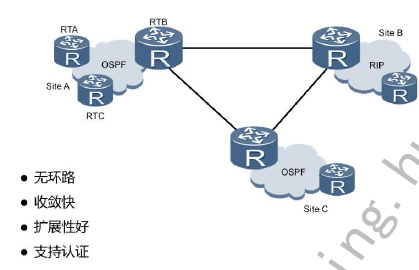
开放式最短路由优先OSPF(Open Shortest Path First)协议是IETF定义的一种基于链路状态的内部网关路由协议。   
RIP是一种基于距离矢量算法的路由协议，存在着收敛慢、易产生路由环路、可扩展性差等问题，目前已逐渐被OSPF取代。

学习目标

1. 掌握OSPF的工作原理
2. 掌握OSPF的基本配置

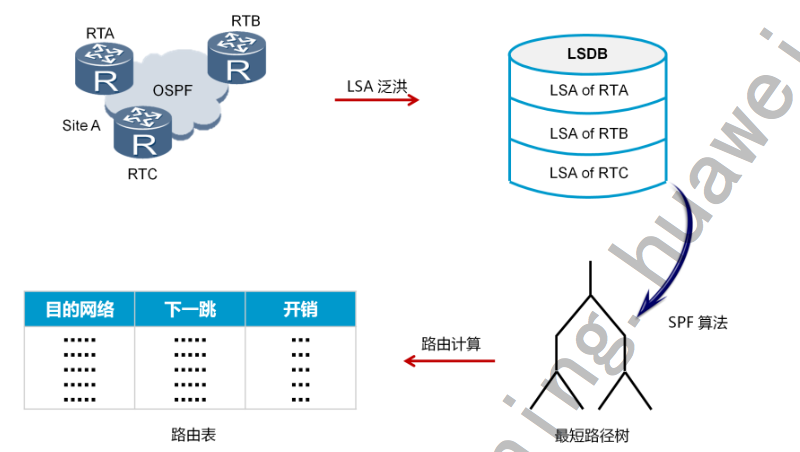
开放式最短路径优先（OSPF）

   
OSPF是一种基于链路状态的路由协议,它从设计上就保证了无路由环路。OSPF支持区域的划分,区域内部的路由器使用SPF最短路径算法保证了区域内部的无环路。OSPF还利用区域间的连接规则保证了区域之间无路由环路。   
OSPF支持触发更新,能够快速检测并通告自治系统内的拓扑变化。   
OSPF可以解决网络扩容带来的问题。当网络上路由器越来越多,路由信息流量急剧增长的时候,OSPF可以将每个自治系统划分为多个区域, 并限制每个区域的范围。OSPF这种分区域的特点,使得OSPF特别适用于大中型网络。OSPF还可以同其他协议(比如多协议标记切换协议 MPLS)同时运行来支持地理覆盖很广的网络。   
OSPF可以提供认证功能。OSPF路由器之间的报文可以配置成必须经过认证才能进行交换。

与RIP协议的比较

| **\** | **OSPF** | **RIPv2** | **RIPv1** |
| --- | --- | --- | --- |
| 协议类型 | 链路状态 | 距离矢量 | 距离矢量 |
| CIDR | 支持 | 支持 | 不支持 |
| VLSM | 支持 | 支持 | 不支持 |
| 自动聚合 | 不支持 | 支持 | 支持 |
| 手动聚合 | 支持 | 支持 | 不支持 |
| 路由泛洪 | 组播更新 | 周期组播更新 | 周期广播 |
| 路径开销 | 带宽 | 跳数 | 跳数 |
| 路由收敛 | 快 | 慢 | 慢 |
| 跳数限制 | 无 | 15 | 15 |
| 邻居认证 | 支持 | 支持 | 不支持 |
| 分级网络 | 支持（区域） | 不支持 | 不支持 |
| 更新 | 事件触发更新 | 路由表更新 | 路由表更新 |
| 路由计算 | Dijkstra | Bellman-ford | Bellman-ford |

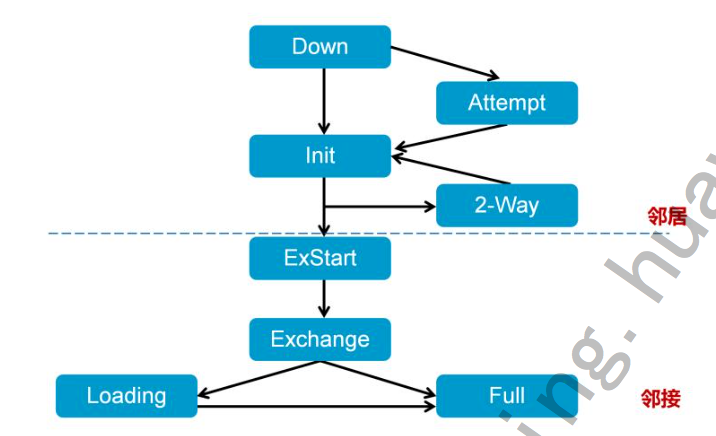
OSPF原理介绍

   
OSPF要求每台运行OSPF的路由器都了解整个网络的链路状态信息, 这样才能计算出到达目的地的最优路径。OSPF的收敛过程由链路状态公告LSA(Link State Advertisement)泛洪开始,LSA中包含了路由器已知的接口IP地址、掩码、开销和网络类型等信息。收到LSA的路由器都可以根据LSA提供的信息建立自己的链路状态数据库LSDB(Link State Database),并在LSDB的基础上使用SPF算法进行运算,建立 起到达每个网络的最短路径树。最后,通过最短路径树得出到达目的网络的最优路由,并将其加入到IP路由表中。

OSPF报文

   
OSPF直接运行在IP协议之上,使用IP协议号89。   
OSPF有五种报文类型,每种报文都使用相同的OSPF报文头。   
1. Hello报文:最常用的一种报文,用于发现、维护邻居关系。并在广播和NBMA(None-Broadcast Multi-Access)类型的网络中选举指定路由器DR(Designated Router)和备份指定路由器BDR( Backup Designated Router)。   
2. DD报文:两台路由器进行LSDB数据库同步时,用DD报文来描述自己的LSDB。DD报文的内容包括LSDB中每一条LSA的头部(LSA的头部可以唯一标识一条LSA)。LSA头部只占一条LSA的整个数据量的一小部分,所以,这样就可以减少路由器之间的协议报文流量。   
3. LSR报文:两台路由器互相交换过DD报文之后,知道对端的路由器有哪些LSA是本地LSDB所缺少的,这时需要发送LSR报文向对方请求缺少的LSA,LSR只包含了所需要的LSA的摘要信息。   
4. LSU报文:用来向对端路由器发送所需要的LSA。   
5. LSACK报文:用来对接收到的LSU报文进行确认。

邻居状态机

   
邻居和邻接关系建立的过程如下:   
1. Down:这是邻居的初始状态,表示没有从邻居收到任何信息。   
2. Attempt:此状态只在NBMA网络上存在,表示没有收到邻居的任何信息,但是已经周期性的向邻居发送报文,发送间隔为HelloInterval（一般是10s） 。如果RouterDeadInterval（一般为HelloInterval的四倍）间隔内未收到邻居的Hello报文,则转为Down状态。   
3. Init:在此状态下,路由器已经从邻居收到了Hello报文,但是自己不在所收到的Hello报文的邻居列表中,尚未不邻居建立双向通信关系 。   
4. 2-Way:在此状态下,双向通信已经建立,但是没有不邻居建立邻接关系。这是建立邻接关系以前的最高级状态。   
5. ExStart:这是形成邻接关系的第一个步骤,邻居状态变成此状态以后,路由器开始向邻居发送DD报文。主从关系是在此状态下形成的 ,初始DD序列号也是在此状态下决定的。在此状态下収送的DD报文不包含链路状态描述。   
6. Exchange:此状态下路由器相互发送包含链路状态信息摘要的DD报文,描述本地LSDB的内容。   
7. Loading:相互发送LSR报文请求LSA,发送LSU报文通告LSA。   
8. Full:路由器的LSDB已经同步。

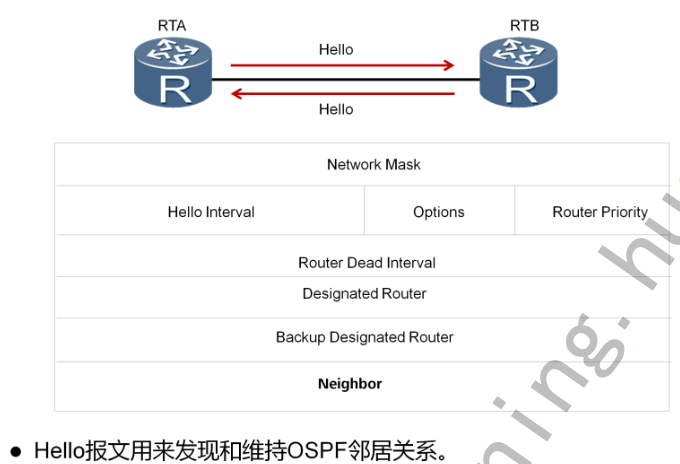
Router ID、邻居和邻接

Router ID是一个32位的值,它唯一标识了一个自治系统内的路由器,可以为每台运行OSPF的路由器上可以手动配置一个Router ID,或者指定 一个IP地址作为Router ID。如果设备存在多个逻辑接口地址,则路由器使用逻辑接口中最大的IP地址作为Router ID;如果没有配置逻辑接口, 则路由器使用物理接口的最大IP地址作为Router ID。在为一台运行 OSPF的路由器配置新的Router ID后,可以在路由器上通过重置OSPF进程来更新Router ID。通常建议手动配置Router ID,以防止Router ID 因为接口地址的变化而改变。

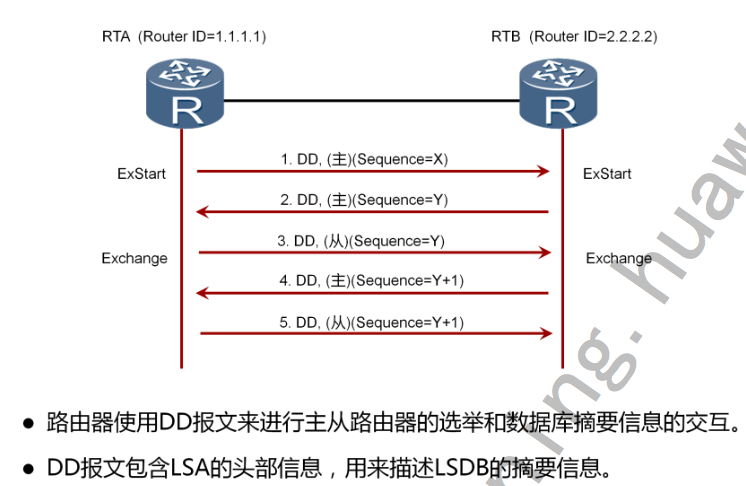
**OSPF Router-ID的确定顺序：**   
1. ospf 1 router id命令   
2. router id 命令   
3. 最大的Loopback接口IP地址   
4. 最大的物理接口IP地址

运行OSPF的路由器之间需要交换链路状态信息和路由信息,在交换这些信息之前路由器之间首先需要建立邻接关系。   
邻居(Neighbor):   
OSPF路由器启动后,便会通过OSPF接口向外发送Hello报文用于发现邻居。收到Hello报文的OSPF路由器会检查报文中所定义的一些参数, 如果双方的参数一致,就会彼此形成邻居关系。   
邻接(Adjacency):   
形成邻居关系的双方不一定都能形成邻接关系,这要根据网络类型而定。 只有当双方成功交换DD报文,并能交换LSA之后,才形成真正意义上的邻接关系。   
路由器在发送LSA之前必须先发现邻居并建立邻居关系。   
本例中,RTA通过以太网连接了三个路由器,所以RTA有三个邻居,但不能说RTA有三邻接关系。

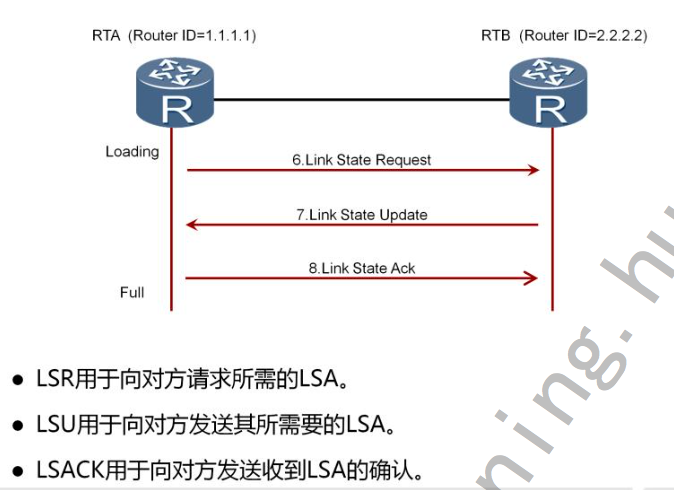
邻居发现

   
OSPF的邻居发现过程是基于Hello报文来实现的,Hello报文中的重要字段解释如下:   
1. Network Mask:发送Hello报文的接口的网络掩码。   
2. HelloInterval:发送Hello报文的时间间隔,单位为秒。   
3. Options:标识发送此报文的OSPF路由器所支持的可选功能。具体的可选功能已超出这里的讨论范围。   
4. Router Priority:发送Hello报文的接口的Router Priority,用于选举DR和BDR。   
5. RouterDeadInterval:失效时间。如果在此时间内未收到邻居发来的 Hello报文,则认为邻居失效;单位为秒,通常为四倍HelloInterval 。   
6. Designated Router:发送Hello报文的路由器所选举出的DR的IP地址。如果设置为0.0.0.0,表示未选举DR路由器。   
7. Backup Designated Router:发送Hello报文的路由器所选举出的BDR的IP地址。如果设置为0.0.0.0,表示未选举BDR。   
8. Neighbor:邻居的Router ID列表,表示本路由器已经从这些邻居收到了合法的Hello报文。   
如果路由器发现所接收的合法Hello报文的邻居列表中有自己的Router ID,则认为已经和邻居建立了双向连接,表示邻居关系已经建立。   
验证一个接收到的Hello报文是否合法包括:   
1. 如果接收端口的网络类型是广播型,点到多点或者NBMA,所接收的Hello报文中Network Mask字段必须和接收端口的网络掩码一致, 如果接收端口的网络类型为点到点类型或者是虚连接,则不检查 Network Mask字段;   
2. 所接收的Hello报文中Hello Interval字段必须和接收端口的配置一致 ;   
3. 所接收的Hello报文中Router Dead Interval字段必须和接收端口的配置一致;   
4. 所接收的Hello报文中Options字段中的E-bit(表示是否接收外部路由信息)必须和相关区域的配置一致

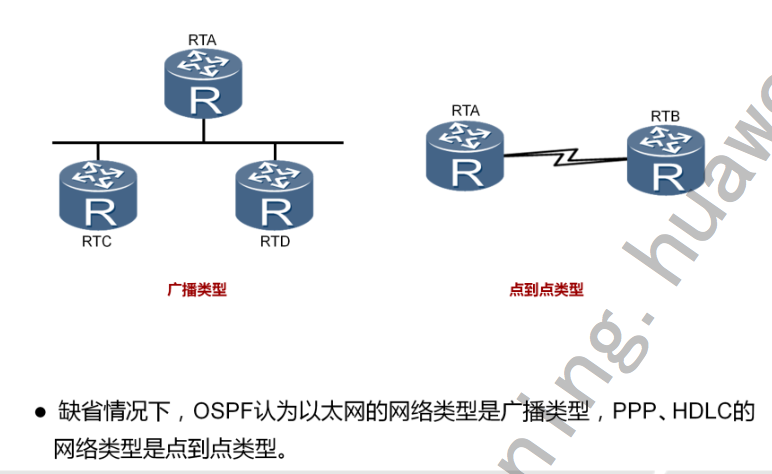
数据库同步

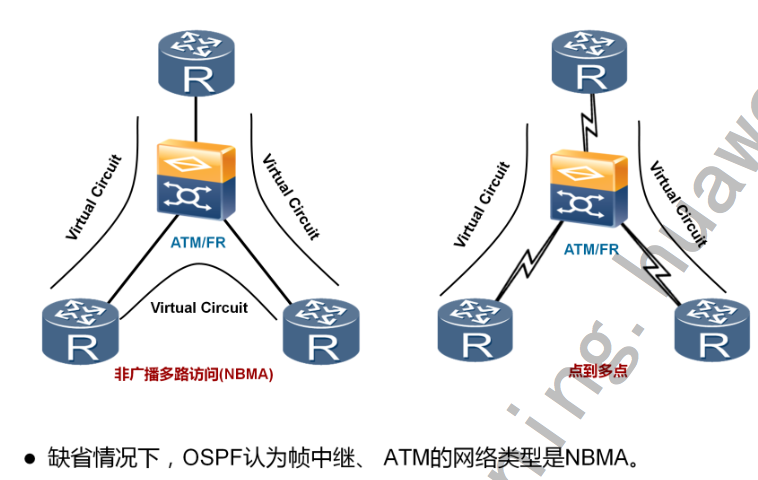
   
如图所示,路由器在建立完成邻居关系之后,便开始进行数据库同步, 具体过程如下:   
1. 邻居状态变为ExStart以后,RTA向RTB发送第一个DD报文,在这个报文中,DD序列号被设置为X(假设),RTA宣告自己为主路由器 。   
2. RTB也向RTA发送第一个DD报文,在这个报文中,DD序列号被设置为Y(假设)。RTB也宣告自己为主路由器。由于RTB的Router ID比RTA的大,所以RTB应当为真正的主路由器。   
3. RTA发送一个新的DD报文,在这个新的报文中包含LSDB的摘要信息,序列号设置为RTB在步骤2里使用的序列号,因此RTB将邻居状态改变为Exchange。   
4. 邻居状态变为Exchange以后,RTB发送一个新的DD报文,该报文中包含LSDB的描述信息,DD序列号设为Y+1(上次使用的序列号加1)。   
5. 即使RTA不需要新的DD报文描述自己的LSDB,但是作为从路由器 ,RTA需要对主路由器RTB发送的每一个DD报文进行确认。所以, RTA向RTB发送一个内容为空的DD报文,序列号为Y+1。   
发送完最后一个DD报文之后,RTA将邻居状态改变为Loading;RTB收到最后一个DD报文之后,改变状态为Full(假设RTB的LSDB是最新最全的,不需要向RTA请求更新)。

建立完全邻接关系

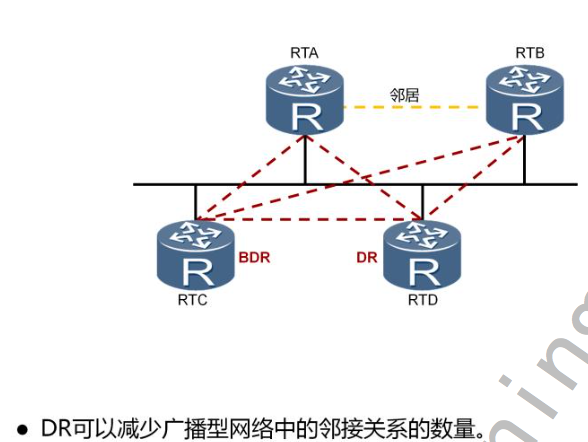
   
6. 邻居状态变为Loading之后,RTA开始向RTB发送LSR报文,请求那些在Exchange状态下通过DD报文发现的,而且在本地LSDB中没有的链路状态信息。   
7. RTB收到LSR报文之后,向RTA发送LSU报文,在LSU报文中,包含了那些被请求的链路状态的详细信息。RTA收到LSU报文之后, 将邻居状态从Loading改变成Full。   
8. RTA向RTB发送LSACK报文,用于对已接收LSA的确认。 此时,RTA和RTB之间的邻居状态变成Full,表示达到完全邻接状态。

OSPF支持的网络类型

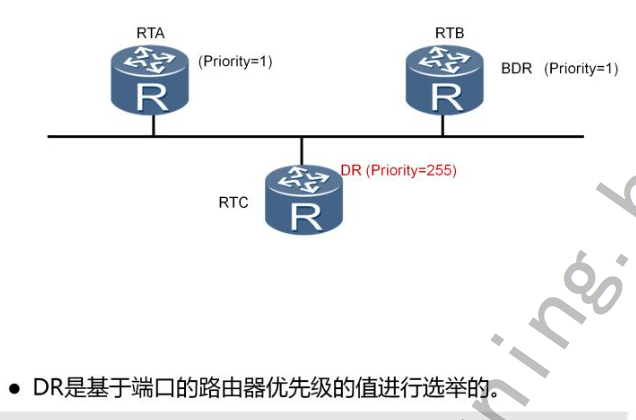
   
OSPF定义了四种网络类型,分别是点到点网络,广播型网络,NBMA 网络和点到多点网络。   
点到点网络是指只把两台路由器直接相连的网络。一个运行PPP的64K 串行线路就是一个点到点网络的例子。   
广播型网络是指支持两台以上路由器,并且具有广播能力的网络。一个含有三台路由器的以太网就是一个广播型网络的例子。

默认情况下，接口的封装协议决定了网络类型   
如果接口的封装协议是PPP，HDLC，帧中继的点到点子接口，默认的网络类型为P2P。   
如果接口的封装协议是以太网的，默认的网络类型为broadcast   
如果接口封装是帧中继或者帧中继的点到多点子接口，ATM接口，默认是NBMA型网络   
任何接口的默认网络类型，可以被人为修改。   
   
OSPF可以在不支持广播的多路访问网络上运行,此类网络包括在hub- spoke拓扑上运行的帧中继(FR)和异步传输模式(ATM)网络,这些网络的通信依赖于虚电路。OSPF定义了两种支持多路访问的网络类型: 非广播多路访问网络(NBMA)和点到多点网络(Point To Multi- Points)。   
1. NBMA:在NBMA网络上,OSPF模拟在广播型网络上的操作,但是 每个路由器的邻居需要手动配置。NBMA方式要求网络中的路由器组成全连接。   
2. P2MP:将整个网络看成是一组点到点网络。对于不能组成全连接的网络应当使用点到多点方式,例如只使用PVC的不完全连接的帧中继网络。

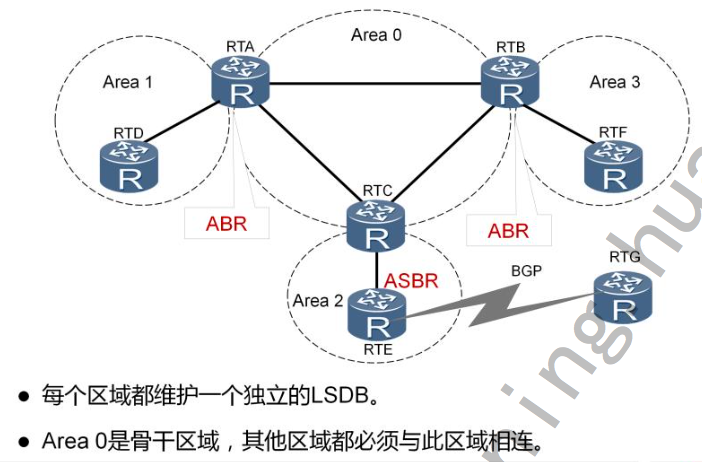
DR&BDR

   
每一个含有至少两个路由器的广播型网络和NBMA网络都有一个DR和BDR。   
DR和BDR可以减少邻接关系的数量,从而减少链路状态信息以及路由信息的交换次数,这样可以节省带宽,降低对路由器处理能力的压力。 一个既不是DR也不是BDR的路由器只与DR和BDR形成邻接关系并交换链路状态信息以及路由信息,这样就大大减少了大型广播型网络和 NBMA网络中的邻接关系数量。在没有DR的广播网络上,邻接关系的数量可以根据公式n(n-1)/2计算出,n代表参不OSPF的路由器接口的数量。 在本例中,所有路由器之间有6个邻接关系。当指定了DR后,所有的路由器都不DR建立起邻接关系,DR成为该广播网络上的中心点。   
BDR在DR収生故障时接管业务,一个广播网络上所有路由器都必须同BDR建立邻接关系。本例中使用DR和BDR将邻接关系从6减少到了5, RTA和RTB都只需要同DR和BDR建立邻接关系,RTA和RTB之间建立的是邻居关系。   
此例中,邻接关系数量的减少效果并不明显。但是,当网络上部署了大量路由器时,比如100台,那么情况就大不一样了。

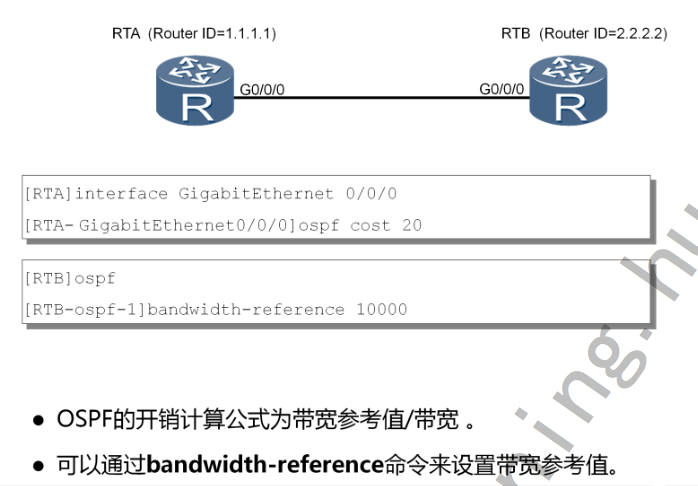
DR&BDR选举

   
在邻居収现完成之后,路由器会根据网段类型进行DR选举。在广播和NBMA网络上,路由器会根据参不选举的每个接口的优先级进行DR选举 。优先级取值范围为0-255,值越高越优先。缺省情况下,接口优先级为1。如果一个接口优先级为0,那么该接口将不会参与DR或者BDR的选举。如果优先级相同时,则比较Router ID,值越大越优先被选举为DR。   
为了给DR做备份,每个广播和NBMA网络上还要选举一个BDR。BDR也会与网络上所有的路由器建立邻接关系。   
为了维护网络上邻接关系的稳定性,如果网络中已经存在DR和BDR, 则新添加进该网络的路由器不会成为DR和BDR,不管该路由器的 Router Priority是否最大。如果当前DR发生故障,则当前BDR自动成为新的DR,网络中重新选举BDR;如果当前BDR发生故障,则DR不变, 重新选举BDR。这种选举机制的目的是为了保持邻接关系的稳定,使拓扑结构的改变对邻接关系的影响尽量小。

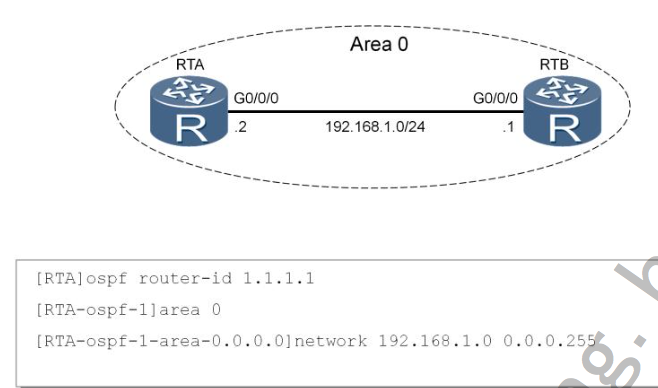
OSPF区域

   
OSPF支持将一组网段组合在一起,这样的一个组合称为一个区域。   
划分OSPF区域可以缩小路由器的LSDB规模,减少网络流量。   
区域内的详细拓扑信息不向其他区域収送,区域间传递的是抽象的路由信息,而不是详细的描述拓扑结构的链路状态信息。每个区域都有自己的LSDB,不同区域的LSDB是不同的。路由器会为每一个自己所连接到的区域维护一个单独的LSDB。由于详细链路状态信息不会被发布到区域以外,因此LSDB的规模大大缩小了。   
Area 0为骨干区域,为了避免区域间路由环路,非骨干区域之间不允许直接相互发布路由信息。因此,每个区域都必须连接到骨干区域。   
运行在区域之间的路由器叫做区域边界路由器ABR(Area Boundary Router),它包含所有相连区域的LSDB。自治系统边界路由器ASBR( Autonomous System Boundary Router)是指和其他AS中的路由器交换路由信息的路由器,这种路由器会向整个AS通告AS外部路由信息。   
在规模较小的企业网络中,可以把所有的路由器划分到同一个区域中, 同一个OSPF区域中的路由器中的LSDB是完全一致的。OSPF区域号可以手动配置,为了便于将来的网络扩展,推荐将该区域号设置为0,即骨干区域。

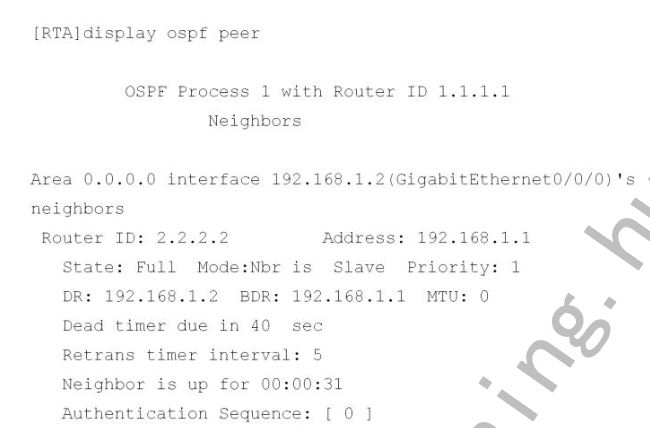
OSPF开销

   
OSPF基于接口带宽计算开销,计算公式为:接口开销=带宽参考值÷带 宽。带宽参考值可配置,缺省为100Mbit/s。以此,一个64kbit/s串口的开销为1562,一个E1接口(2.048 Mbit/s)的开销为48。   
命令bandwidth-reference可以用来调整带宽参考值,从而可以改变接口开销,带宽参考值越大,开销越准确。在支持10Gbit/s速率的情况下, 推荐将带宽参考值提高到10000Mbit/s来分别为10 Gbit/s、1 Gbit/s和 100Mbit/s的链路提供1、10和100的开销。注意,配置带宽参考值时, 需要在整个OSPF网络中统一进行调整。   
另外,还可以通过ospf cost命令来手动为一个接口调整开销,开销值范围是1~65535,缺省值为1。

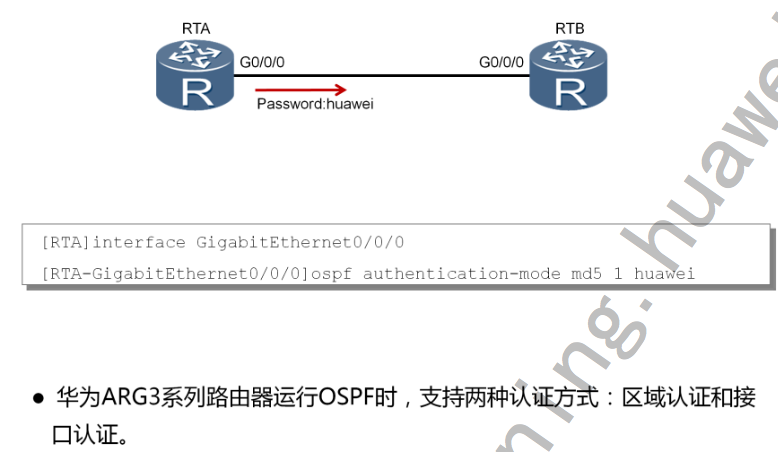
OSPF配置

   
在配置OSPF时,需要首先使能OSPF迚程。   
命令ospf [process id]用来使能OSPF,在该命令中可以配置迚程ID。如   
果没有配置进程ID,则使用1作为缺省迚程ID。   
命令ospf [process id] [router-id ]既可以使能OSPF进程,还   
同时可以用于配置Router ID。在该命令中,router-id代表路由器的ID。   
命令network用于指定运行OSPF协议的接口,在该命令中需要指定一 个反掩码。反掩码中,“0”表示此位必须严格匹配,“1”表示该地址可以为任意值。

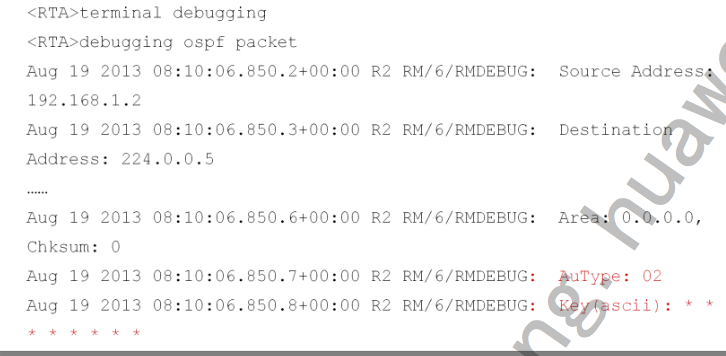
配置验证

   
命令display ospf peer可以用于查看邻居相关的属性,包括区域、邻居 的状态、邻接协商的主从状态以及DR和BDR情况。

OSPF认证

   
OSPF支持简单认证及加密认证功能,加密认证对潜在的攻击行为有更强的防范性。OSPF认证可以配置在接口或区域上,配置接口认证方式的优先级高于区域认证方式。   
接口戒区域上都可以运行ospf authentication-mode { simple [ [ plain ] | cipher ] | null } 命令来配置简单认证,参数plain表示使用显示密码,参数cipher表示使用密文密码,参数null表示不认证。   
命令ospf authentication-mode { md5 | hmac-md5 } [ key-id { plain | [ cipher ] } ] 用于配置加密认证,MD5是一种保证链路认证安全的加密算法(具体配置已在举例中给出),参数 key-id表示接口加密认证中的认证密钥ID,它必须不对端上的key-id一致 。

配置验证

   
在启用认证功能之后,可以在终端上进行调试来查看认证过程。   
debugging ospf packet命令用来指定调试OSPF报文,然后便可以查看认证过程,以确定认证配置是否成功。

总结

1. OSPF Hello报文章那个Router Dead Interval字段的作用是什么?
2. 在广播网络中，DR和BDR用来接收链路状态更新报文的地址是什么？   
   答：
3. Hello报文中的Router Dead Interval字段代表死亡间隔,如果在此时间内未收到邻居収来的Hello报文,则认为邻居失效。死亡间隔是 Hello间隔的4倍,在广播网络上缺省为40秒(因为Hello间隔缺省为 10秒)。
4. 在广播网络上,DR和BDR都使用组播地址224.0.0.6来接收链路状态更新报文。