放式最短路径优先

这是一个基于链路状态的内部网关协议（IGP），广泛应用于大型网络中。OSPF 使用 Dijkstra 算法来计算网络的最短路径，并且能够快速收敛，适用于各种规模的网络。

OSPF 是一种基于 IP 协议的路由协议。它是大中型网络上使用较为广泛的 IGP 协议。OSPF 是对链路状态路由协议的一种实现，运作于自治系统内部。OSPF 分为 OSPFv2 和 OSPFv3 两个版本：OSPFv2 定义于 RFC 2328 (1998)，支持 IPv4 网络；而 OSPFv3 定义于 RFC 5340 (2008)，支持 IPv6 网络。

它采用戴克斯特拉算法来计算最短路径树。它使用“开销 (Cost)”作为路由度量。链路状态数据库 (LSDB) 用来保存当前网络拓扑结构，路由器上属于同一区域的链路状态数据库是相同的（属于多个区域的路由器会为每个区域维护一份链路状态数据库）。

OSPF 提出了“区域 (Area)”的概念，一个网络可以由单一区域或者多个区域组成。其中，一个特别的区域被称为骨干区域 (Backbone Area)，该区域是整个 OSPF 网络的核心区域，并且所有其他的区域都与之直接连接。所有的内部路由都通过骨干区域传递到其他非骨干区域。所有的区域都必须直接连接到骨干区域，如果不能建立直接连接，那么可以通过虚链路 (virtual link) 和骨干区域建立虚拟连接。

同一个广播域 (Broadcast Domain) 的路由器或者一个点对点 (Point To Point) 连接的两端的路由器，在发现彼此的时候，建立邻接 (Adjacencies) 多路访问网络以及非广播多路访问网络的路由器会选举指定路由器 (Designated Router, DR) 和备份指定路由器 (Backup Designated Router, BDR)，DR 和 BDR 作为网络的中心负责路由器之间的信息交换从而降低了网络中的信息流量。OSPF 协议同时使用单播 (Unicast) 和群播 (Multicast) 来发送 Hello 包和链路状态更新 (Link State Updates)，使用的群播地址为 224.0.0.5 和 224.0.0.6。与 RIP 和 BGP 不同的是，OSPF 协议不使用 TCP 或者 UDP 协议而是直接承载在 IP 协议之上，IP 协议号为 89。

区域是以接口 (Interface) 为单位来划分的，所以一台多接口路由器可能属于多个区域。相同区域内的所有路由器都维护一份相同的链路状态数据库 (LSDB)，如果一台路由器属于多个区域，那么它将为每一个区域维护一份 LSDB。将一个网络划分为多个区域有以下优点：

某一区域内的路由器只需要维护该区域的链路状态数据库，而不用维护整个 OSPF 网络的链路状态数据库。

将某一区域网络拓扑变化的影响限制在该区域内，不会影响到整个 OSPF 网络，从而减小 OSPF 计算的频率。

将链路状态通告 (LSA) 的洪泛限制在本区域内，从而降低 OSPF 协议产生的数据量。

划分区域可以对网络进行层次化结构设计。

划分区域有利于资源合理调配，核心区域部署性能较好的设备资源，边缘区域部署性能较差的设备资源即可。

OSPF 定义了以下 4 种网络类型：[1]

点到点网络 (point-to-point)

广播网络 (broadcast)

非广播多路访问网络 (non-broadcast multi-access, NBMA)

点到多点网络 (Point-to-MultiPoint)

点到点网络

点到点网络，例如 E1、SONET，是单独连接一对路由器的网络。点到点网络上的一对 OSPF 路由器形成完全邻接关系 (Full Adjacency)，并且不进行 DR 和 BDR 的选举。点到点网络上的路由器使用群播地址 224.0.0.5 发送 OSPF 协议数据包。广播网络

广播网络即可以同时连接多于两台设备的网络，如以太网、令牌环网、FDDI。广播网络上的路由器发送的群播/广播数据包会被其他与之相连的路由器收到。在广播网络上的 OSPF 路由器会选举一台指定路由器 (DR) 和一台备份指定路由器 (BDR)。所有始发于 DR 和 BDR 的 OSPF 数据包使用目的地址 224.0.0.5，以群播方式发送到所有其他 OSPF 路由器，所有其他的路由器都将使用目的地址 224.0.0.6，以群播方式发送 OSPF 数据包到 DR 和 BDR。所有其他的路由器只与 DR 和 BDR 建立完全邻接关系。非广播多路访问网络

NBMA 网络，诸如 X.25、帧中继、ATM 等，可以同时连接两台以上的路由器，但是这种网络没有广播数据包的能力。一台处于 NBMA 网络上的路由器发送的群播/广播数据包将不能被其他与之相连的路由器收到。在 NBMA 网络上需要选举 DR 和 BDR，并且所有的 OSPF 数据包都是单播发送的。点到多点网络点到多点网络是 NBMA 网络的一个特殊设置，可以看作是一群点到点链路的集合，因此在该种网络上不必选举 DR 和 BDR。点到多点网络上 OSPF 的行为和点到点网络 OSPF 的行为一样，也使用群播地址 224.0.0.5 发送 OSPF 协议数据包。

在 OSPF 路由器之间互相交换信息之前，必须先建立邻接关系。

两台 OSPF 路由器要建立完全邻接关系，以下参数必须相同：

Hello时间间隔  
Dead时间间隔  
区域编号  
认证（如果启用了认证）  
链路MTU大小  
子网掩码  
子网号  
末梢区域设置

一般来说，建立 OSPF 完全邻接时会经过以下状态：

失效状态（Down）：这是邻居会话的初始状态，表示最近没有从邻居收到信息。在NBMA网络上，可能仍然会以较低频率向处于Down状态的邻居发送Hello数据包。  
尝试状态（Attempt）：该状态仅仅适用于连接在NBMA网络上的邻居。该状态表示最近没有从邻居收到信息，但仍需要作进一步的尝试。  
初始状态（Init）：在此状态下，表示最近收到了从邻居发来的Hello数据包。但是，仍然没有和邻居建立双向通信（Bidirectional Communication）。  
双向通信状态（2-Way）：此状态意味着两台路由器之间建立了双向通信。在此状态下还将进行DR和BDR的选举（只有处于2-Way状态的邻居才有资格成为DR或BDR）。  
信息交换初始状态（ExStart）：这个状态是建立邻接关系的第一步。该状态的目标是决定信息交换时路由器的主从关系，并确定要交换的数据包的范围。  
信息交换状态（ExChange）：在此状态的路由器通过向邻居发送DD数据包来描述其完整的链路状态数据库。每一个DD数据包都有序列号和长度。  
信息加载状态（Loading）：在此状态下，路由器将会向邻居路由器发送链路状态请求数据包，用来请求信息交换状态发现的最新链路状态信息。  
完全邻接状态（Full）：在此状态下，邻居路由器形成完全邻接关系。这些邻接关系将会在路由器LSA和网络LSA中被描述。

**Hello 包** Hello 包的 OSPF 包类型为 1。这些包被周期性的从各个接口（包括虚链路）发出，用来建立和维护邻居关系。另外，在支持群播或广播的物理网络上，Hello 包使用群播地址（通常为 224.0.0.5）发送。Hello 包的发送间隔由 HelloInterval 指定（通常为 10s），路由器将会先发出不包含邻居字段的空 Hello 包，当收到邻接路由器的 Hello 包之后，将对方的路由器标识放入本机的 Hello 包中进行群播，这种包含邻居字段的包，也被称为 HelloSeen 包。假如在间隔达到 RouterDeadInterval 所规定的时长（通常为 40s）内仍未收到一个已建立连接的路由器的 Hello 包，路由器将会终止这一连接。将对方的状态转为 Down。