

组播扩展OSPF

编辑^[1] 锁定^[2] 同义词^[3] 一般指组播扩展OSPF

本词条由“科普中国”百科科学词条编写与应用工作项目^[4] 审核。

OSPF(Open Shortest Path First开放式最短路径优先^[5]) 是一个内部网关协议^[6](Interior Gateway Protocol，简称IGP)，用于在单一自治系统^[7]（autonomous system,AS）内决策路由^[8]。是对链路状态路由协议^[9]的一种实现，隶属内部网关协议（IGP），故运作于自治系统内部。著名的迪克斯加算法(Dijkstra)被用来计算最短路径树。OSPF分为OSPFv2和OSPFv3两个版本,其中OSPFv2用在IPv4^[10]网络，OSPFv3用在IPv6^[11]网络。OSPFv2是由RFC 2328定义的，OSPFv3是由RFC 5340定义的。与RIP^[12]相比，OSPF是链路状态协议，而RIP是距离矢量协议^[13]。

不同厂商管理距离不同，思科OSPF的协议管理距离^[14]（AD）是110，华为OSPF的协议管理距离是10。。。。[1]

中文名
开放最短路径优先
外文名
OSPF（Open Shortest Path First）
类 型
路由器选择协议

组播扩展 OSPFOSPF起源

编辑^[15]

IETF^[16]为了满足建造越来越大基于IP网络的需要，形成了一个工作组，专门用于开发开放式的链路状态路由协议^[17]，以便用在大型、异构的IP网络中。新的路由协议^[18]已经取得一些成功的一系列私人的、和生产商相关的、最短路径^[19]优先（SPF）路由协议为基础，在市场上广泛使用。包括OSPF在内，所有的S P F路由协议基于一个数学算法—Dijkstra算法^[20]。这个算法能使路由选择^[21]基于链路状态，而不是距离向量^[22]。OSPF由IETF在20世纪80年代末期开发，OSPF是SPF类路由协议中的开放式版本。最初的OSPF规范体如今RFC1131中。这个第1版（OSPF版本1）很快被进行了重大改进的版本所代替，这个新版本体如今RFC1247文档中。RFC 1247OSPF称为OSPF版本2是为了明确指出其在稳定性和功能性方面的实质性改进。这个OSPF版本有许多更新文档，每一个更新都是对开放标准的精心改进。接下来的一些规范出如今RFC 1583、2178和2328中。OSPF版本2的最新版体如今RFC 2328中。最新版只会和由RFC 2138、1583和1247所规范的本体进行互操作。

链路是路由器^[23]接口的另一种说法，因此OSPF也称为接口状态路由协议^[24]。OSPF通过路由器之间通告网络接口的状态来建立链路状态数据库，生成最短路径^[25]树，每个OSPF路由器使用这些最短路径构造路由表^[26]。

OSPF路由协议是一种典型的链路状态（Link-state）的路由协议，一般用于同一个路由域内。在这里，路由域^[27]是指一个自治系统^[28]（**Autonomous System**），即AS，它是指一组通过统一的路由政策或路由协议互相交换路由信息的网络。在这个AS中，所有的OSPF路由器^[29]都维护一个相同的描述这个AS结构的数据库，该数据库中存放的是路由域中相应链路的状态信息，OSPF路由器正是通过这个数据库计算出其OSPF路由表的。

作为一种链路状态的路由协议^[30]，OSPF将链路状态组播数据LSA（Link State Advertisement）传送给在某一区域内的所有路由器，这一点与距离矢量路由协议^[31]不同。运行距离矢量路由协议的路由器是将部分或全部的路由表^[32]传递给与其相邻的路由器。

组播扩展 OSPF术语与协议

编辑^[33]

组播扩展 OSPFOSPF术语

Router-ID

假设这个世界上的人名字是没有重复的，每个人的名字都不相同，当有一天，遇上个陌生人告诉你，有任何麻烦可以找他，他一定能够帮你解决；等到你有麻烦的时候，你想找那个人帮忙，可是如果你连那个人的名字都不知道，那么也就不可能找到那个人帮忙了。OSPF就类似于上述情况，网络中每台OSPF路由器^[34]都相当于一个人，OSPF路由器之间相互通告链路状态，就等于是告诉别人可以帮别人的忙，如此一来，如果路由器之间分不清谁是谁，没有办法确定各自的身份，那么通告的链路状态就是毫无意义的，所以必须给每一个OSPF路由器定义一个身份，就相当于人的名字，这就是Router-ID，并且Router-ID在网络中绝对不可以有重名，否则路由器收到的链路状态，就无法确定发起者的身份，也就无法通过链路状态信息确定网络位置，OSPF路由器发出的链路状态都会写上自己的Router-ID，可以理解为该链路状态的签名，不同路由器产生的链路状态，签名绝不会相同。

每一台OSPF路由器只有一个Router-ID，Router-ID使用IP地址的形式来表示，确定Router-ID的方法为：

★1.手工指定Router-ID。

★2.路由器上活动Loopback接口中IP地址最大的，也就是数字最大的，如C类地址优先于B类地址，一个非活动的接口的IP地址是不能被选为Router-ID的。

★3.如果没有活动的Loopback接口，则选择活动物理接口IP地址最大的。

注：如果一台路由器收到一条链路状态，无法到达该Router-ID的位置，就无法到达链路状态中的目标网络。

Router-ID只在OSPF启动时计算，或者重置OSPF进程后计算。

COST

OSPF使用接口的带宽来计算Metric^[35]，例如一个10 Mbit/s的接口，计算Cost的方法为：

将10 Mbit换算成bit，为10 000 000 bit，然后用10000 0000除以该带宽，结果为 $10000\ 0000/10\ 000\ 000\ \text{bit} = 10$ ，所以一个10 Mbit/s的接口，OSPF认为该接口的Metric值为10，需要注意的是，计算中，带宽的单位取bit/s，而不是Kbit/s，例如一个100 Mbit/s的接口，Cost 值为 $10000\ 0000/100\ 000\ 000=1$ ，因为Cost值必须为整数，所以即使是一个1000 Mbit/s（1Gbit/s）的接口，Cost值和100Mbit/s一样，为1。如果路由器要经过两个接口才能到达目标网络，那么很显然，两个接口的Cost值要累加起来，才算是到达目标网络的Metric值，所以OSPF路由器计算到达目标网络的Metric值，必须将沿途中所有接口的Cost值累加起来，在累加时，同EIGRP一样，只计算出接口，不计算进接口。

OSPF会自动计算接口上的Cost值，但也可以通过手工指定该接口的Cost值，手工指定的优先于自动计算的值。

OSPF计算的Cost，同样是和接口带宽成反比，带宽越高，Cost值越小。到达目标相同Cost值的路径，可以执行负载均衡，最多6条链路同时执行负载均衡。

链路（Link）

就是路由器上的接口，在这里，应该指运行在OSPF进程下的接口。

链路状态（Link-State）

链路状态（LSA）就是OSPF接口上的描述信息，例如接口上的IP地址，子网掩码^[36]，网络类型，Cost值等等，OSPF路由器之间交换的并不是路由表，而是链路状态（LSA），OSPF通过获得网络中所有的链路状态信息，从而计算出到达每个目标精确的网络路径。OSPF路由器会将自己所有的链路状态毫不保留地全部发给邻居，邻居将收到的链路状态全部放入链路状态数据库（Link-State Database），邻居再发给自己的所有邻居，并且在传递过程中，绝对不会有任意更改。通过这样的过程，最终，网络中所有的OSPF路由器都拥有网络中所有的链路状态，并且所有路由器的链路状态应该能描绘出相同的网络拓扑。比如如今要计算一条地铁线路图，如上海地铁二号线某段的图，如果不直接将该图给别人看，图好比是路由表，如今只是报给别人各个站的信息，该信息好比是链路状态，通过告诉别人各个站左边一站是什么，右边一站是什么，别人也能通过该信息（链路状态），画出完整的线路图（路由表），如得到如下各站信息（链路状态）：

★南京东路-站（左边一站是人民广场，右边一站是陆家嘴）

★南京西路-站（左边一站是静安寺，右边一站是人民广场）

★静安寺-站（右边一站是南京西路）

★人民广场-站（左边一站是南京西路，右边一站是南京东路）

★陆家嘴-站 （左边一站是南京东路）

还原线路图（路由表）如下：

根据分析以下两站信息（两条链路状态）：

★南京西路-站 （左边一站是静安寺，右边一站是人民广场）

★静安寺-站 （右边一站是南京西路）

计算 因为静安寺右边是南京西路，而南京西路左边是静安寺^[37]，所以静安寺和南京西路是相邻的，为 静安寺 — 南京西路，并且由于南京西路右边是人民广场，所以通过这两条信息，得出线路为 静安寺 — 南京西路 — 人民广场，继续往下

再根据如下两站信息（链路状态）：

★人民广场-站 （左边一站是南京西路，右边一站是南京东路）

★南京东路-站 （左边一站是人民广场，右边一站是陆家嘴）

计算 因为之前南京西路右边是人民广场，人民广场左边是南京西路，所以南京西路和人民广场是相邻的两站，并且人民广场右边是南京东路，得出线路为 南京西路 — 人民广场 — 南京东路，并且因为南京东路右边是陆家嘴，所以 这部分线路得知为南京西路 — 人民广场 — 南京东路 — 陆家嘴，继续往下

再根据如下一站信息（链路状态）：

★陆家嘴-站 （左边一站是南京东路）

计算 因为南京东路右边是陆家嘴，而陆家嘴左边是南京东路，所以两站相邻，得出为 南京东路 — 陆家嘴，

通过以上各部分的线路：

静安寺 — 南京西路 — 人民广场

南京西路 — 人民广场 — 南京东路 — 陆家嘴

南京东路 — 陆家嘴

所以很轻松的就画出该段地铁线路图为：

静安寺 — 南京西路 — 人民广场 — 南京东路 — 陆家嘴

从以上计算过程可以知道，因为得到各站的信息，就能画出整条线路图，而OSPF也同样根据路由器各接口的信息（链路状态），计算出网络拓扑图，OSPF之间交换链路状态，就像上面交换各站信息，而不像RIP和EIGRP直接交换路由表，交换路由表，就等于直接给人看线路图，可见OSPF的智能算法，比距离矢量协议对网络有更精确的认知。

OSPF区域

因为OSPF路由器之间会将所有的链路状态（LSA）相互交换，毫不保留，当网络规模达到一定程度时，LSA将形成一个庞大的数据库，势必会给OSPF计算带来巨大的压力；为了能够降低OSPF计算的复杂程度，缓存计算压力，OSPF采用分区域计算，将网络中所有OSPF路由器划分成不同的区域，每个区域负责各自区域精确的LSA传递与路由计算，然后再将一个区域的LSA简化和汇总之后转发到另外一个区域，这样一来，在区域内部，拥有网络精确的LSA，而在不同区域，则传递简化的LSA。区域的划分为了能够尽量设计成无环网络，所以采用了Hub-Spoke的拓扑架构，也就是采用核心与分支的拓扑^[38]，如下图：

[39]

区域的命名可以采用整数数字，如1、2、3、4，也可以采用IP地址的形式，0.0.0.1、0.0.0.2，因为采用了Hub-Spoke的架构，所以必须定义出一个核心，然后其它部分都与核心相连，OSPF的区域0就是所有区域的核心，称为BackBone 区域（骨干区域），而其它区域称为Normal 区域（常规区域），在理论上，所有的常规区域应该直接和骨干区域相连，常规区域只能和骨干区域交换LSA，常规区域与常规区域之间即使直连也无法互换LSA，如上图中Area 1、Area 2、Area 3、Area 4只能和Area 0互换LSA，然后再由Area 0转发，Area 0就像是一个中转站，两个常规区域需要交换LSA，只能先交给Area 0，再由Area 0转发，而常规区域之间无法互相转发。

OSPF区域是基于路由器的接口划分的，而不是基于整台路由器划分的，一台路由器可以属于单个区域，也可以属于多个区域，如下图：
[40]

如果一台OSPF路由器属于单个区域，即该路由器所有接口都属于同一个区域，那么这台路由器称为Internal Router（IR），如上图中的R2，R3和R4；如果一台OSPF路由器属于多个区域，即该路由器的接口不都属于一个区域，那么这台路由器称为Area Border Router（ABR），如上图中的R1，ABR可以将一个区域的LSA汇总后转发至另一个区域；如果一台OSPF路由器将外部路由协议重分布进OSPF，那么这台路由器称为Autonomous System Boundary Router（ASBR），如上图，R5将EIGRP重分布进OSPF，那么R5就是ASBR，但是如果只是将OSPF重分布进其它路由协议，则不能称为ASBR。

可以配置任何OSPF路由器成为ABR或ASBR。

由于OSPF有着多种区域，所以OSPF的路由在路由表中也以多种形式存在，共分以下几种：

如果是同区域的路由，叫做Intra-Area Route，在路由表中使用O来表示；

如果是不同区域的路由，叫做Inter-Area Route或Summary Route，在路由表中使用O IA来表示；

如果并非OSPF的路由，或者是不同OSPF进程的路由，只是被重分布到OSPF的，叫做External Route，在路由表中使用O E2或OE 1来表示。

当存在多种路由可以到达同一目的地时，OSPF将根据先后顺序来选择要使用的路由，所有路由的先后顺序为：

Intra-Area — Inter-Area — External E1 — External E2，即 O — O IA — O E1 — O E2。

注：

★一台路由器可以运行多个OSPF进程，不同进程的OSPF，可视为没有任何关系，如需要获得相互的路由信息，需要重分布。

★每个OSPF进程可以有多个区域，而路由器的链路状态数据库是分进程和分区域存放的。

邻居（Neighbor）

OSPF只有邻接状态才会交换LSA，路由器会将链路状态数据库中所有的内容毫不保留地发给所有邻居，要想在OSPF路由器之间交换LSA，必须先形成OSPF邻居，OSPF邻居靠发送Hello包来建立和维护，Hello包会在启动了OSPF的接口上周期性发送，在不同的网络中，发送Hello包的间隔也会不同，当超过4倍的Hello时间，也就是Dead时间过后还没有收到邻居的Hello包，邻居关系将被断开。

两台OSPF路由器必须满足4个条件，才能形成OSPF邻居，4个必备条件如下：

Area-id（区域号码）

即路由器之间必须配置在相同的OSPF区域，否则无法形成邻居。

Hello and Dead Interval（Hello时间与Dead时间）

即路由器之间的Hello时间和Dead时间必须一致，否则无法形成邻居。

Authentication（认证）

路由器之间必须配置相同的认证密码，如果密码不同，则无法形成邻居。

Stub Area Flag（末节标签）

路由器之间的末节标签必须一致，即处在相同的末节区域内，否则无法形成邻居。

注：

★OSPF只能使用接口的Primary地址建立邻居，不能使用Secondary建立邻居。

★路由器双方接口要么都为手工配置地址（Numbered），要么都为借用地址（Unnumbered），否则无法建立邻居。

邻接（Adjacency）

两台OSPF路由器能够形成邻居，但并不一定能相互交换LSA，只要能交换LSA，关系则称为邻接（Adjacency）。邻居之间只交换Hello包，而邻接（Adjacency）之间不仅交换Hello包，还要交换LSA。

DR/BDR

当多台OSPF路由器连到同一个多路访问网段时，如果每两台路由器之间都相互交换LSA，那么该网段将充满着众多LSA条目，为了能够尽量减少LSA的传播数量，通过在多路访问网段中选择出一个核心路由器，称为DR（Designated Router），网段中所有的OSPF路由器都和DR互换LSA，这样一来，DR就会拥有所有的LSA，并且将所有的LSA转发给每一台路由器；DR就像是该网段的LSA中转站，所有的路由器都与该中转站互换LSA，如果DR失效后，那么就会造成LSA的丢失与不完整，所以在多路访问网络中除了选举出DR之外，还会选举出一台路由器作为DR的备份，称为BDR（Backup Designated Router），BDR在DR不可用时，代替DR的工作，而既不是DR，也不是BDR的路由器称为Drother，事实上，Drother除了和DR互换LSA之外，同时还会和BDR互换LSA。

其实不难看出，DR与BDR并没有任何本质与功能的区别，只有在多路访问的网络环境，才需要DR和BDR，DR与BDR的选举是在一个二层网段内选举的，即在多个路由器互连的接口范围内，与OSPF区域没有任何关系，一个区域可能有多个多路访问网段，那么就会存在多个DR和BDR，但一个多路访问网段，只能有一个DR和BDR；选举DR和BDR的规则为：

★比较接口优先级

选举优先级最高的成为DR，优先级数字越大，表示优先级越高，被选为DR的几率就越大，次优先级的为BDR，优先级范围是0-255，默认为1，优先级为0表示没有资格选举DR和BDR。

★Route-Id大小

如果在优先级都相同的情况下，Route-Id 最大的成为DR，其次是BDR，数字越大，被选为DR的几率就越大。

因为所有路由器都能与DR和BDR互换LSA，所以所有路由器都与DR和BDR是邻接（Adjacency）关系，而Drother与Drother之间无法互换LSA，所以Drother与Drother之间只是邻居关系。

在一个多路访问网络中，选举DR和BDR是有时间限制的，该时间为Wait时间，默认为4倍的Hello时间，即与Dead时间相同，如果OSPF路由器在超过Wait时间后也没有其它路由器与自己竞争DR与BDR的选举，那么就选自己为DR；当一个多路访问网络中选举出DR与BDR之后，在DR与BDR没有失效的情况下，不会进行重新选举，也就是在选举出DR与BDR之后，即使有更高优先级的路由器加入网络，也不会影响DR与BDR的角色，在越出选举时间（Wait时间）后，只有DR与BDR失效后，才会重新选举。DR失效后，会同时重新选举DR与BDR，而在BDR失效后，只会重新选举BDR。

DR和BDR与Drother的数据包处理会有所不同，

所有OSPF路由器，包括DR与BDR，都能够接收和传递目标地址为224.0.0.5的数据包。

只有DR和BDR才能接收和传递目标地址为224.0.0.6的数据包。

由此可见，Drother路由器将数据包发向目标地址224.0.0.6，只能被DR和BDR接收，其它Drother不能接收；而DR和BDR将数据包发向目标地址224.0.0.5，可以被所有路由器接收。

组播扩展OSPF协议

一、Hello协议的目的：

用于发现邻居

在成为邻居之前，必须对Hello包里的一些参数进行协商

Hello包在邻居之间扮演着keepalive的角色

允许邻居之间的双向通信

用于在NBMA(Nonbroadcast Multi-access)、广播网络(以太网)中选举DR和BDR

二、Hello Packet包含以下信息：

源路由器^[41]的RID

源路由器的Area ID

源路由器接口的掩码^[42]

源路由器接口的认证类型和认证信息

源路由器接口的Hello包发送的时间间隔

源路由器接口的无效时间间隔

DR/BDR接口IP地址

五个标记位（flag bit）

源路由器的所有邻居的RID

组播扩展 OSPF网络类型

编辑^[43]

OSPF定义的5种网络类型：

- 1. 点到点网络^[44](point-to-point) ， 由cisco^[45]提出的网络类型， 自动发现邻居， 不选举DR/BDR， hello时间10s。
 - 1.1点到点网络,比如T1线路， 是连接单独的一对路由器^[46]的网络， 点到点网络上的有效邻居总是可以形成邻接关系的， 在这种网络上， OSPF包的目标地址使用的是**224.0.0.5**， 这个组播地址^[47]称为**AllSPFRouters**.
- 2.广播型网络^[48](broadcast) ， 由cisco提出的网络类型， 自动发现邻居， 选举DR/BDR， hello时间10s。
 - 2.1广播型网络， 比如以太网， *Token Ring*和*FDDI*， 这样的网络上会选举一个DR和BDR,DR/BDR的发送的OSPF包的目标地址为224.0.0.5， 运载这些OSPF包的帧的目标MAC地址为0100.5E00.0005； 而除了DR/BDR以外发送的OSPF包的目标地址为**224.0.0.6**， 这个地址叫**AllDRouters**.
- 3.非广播型（NBMA^[49]）网络 (non-broadcast) ， 由RFC提出的网络类型， 手工配置邻居， 选举DR/BDR， hello时间30s。
 - 3.1.NBMA 网络， 比如X.25,Frame Relay， 和ATM， 不具备广播的能力， 因此邻居要人工来指定， 在这样的网络上要选举DR和BDR,OSPF包采用unicast的方式
- 4.点到多点网络 (point-to-multipoint) ， 由RFC^[50]提出， 自动发现邻居， 不选举DR/BDR， hello时间30s。
 - 4.1点到多点网络 是NBMA网络的一个特殊配置， 可以看成是点到点链路的集合. 在这样的网络上不选举DR和BDR.
- 5.点到多点非广播，由cisco提出的网络类型， 手动配置邻居， 不选举DR/BDR， hello时间30s。
- 6.虚链接： OSPF包是以unicast的方式发送

所有的网络也可以归纳成2种网络类型：

- 1. 传输网络（Transit Network）
- 2. 末梢网络（Stub Network）

OSPF模式

OSPF模式	NBMA 推荐拓扑	Hello 时间	RFC/Cisco
Broadcast	Full or partial mesh	10 sec	Automatic, DR/BDR elected Cisco
Nonbroadcast (NBMA)	Full or partial mesh	30 sec	Manual configuration, DR/BDR elected

Point-to-multipoint	Partial-mesh		30 Sec	Automatic,	
	or star			no DR/BDR	
Point-to-multipoint nonbroadcast	partial-mesh		30 sec	Manual configuration, no/DR/BDR	Cisco
	or star				
Point-to-point	Partial-mesh or star, using subinterface	Different for Each Subinterface	10 sec	Automatic, no DR/BDR	Cisco

组播扩展 OSPFDR、BDR

编辑^[51]

在DR和BDR出现之前，每一台路由器^[52]和他的所有邻居成为完全网状的OSPF邻接关系，这样5台路由器之间将需要形成10个邻接关系，同时将产生20条LSA.而且在多址网络中，还存在自己发出的LSA 从邻居的邻居发回来，导致网络上产生很多LSA的拷贝，所以基于这种考虑，产生了DR和BDR.

DR将完成如下工作

- 1. 描述这个多址网络 and 该网络上剩下的其他相关路由器^[53].
- 2. 管理这个多址网络上的flooding过程.
- 3. 同时为了冗余^[54]性，还会选取一个BDR，作为双备份之用.

DR BDR选取规则：DR BDR选取是以接口状态机^[55]的方式触发的.

- 1. 路由器^[56]的每个多路访问（multi-access）接口都有个路由器优先级（Router Priority),8位长的一个整数，范围是0到255,Cisco路由器默认的优先级是1优先级为0的话将不能选举为DR/BDR.优先级可以通过命令ip ospf priority进行修改.
- 2. Hello包里包含了优先级的字段，还包括了可能成为DR/BDR的相关接口的IP地址.
- 3. 当接口在多路访问网络上初次启动的时候，它把DR/BDR地址设置为0.0.0.0，同时设置等待计时器（wait timer）的值等于路由器^[57]无效间隔（Router Dead Interval).

DR BDR选取过程：

- 1. 路由器^[58]X在和邻居建立双向（2-Way）通信之后，检查邻居的Hello包中Priority,DR和BDR字段，列出所有可以参与DR/BDR选举的邻居（priority不为0).
- 2. 如果有一台或多台这样的路由器宣告自己为BDR（也就是说，在其Hello包中将自己列为BDR，而不是DR），选择其中拥有最高路由器优先级的成为BDR；如果相同，选择拥有最大路由器标识的。如果没有路由器宣告自己为BDR，选择列表中路由器拥有最高优先级的成为BDR，（同样排除宣告自己为DR的路由器），如果相同，再根据路由器标识。
- 3. 按如下计算网络上的DR。如果有一台或多台路由器^[59]宣告自己为DR（也就是说，在其Hello包中将自己列为DR），选择其中拥有最高路由器优先级的成为DR；如果相同，选择拥有最大路由器标识的。如果没有路由器宣告自己为DR，将新选举出的BDR设定为DR。
- 4. 如果路由器X新近成为DR或BDR，或者不再成为DR或BDR，重复步骤2和3，然后结束选举。这样做是为了确保路由器不会同时宣告自己为DR和BDR。
- 5. 要注意的是，当网络中已经选举了DR/BDR后，又出现了1台新的优先级更高的路由器^[60]，DR/BDR是不会重新选举的。
- 6. DR/BDR选举完成后，DRother只和DR/BDR形成邻接关系.所有的路由器将组播Hello包到AllSPFRouters地址224.0.0.5以便它们能跟踪其他邻居的信息，即DR将泛洪update packet到224.0.0.5;DRother只组播update packet到AllDRRouter地址224.0.0.6,只有**DR/BDR**监听这个地址。

简洁的说：**DR**的筛选过程

- 1. 优先级为0的不参与选举；
- 2. 优先级高的路由器^[61]为DR；
- 3. 优先级相同时，以router ID 大为DR；

router-ID选举规则

首先会在该路由器上的所有 环回接口里选一个IP地址最大的，如果没有环回接口就选活动的物理接口IP地址最大的。

（只有在IPV4的环境下可以自动选举route-ID 如果是纯IPV6环境必须手工指定）

邻接关系建立的4个阶段：

- 1. 邻居发现阶段。
- 2. 双向通信阶段：**Hello**报文都列出了对方的**RID**，则**BC**完成。
- 3. 数据库同步阶段：主从协商；DD交换；LSA请求；LSA传播；LSA应答。
- 4. 完全邻接阶段：**full adjacency**。

邻居关系的建立和维持都是靠Hello包完成的，在一般的网络类型中，Hello包周期性的以HelloInterval^[62]秒发送，有1个例外：在NBMA网络中，路由器^[63]每经过一个PollInterval周期发送Hello包给状态为down的邻居（其他类型的网络是不会把Hello包发送给状态为down的路由器的）。Cisco路由器上PollInterval默认120s Hello Packet以组播的方式发送给224.0.0.5，在NBMA类型，点到多点和虚链路类型网络，以单播^[64]发送给邻居路由器。邻居可以通过手工配置或者Inverse-ARP发现。

OSPF路由器在完全邻接之前，所经过的几个状态：

- 1. **Down**：此状态还没有与其他路由器^[65]交换信息。首先从其ospf接口向外发送hello分组，还并不知道DR（若为广播网络）和任何其他路由器。发送hello分组使用组播地址^[66]224.0.0.5。
- 2. **Attempt**：只适于NBMA网络，在NBMA网络中邻居是手动指定的，在该状态下，路由器^[67]将使用HelloInterval^[68]取代PollInterval来发送Hello包。
- 3. **Init**：表明在DeadInterval里收到了Hello包，但是2-Way通信仍然没有建立起来。
- 4. **two-way**：双向会话建立，而RID彼此出现在对方的邻居列表中。（若为广播网络：例如：以太网。在这个时候应该选举DR,BDR。）
- 5. **ExStart**：信息交换初始状态，在这个状态下，本地路由器^[69]和邻居将建立Master/Slave关系，并确定DD Sequence Number，路由器ID大的的成为Master。
- 6. **Exchange**：信息交换状态，本地路由器和邻居交换一个或多个DBD分组（也叫DDP），DBD包含有关LSDB中LSA条目的摘要信息。
- 7. **Loading**：信息加载状态：收到DBD后，将收到的信息同LSDB中的信息进行比较。如果DBD中有更新的链路状态条目，则向对方发送一个LSR，用于请求新的LSA。
- 8. **Full**：完全邻接状态，邻接间的链路状态数据库同步完成，通过邻居链路状态请求列表为空且邻居状态为Loading判断。

组播扩展 OSPF无法形成邻接关系的常见原因

编辑^[70]

- 1. Hello间隔和Dead间隔必须相同才能建立邻接关系。
- 2. 区域号码不一致。
- 3. 特殊区域（如stub和nssa等）区域类型不匹配。
- 4. 认证类型和密码不一致。
- 5. 路由器ID相同。
- 6. Hello包被ACLdeny。
- 7. 链路上的MTU不匹配。
- 8. 接口下OSPF网络类型不匹配。

组播扩展 OSPF泛洪与区域

编辑^[71]

组播扩展 OSPFOSPF泛洪

Flooding采用2种报文

LSU Type 4---链路状态更新报文

LSA Type 5---链路状态确认报文

（补充下）

{

Hello Type 1 ---Hello协议报文

DD(Data Description) Type 2----链路数据描述报文

LSR Type 3----链路状态请求报文

}

在**P-P**网络，路由器是以组播方式将更新报文发送到组播地址**224.0.0.5**。

在**P-MP**和虚链路网络，路由器以单播方式将更新报文发送至邻接邻居的接口地址。

在广播型网络，DRother路由器只能和DR&BDR形成邻接关系，所以更新报文将发送到224.0.0.6，相应的DR以224.0.0.5泛洪LSA并且BDR只接收LSA，不会确认和泛洪这些更新，除非DR失效 在NBMA型网络，LSA以单播方式发送到DR和BDR，并且DR以单播方式发送这些更新。

LSA通过LS类型、LS标识和宣告路由器来识别，并通过序列号、校验和、老化时间判断LSA新旧。

Seq: 序列号（Seq）的范围是0x80000001到0x7fffffff。

Checksum: 校验和（Checksum）计算除了Age字段以外的所有字段，每5分钟校验1次。

Age: 范围是0到3600秒，16位长.当路由器发出1个LSA后，就把Age设置为0，当这个LSA经过1台路由器以后，Age就会增加，1个LSA保存在LSDB中的时候，老化时间也会增加。

当收到相同的LSA的多个实例（LS类型、LS标识、宣告路由器相同）的时候，将通过下面的方法来确定哪个LSA是最新的：

1. 比较LSA实例的序列号，越大的越新。
2. 如果序列号相同，就比较校验和，越大越新。
3. 如果校验和也相同，就比较老化时间，如果只有1个LSA拥有MaxAge(3600秒)的老化时间，它就是最新的。
4. 如果LSA老化时间相差15分钟以上，（叫做MaxAgeDiff），老化时间越小的越新。
5. 如果上述都无法区分，则认为这2个LSA是相同的。

组播扩展OSPFOSPF区域

区域长度32位，可以用10进制，也可以类似于IP地址的点分十进制，分3种通信量

1. Intra-Area Traffic：域内通信量
2. Inter-Area Traffic：域间通信量
3. External Traffic：外部通信量

路由器类型

1. **Internal Router**：域内路由器
2. **ABR(Area Border Router)**：区域边界路由器
3. **Backbone Router(BR)**：骨干路由器
4. **ASBR(Autonomous System Boundary Router)**：自治系统边界路由器。

虚链路（Virtual Link）

以下2中情况需要使用到虚链路：

1. 通过一个非骨干区域连接到一个骨干区域。
2. 通过一个非骨干区域连接一个分段的骨干区域两边的部分区域。

虚链接是一个逻辑的隧道（Tunnel），配置虚链接的一些规则：

1. 虚链接必须配置在2个ABR之间。
2. 虚链接所经过的区域叫**Transit Area**，它必须拥有完整的路由信息。
3. Transit Area不能是Stub Area^[72]。
4. 尽可能的避免使用虚链接，它增加了网络的复杂程度和加大了排错的难度。

OSPF区域—OSPF的精华

Link-state 路由在设计时要求需要一个层次性的网络结构.

OSPF网络分为以下2个级别的层次：

骨干区域 (backbone or area 0)

非骨干区域 (nonbackbone areas)

在一个OSPF区域中只能有一个骨干区域，可以有多个非骨干区域，骨干区域的区域号为0。

为了避免回环的产生，各非骨干区域间是不可以交换LSA信息的，他们只有与骨干区域相连，通过骨干区域相互交换信息。

非骨干区域和骨干区域之间相连的路由叫边界路由（ABRs-Area Border Routers），只有ABRs记载了接入各区域的所有路由信息。各非骨干区域内的非ABRs只记载了本区域内的路由表，若要与外部区域中的路由相连，只能通过本区域的ABRs，由ABRs连到骨干区域的BR，再由骨干区域的BR连到要到达的区域。

组播扩展OSPF两种类型

编辑^[73]

组播扩展OSPFOSPF LSA类型

- 1.类型 1:Router LSA：每个路由器都将产生Router LSA，这种LSA只在本区域内传播，描述了路由器所有的链路和接口，状态和开销。
- 2.类型 2:Network LSA：在每个多路访问网络中，DR都会产生这种Network LSA，它只在产生这条Network LSA的区域泛洪描述了所有和它相连的路由器（包括DR本身）。
- 3.类型 3:Network Summary LSA：由ABR路由器始发，用于通告该区域外部的目的地址.当其他的路由器收到来自ABR的Network Summary LSA以后，它不会运行SPF算法，它只简单的加上到达那个ABR的开销和Network Summary LSA中包含的开销,通过ABR，到达目标地址的路由和开销一起被加进路由表里，这种依赖中间路由器来确定到达目标地址的完全路由（full route）实际上是距离矢量路由协议的行为。
- 4.类型 4:ASBR Summary LSA:由ABR发出，ASBR汇总LSA除了所通告的目的地是一个ASBR而不是一个网络外，其他同Network Summary LSA.
- 5.类型 5:AS External LSA:发自主ASBR路由器，用来通告到达OSPF自治系统外部的目的地，或者OSPF自治系统那个外部的缺省路由^[74]的LSA.这种LSA将在全AS内泛洪（4个特殊区域除外）
- 6.类型 6:Group Membership LSA
- 7.类型 7:NSSA External LSA:来自非完全Stub区域（not-so-stubby area）内ASBR路由器始发的LSA通告它只在NSSA区域内泛洪，这是与LSA-Type5的区别。
- 8.类型 8:External Attributes LSA
- 9.类型 9:Opaque LSA(link-local scope,)
- 10.类型 10:Opaque LSA(area-local scope)
- 11.类型 11:Opaque LSA(AS scope)

组播扩展OSPFOSPF路由类型

- 1. 所有路由器都计算前往其所有在区域中每个目的地的最佳路径，并将他们加入到路由表中。这些是1类LSA和2类LSA,在路由表中用路由指示符O(OSPF)表示。
- 2. 所有路由器都计算前往互联网络中其他区域的最佳路径。在这些路径是区域间路由（3类和4类LSA）,在路由表中用路由指示灯IA（区域间）表示。

3. 除末节区域内的路由器外，所有路由器都计算前往外部自治系统中目标网络的最佳路径（5类）。这些路由是1类外部路由（E1）还是2类外部路由（E2）取决于配置。在路由表中，1类外部路由用O E1表示，2类外部路由用O E2表示。类似的还有ON1/ON2，与OE1/OE2相同，不过是来自NSSA区域的路由。

组播扩展 OSPFOSPF度量值

编辑^[75]

在Cisco路由器中，使用公式100Mbit/带宽（单位为Mbit）来计算的，但是，在带宽等于100Mbits的链路上，成本为1.在大于100Mbits的链路上这个值就不是很好

RouterA(config-if)#ip ospf cost interface-cost

成本越低，链路越好

RouterA(config-router)#atuoo-cost reference-bandwidth ref-bw

其中cost：1～65535 ref-bw：1～4294967

组播扩展 OSPF特殊区域

编辑^[76]

- 1. 特殊区域是指人为定义的一些区域，它们在逻辑中一般位于OSPF区域的边缘只与骨干区域相连。
- 2. 常见的特殊区域有以下几类：

(1)Stub区域

由于并不是每个路由器都需要外部网络的信息，为了减少LSA泛洪量和路由表条目，就创建了末梢区域，位于Stub边界的ABR将宣告一条默认路由到所有的Stub区域内的内部路由器，不接受自治系统外部信息，同时也拒绝4类LSA，Stub区域不能包含ASBR(除非ABR也是ASBR)。

(2)Totally Stub区域

Cisco专用区域，不接受自治系统外部路由和自治系统内其他区域汇总路由。

(3)NSSA区域

不接受自治系统外部信息，但NSSA可以包含ASBR。

(4)Totally NSSA区域

Cisco专用区域，不接受自治系统外部路由和自治系统内其他区域汇总路由

区域限制

- 1. 所有位于stub area^[77]的路由器必须保持LSDB信息同步，并且它们会在它的Hello包中设置一个值为0的E位（E-bit），因此这些路由器是不会接收E位为1的Hello包，也就是说在stub area里没有配置成stub router的路由器将不能和其他配置成stub router的路由器建立邻接关系。
- 2. 不能在stub area中配置虚链接（virtual link），并且虚链接不能穿越stub area。
- 3. stub area里的路由器不可以是ASBR。
- 4. stub area可以有多个ABR，但是由于默认路由的缘故，内部路由器无法判定哪个ABR才是到达ASBR的最佳选择。
- 5. NSSA允许外部路由被宣告OSPF域中来，同时保留Stub Area的特征，因此NSSA里可以有ASBR,ASBR将使用type7-LSA来宣告外部路由，但经过ABR,Type7被转换为Type5.7类LSA通过OSPF报头的一个P-bit作Tag，如果NSSA里的ABR收到P位设置为1的NSSA External LSA，它将把LSA类型7转换为LSA类型5.并把它洪泛到其他区域中；如果收到的是P位设置为0的NSSAExternal LSA，它将不会转换成类型5的LSA，并且这个类型7的LSA里的目标地址也不会被宣告到NSSA的外部NSSA在IOS11.2后支持。
- 6. totally stub area完全的stub区域，连类型3的LSA也不接收^[78]。

组播扩展 OSPF数据包类型和格式

编辑^[79]

类型号 包 作用 可靠性

1 HELLO 1、用于发现邻居2、建立邻接关系3、维持邻接关系4、确保双向通信5、选举DR和BDR

2 Database Description 数据库的描述 DBD 可靠

3 Link-state Request 链路状态请求包 LSR 可靠

4 Link-state Update 链路状态更新包 LSU 可靠

5 Link-state Acknowledgment 链路状态确认包 LSACK

Hello包的格式如下：

[80]

数据库描述包的格式如下：

[81]

链路状态请求包格式如下：

[82]

链路状态更新包的格式如下：

[83]

链路状态确认包的格式如下：

[84]

AS 自治系统（autonomous system）：一组相互管理下的网络，它们共享同一个路由选择方法，自治系统由地区再划分并必须由IANA分配一个单独的16位数字。地区通常连接到其他地区，使用路由器创建一个自治系统。

组播扩展 OSPFOSPF基本配置命令

编辑^[85]

配置LOOPBACK接口地址

ROUTER(config)#interface loopback 0

ROUTER(config)#ip address IP地址 掩码

- ospf区域的配置

router ospf 100 /* 区域号在锐捷低端设备(交换机S3550，路由器R2621)里是不需要指定区域号的*/

network 192.168.1.0 0.0.0.255 area 0

router-id 192.168.2.1 手动设置router-id

area 1 default-cost 50 手动设置开销

#clean ip ospf process

- 配置ospf明文认证

interface s0

ip ospf authentication

ip ospf authentication-key <; 密码>

- 配置ospf密文认证

interface s0

ip ospf authentication

ip ospf message-digest-key 1 md5 7 <; 密码>

- debug ip ospf adj 开启ospf调试

show ip protocols

show ip ospf interface s0

- 手动配置接口开销，带宽，优先级

inter s0

ip ospf cost 200

bandwidth 100

ip ospf priority 0

- 虚链路的配置

router ospf 100

area <area-id> virtual-link <router-id>

show ip ospf virtual-links

Show ip ospf border-routers

Show ip ospf process-id

Show ip ospf database

show ip ospf database nssa-external

- OSPF路由归纳

Router ospf 1\\对ASBR外部的路由进行路由归纳

Summary-address 200.9.0.0 255.255.0.0

Router ospf 1\\执行AREA1到AREA0的路由归纳

Area 1 range 192.168.16.0 255.255.252.0

- 配置末节区域

IR area <area-id> stub

ABR area <area-id> stub

- 配置完全末节区域

IR area <area-id> stub

ABR area <area-id> stub no-summary

- 配置NSSA

ASBR router ospf 100

area 1 nssa

ABR router ospf 100

area 1 nssa default-information-originate

OSPF协议主要优点

- 1、OSPF是真正的LOOP- FREE（无路由自环）路由协议。源自其算法本身的优点。（链路状态及最短路径树算法）
- 2、OSPF收敛速度快：能够在最短的时间内将路由变化传递到整个自治系统。
- 3、提出区域（area）划分的概念，将自治系统划分为不同区域后，通过区域之间的对路由信息的摘要，大大减少了需传递的路由信息数量。也使得路由信息不会随网络规模的扩大而急剧膨胀。
- 4、将协议自身的开销控制到最小。见下：
 - 1) 用于发现和维护邻居关系的是定期发送的是不含路由信息的hello报文，非常短小。包含路由信息的报文时是触发更新的机制。（有路由变化时才会发送）。但为了增强协议的健壮性，每1800秒全部重发一次。
 - 2) 在广播网络中，使用组播地址（而非广播）发送报文，减少对其它不运行ospf 的网络设备的干扰。
 - 3) 在各类可以多址访问的网络中（广播，NBMA），通过选举^[86]DR，使同网段的路由器之间的路由交换（同步）次数由 O（N*N）次减少为 O（N）次。
 - 4) 提出NSSA区域的概念，使得NSSA区域内不再传播引入的ASE路由。
 - 5) 在ABR（区域边界路由器）上支持路由聚合，进一步减少区域间的路由信息传递。
 - 6) 在点到点接口类型中，通过配置按需拨号属性（OSPF over On Demand Circuits），使得ospf不再定时发送hello报文及定期更新路由信息。只在网络拓扑真正变化时才发送更新信息。
- 5、通过严格划分路由的级别（共分四级），提供更可信的路由选择。
- 6、良好的安全性，ospf支持基于接口的明文及md5 验证。
- 7、OSPF适应各种规模的网络，最多可达数千台。

更多图册^[87]

参考资料

- 组播扩展 OSPF 介绍^[88] . 百度百科[引用日期2013-08-2]

中国通信学会^[89] ,

Links

1. javascript;;
2. <http://baike.baidu.com/view/10812319.htm>