

OSPF 邻居建立

扩展问题 1:为什么需要选举主从关系？

OSPF 是基于 IP 工作的，IP 是一个无连接的协议。同时，OSPF 又没有其它的确认机制来保证 DD 报文可靠的传输。所以，选举主从是为了保证 DD 报文有序、可靠的交互。另外 DD 的序列号还能充当确认的作用。

扩展问题 2:mtu 在哪里协商，是怎么协商的？如果协商不通过，会出现什么问题？

exstart 阶段会进行协商。（华为默认不协商。报文中 MTU 的字段填充为 0）

MTU 协商规则：对方 DD 报文中 mtu 字段的值 \leq 接收报文接口的 mtu 值。

如果协商不通过会导致邻接关系停留在 exstart 状态，无法建立 full 的邻接关系。

扩展问题 3:如果 mtu 不协商会带来什么问题？

如果不协商 mtu，可能会出现接口收到的 OSPF 报文的比接口 mtu 大的情况。

此时路由器就会丢弃这个报文，导致数据同步无法完成。

邻居状态停留在 exchange 或 loading 中。

扩展问题 4：ospf 在设计的时候为什么不直接泛洪 LSU？

路由器直接泛洪 lsdb 里所有的 lsa，邻居重复接收到已经存在的 lsa。造成带宽资源的浪费与设备性能的消耗。即 OSPF 在设计的时候，先与邻居交互发送携带 LSA 摘要信息的 DD 报文描述自身的数据库，然后再向邻居请求缺少的 LSA 信息，减少对带宽资源的消耗。

扩展问题 5：ospf 为什么用组播地址来建立邻居？

(1) 便于发现邻居

(2) 只有使能 ospf 协议的路由器才会解析 ospf 报文

(监听 224.0.0.5/224.0.0.6)，减小其它设备的性能消耗

OSPF 邻居关系

邻接关系建立过程中有哪些状态机？

邻居关系的状态：

Down：邻居会话的初始阶段，表明没有在邻居失效时间间隔内收到来自邻居路由器的 Hello 数据包。在 NBMA 网络上，此状态下仍然可以向静态配置的邻居发送 Hello 报文，

Init：从邻居收到的 hello 报文，但是活跃邻居列表没有自己的 Router-ID

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。

2-way：收到的 Hello 报文中包含有自己的 Router-ID；

邻接关系的状态：

Exstart：开始协商主从关系，并确定 DD 的序列号。

Exchange：主从关系协商完毕后继续交换 DD 报文（带 LSA 头部）

Loading：使用 LSR,LSU 以及 LSACK 报文，交互 LSA 的详细信息，进行数据库的同步

Full：LSR 重传列表为空。表示与邻居的数据库已经同步完成，邻接建立完成。收到 LSACK 不一定进入 full 状态。

影响邻居关系建立的因素有哪些（至少五种）？

(1) Router-ID 出现一致

(2) Area ID 不一致

(3) 认证不一致，认证字段放在 ospf 报文的头部，针对所有报文都做认证的

(4) hello/dead timer 不一致

(5) 区域类型不一致

option 位中，e=1、n=0 为普通/骨干区域，

e=0、n=0 为 stub 区域，

e=0、n=1 为 nssa 区域。

total stub 与 stub 可以建立邻居关系，

total nssa 与 nssa 可以建立邻居关系

(6) 接口的掩码不一致

MA 网络类型邻居接口的地址掩码要一致

p2p 网络类型默认不检查网段和掩码信息，网段不同可以建立邻居

(7) 网络类型为 NBMA 的时候，没有手工指定 PEER

(8) 设置 OSPF 静默接口

以下为影响邻接关系建立的因素

(1) 接口 mtu 不一致的情况下，开启了 mtu 检查

华为默认不开启检查，开启检查协商失败后邻居状态会停留在 Exstart

(2) MA 网络类型下，所有接口的 dr 优先级都调整为 0

扩展问题 1：SPF 的 Router-ID 如何配置，有什么不同？

自动选择：优先选择 Loopback 口。其次是物理接口。IP 地址越大越优

手工配置：设置全局 RouterID：Router-ID IP 地址

//设置所有相关路由协议的 RouterID

设置进程 RouterID：ospf X Router-ID XX

//创建一个 OSPF 进程，并配置 RouterID

进程 RouterID 优于全局 RouterID

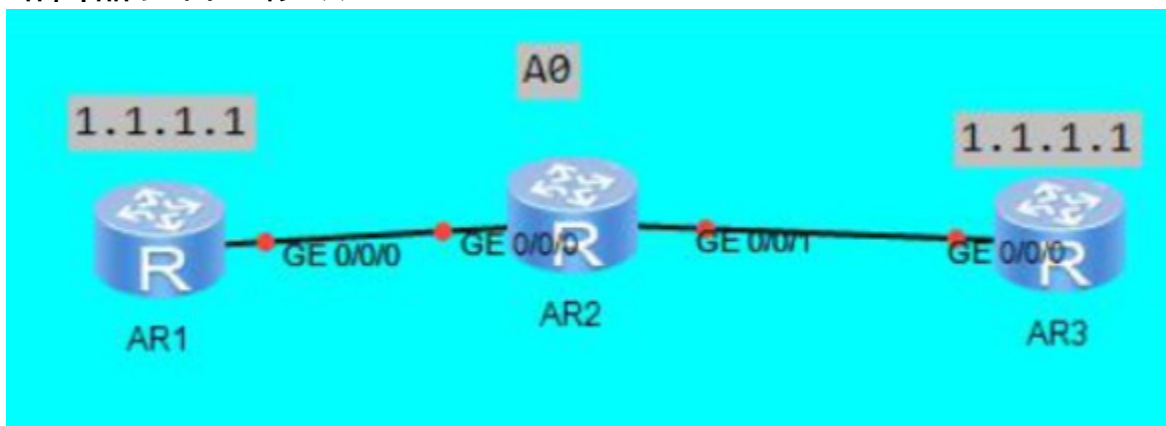
<Router>reset ospf X process (Router id 修改后需重置进程才会生效)

扩展问题 2：直连的两台路由器 router-id 相同会产生什么样的后果，为什么？

邻居无法正常建立。ospf 建立邻居的条件中要求 router-id 不能相同，如果相同则无法正常建立邻居关系。

扩展问题 3：ospf 的某个区域内的两台路由器的 router-id 的一样，会产生后果，为什么？

华为：区域内假如 Router-ID 相同，经过时一段间后其中一台路由器会自动修改 Router-ID



分析：

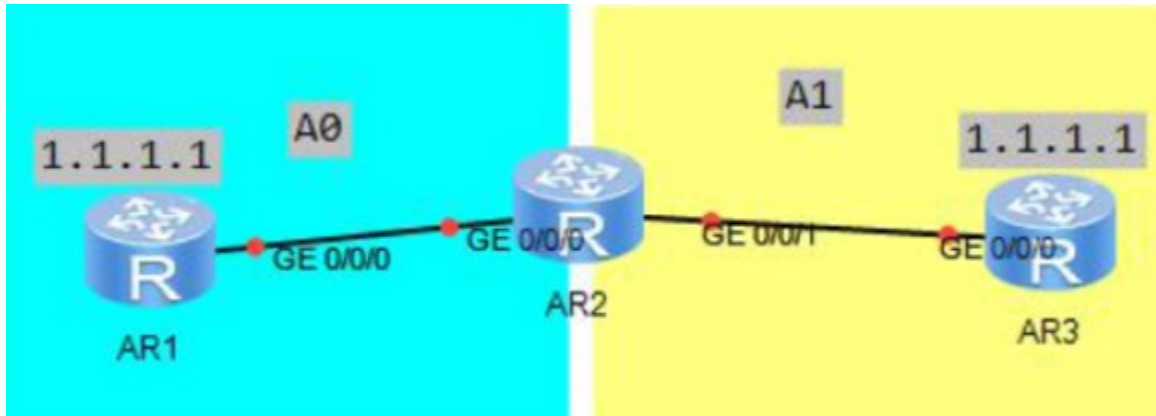
R1-R2 及 R2-R3 都可以正常建立邻居，但同步数据库的时候就会出现问題。

因为 R2 的 lsdb 中，关于 adv 为 1.1.1.1 的 LSA1 或 LSA2 只能有一份。

除此之外路由计算时也会出现问題：

R2 路由一直震荡，时有时无。在一段时间后，其中一台路由器修改自己的 Router-id。

扩展问题 4：ospf 中区域间的两台路由器的 Router-id 的一样，会产生后果，为什么？



假设 R1 和 R3 不引入外部路由的话，不会出现问题。
因为 ospf 在区域间使用 LSA3，LSA3 是由区域的 ABR 根据 LSA1、LSA2 产生的，adv 是 ABR 的 router-id。（R1 与 R3 彼此感知不到对方的存在）
区域间路由只是被当成叶子挂在 ABR 上，本区域内的 spt 树上不会出现在有相同 router-id 的节点，也就不会出现问题。
但是如果在相同 router-id 的设备上做引入的时候会出现问题。
因为 asbr 的 router-id 是需要被 ospf 域内的所有路由器所知道的。（lsa 5 类的传递范围）
若发现 asbr 的 router-id 与本设备的 router-id 一样时，会出现问题。

扩展问题 5：option 选项不一致一定建立不起邻居吗？
不一定，比如 DC 位（按需链路）。一端 DC=1，另一端 DC=0 时，会进行协商，最后 DC 都置位为 0，然后建立邻居；
（场景：cisco 和 huawei 的 Vlink 对接）

扩展问题 6：在 MA 网络中 IP address 不在同一网段中，能否用 Secondary address 建立邻居关系？P2P 和 MA 网络都是

怎么描述 secondary address 的？

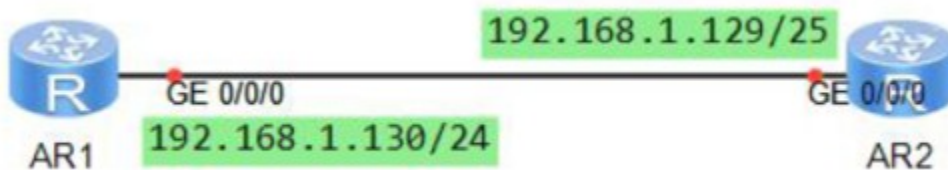
MA 网络中不会为 secondary address 产生 TransNet 类型的 1 类 LSA 和 2 类 LSA，只是将 secondary address 当成 loopback 接口来产生路由信息。

所以即使 secondary address 的网段相同，ip address 不同网段也不能建立邻居关系。

p2p 和 MA 网络的 secondary address 都是使用一个 stub 类型的 1 类 LSA 来描述。

扩展问题 7：为什么 MA 网络需要比较掩码？

如果在 MA 网络中掩码不一致也能建立邻居关系，会导致 DR 无法正确描述当前 MA 网络的路由信息。以下图为例：



当 R1 是 DR 时，2 类 LSA 中携带的路由信息为 192.168.1.0/24；

当 R2 是 DR 时，2 类 LSA 中携带的路由信息为 192.168.1.128/25。

无论哪台路由器成为 DR，所描绘的路由信息都与另一台路由器的接口的网络号不相符。

所以 OSPF 设计时要求 MA 网络中各个路由器的接口必须属于同一网段。

扩展问题 8：为什么 P2P 网络类型可以不用比较掩码？

P2P 网络类型的路由器可以独立表达自己接口的所有网络(使用 stub 类型的 1 类 LSA)，彼此间没有关系，建立邻居关系没有限制，所以 P2P 网络类型建立邻居关系时既不检查掩码，

也不检查源地址。（默认底层为 ppp 链路，不同网段也可访问）

OSPF 网络类型

OSPF 设计有几种网络类型？

网络类型	是否选举 DR	邻居的发现方式	Hello Timer	Dead Timer	报文发送方式
P2P	否	自动	10	40	组播：224.0.0.5
BMA	是	自动	10	40	组播：hello、LSU、LSAck 单播：DD、LSR、
NBMA	是	手动	30	120	所有报文都是单播发送
P2MP	否	自动	30	120	组播：hello 单播：其它报文

扩展问题 1：两端的网络类型不一致时，会不会建立邻居？能建立邻居，路由是否正常？

NBMA 只能与 NBMA 建立 full 邻接关系，与其他网络类型都无法建立 FULL 的邻居关系；（单播发现邻居）

broadcast 与 P2P，能建立 full 邻接关系，但不能正常计算路由；（构建的拓扑不一致）

broadcast 与 P2MP，通过修改 hello/dead timer 可以建立 full 邻接关系，但也不能计算路由；

P2P 与 P2MP，通过修改 hello/dead timer 可以建立 full 邻接关系，路由计算也正常。

总结：

因为 OSPF 的 hello 报文中没有描述网络类型的字段。

能否建立邻居关系主要看 hello 报文的发送间隔、holdtime，和 hello 报文的发送方式；

路由计算是否正常，主要看是否选举 DR；

扩展问题 2：OSPF 为什么会设计那么多网络类型？

OSPF 多用于企业网环境，企业网拓扑结构复杂，链路种类繁多。设计多种的网络类型可以对应不同的链路以及不同的网络拓扑。

扩展问题 5：为什么 P2P 网络不需要选举 DR，而 MA 网络需要选举 DR？

因为 P2P 网络中两台路由器点对点互联（互联设备少），选举 DR，再由 DR 来描绘拓扑信息，只会让拓扑的描绘更加复杂；

而在 MA 网络中，一个接口可能会有多个邻居，选举 DR 后，描绘的拓扑时会比较简单（各个节点连接到 DR 上，而不需要互相连接）；

另外在 MA 网络中 DRother 只和 DR/BDR 建立邻接关系，可以减少邻接关系的数量，减少报文的交互。

扩展问题 6：MA 网络中 DRother 只和 DR/BDR 建立邻接关系，这一机制是如何实现的？

在 exstart 状态中，DD 报文是通过单播方式发送的。

扩展问题 7：为什么 NBMA 无法和其他网络类型建立邻居关系？

因为 NBMA 是通过手动、单播的方式发现邻居、发送 hello 报文的，而其他网络类型即使配置了 peer，也不会单播的发送 hello 报文。

扩展问题 9：以太网链路将网络类型改为 P2P，接口不在同一网段是否能建立 OSPF 邻接关系？



答：可以，因为 P2P 是以 DIP : 224.0.0.5 发送 OSPF 的 hello 报文的，而只要宣告进 OSPF 进程中的接口就等于加入了这个组，所以 AR1 和 AR2 接口即使不在同一网段，也能建立邻居关系，同时也能学习到路由。但是无法互访，因为底层链路是以太网链路，不同网段互访需要有路由以及相应的 2 层封装。

扩展问题 10:网络类型可以更改吗？更改的命令式什么？
可以更改。

```
ospf network-type{ broadcast | nbma | p2mp | p2p [ peer-ip-ignore ] }
```

//配置 OSPF 接口的网络类型

DR 和 BDR

DR 和 BDR 的作用和竞选规则？

DR 在多路访问中可以减少邻接关系和 LSA 的泛洪，BDR 提供备份。

在 BMA 网络中，DRother 会使用目的地址为 224.0.0.6 发送 LSU、LSAck。DR 和 BDR 监听 224.0.0.6。

DR、BDR 使用目的地址为 224.0.0.5 发送 LSU、LSAck。DR other 监听 224.0.0.5。

竞选规则：

1. 接口优先级数字越大越优先 (0-255 , 0 优先级不能参与 D

R、BDR 选举)

2. RouterID 越大越优先

扩展问题 1:wait time 的作用是什么？

在 wait timer 时间内收集 MA 网络中的路由器信息，防止后期发生抢占。

实现方式：发出第一份 hello 报文的时候就开始计时，40s 之内不会认为自己是 DR。如果 40s 超时之后，还是没有收到携带 DR、BDR 参数字段的 Hello 报文，就认为自己是 DR，并发出 hello 包（DR 字段为自己的接口 IP 地址）。

扩展问题 2:wait time 太短会导致什么问题？

可能导致 MA 网络提前收敛，没有让最优的路由器成为 DR；

扩展问题 3:路由器如何判断这个 MA 网络的 DR 和 BDR 已经选举成功？

通过查看 Hello 报文中的 DR/BDR 字段是否为 0.0.0.0

扩展问题 4:如果有 DR 优先级为 255 的路由器接入，是否会发生抢占？

DR 优先级为 255 的路由器接入发送的第一份 hello 报文中 DR/BDR 都为 0.0.0.0

分为两种情况：

1. 不会：当网络已经收敛完成，已经选择出了 DR 时，新接入进来的路由器会等待 40s。这 40s 之内会收到填充有 DR、BDR 字段的 hello 报文。新路由器不会再参与竞选；
2. 会：网络中没有选举出 DR，因为可能原来的网络中没有一台路由器参选 DR（即优先级都为 0），那么新进来的优先级为 255 的路由器，则会在 40s 之后成为 DR；

扩展问题 5:是不是优先级最高的路由器成为 DR , 次优的成为 BDR ?

分情况而定 :

1. MA 网络中的各台路由器同时参与 DR 竞选 , 那么优先级高的会成为 DR。
2. 在一个已经收敛完成的网络中 , 有新的高优先级路由器接入进来 , 此时并不发生抢占。如果 DR down 了 , BDR 会升级成为 DR。这个新的高优先级路由器升级为 BDR。

扩展问题 6 : 两个收敛完成的 MA 网络 , 都有 DR 和 BDR。如果连接了起来 , 新的 MA 网络中 DR 和 BDR 分别是如何选举的 ?

DR 只会从 DR 的合集当中竞选 , BDR 只会从 BDR 的合集当中竞选。如果集合当中有路由器 , 优先从集合当中竞选。

扩展问题 7 : 三台路由器轮流间隔 40s 开启 , 谁是 DR , 谁是 BDR ?

假如某台路由器先开启 , 40s 之后再开启其它的路由器。这个时候最先开启的路由器 , 40s 内没有收集到网络上其它路由器的信息 , 它就成为 DR。

有没有办法让这三台路由器按正常优先级选举 ?

将中间的交换机或者 hub 关机 , 让 3 台路由器同时接入 MA 网络中。

扩展问题 4 : LSA 是如何比较新旧的 ?

1. 比较 Sequence Number , 越大越新
2. 比较 Checksum , 越大越新
3. 比较 Age , 3600s 最优
4. Age Time 之差小于 900s , 不比较 , 先收到先放入 LSDB 数

数据库（保留旧的 LSA）

5. Age Time 之差大于 900s，越小越优（保留新的 LSA）

扩展问题 5：1 类 LSA 有几种链路类型？

一条 LSA 可以描述多条链路。每条链路由 Link ID，Data，Type 和 Metric 组成

LinkID：此链接的对端标识

Data：描述此链接的附加信息

Type：链路类型（P2P、TransNet、StubNet、Virtual）

1 P2P 网段：描述连接到另外一个路由器；

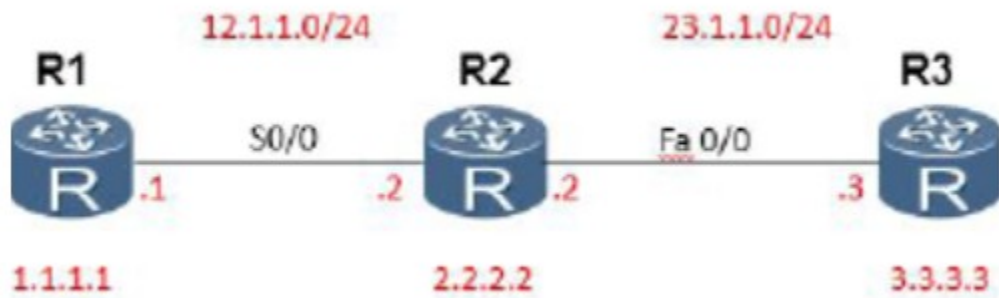
2 Transit 网段：描述使用该接口连接到一个虚节点（DR），例如广播型网段或 NBMA 网段；

3 Stub 网段：描述末节类型链路（网络），例如 Loopback 接口、Sub-IP 地址、P2P 和 P2MP 网络中链路的路由信息；

4 Virtual：一个从本路由器到虚连接对端 ABR 的连接；

Link Type	Link ID	Data	Metric
Point-to-point	邻居的 Router-ID	该网段上本地接口的 IP 地址	互连链路的 Cost
TransNet	DR 的接口 IP 地址	该网段上本地接口的 IP 地址	到伪节点的链路值
StubNet	该 Stub 网段的 IP 网络地址	该 Stub 网段的网络掩码	表示到达该网络的 Cost 值
Virtual	虚连接邻居的 Router-ID	去往该虚连接邻居的本地接口的 IP 地址	发出接口到达该邻居的 Cost 值

扩展问题 7：拓扑中 R1 上会产生几个 1 类 LSA，整个区域有几个 1 类 LSA，画图说明 1 类 LSA 是如何泛洪的？



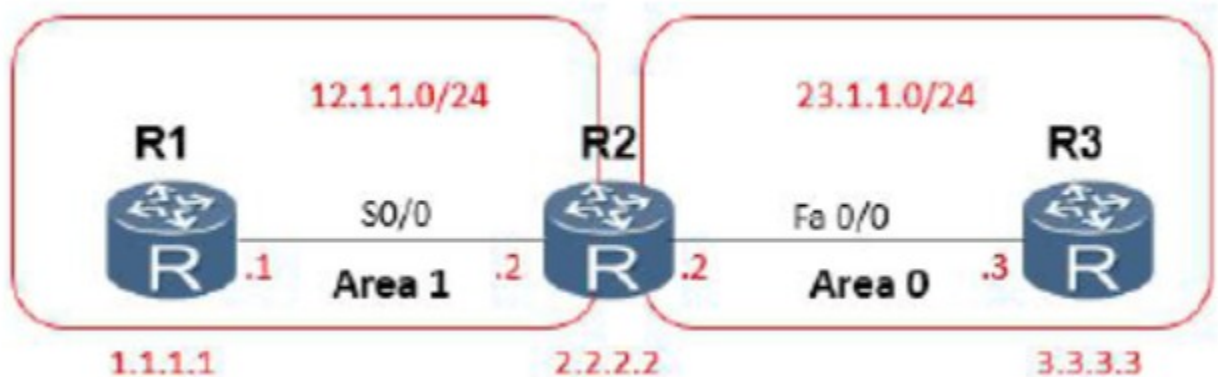
R1 产生 1 份 1 类 LSA，这份 1 类 LSA 里携带多个链路描述信息。在整个 area 内，有 3 份 1 类 LSA，每台路由器产生 1 份。1 类 LSA 在所属的区域内传播。

扩展问题 8：在上图的此区域中有几个 transit 类型的 1 类 LSA？Transit 是指向 DR 的，左侧的 P2P link 没有 P-Node，不需要 transit。

所以在 R2 和 R3 的 LSA1 中都会指向 DR 的 transit；
答案是 2 个 transit。

扩展问题 9：除了 1 类 LSA 外其他的 LSA 有没有 flag 字段呢？没有。因为 Flag 字段是用来在区域内标识路由器角色的。

扩展问题 11：在下图中，area 0 里的路由器上都会产生几类的 LSA。



Area1 中 R1 只会产生 LSA1。

Area0 里面的 R3 的 LSDB 中会有本区域里面 R2 /R3 的 LSA

1。

如果 R3 是 DR，则 R3 会产生 LSA2。LSA3 是 R2 产生的，携带邻居区域的网络。

扩展问题 12：OSPF 查找域间路由时，使用的 SPF 算法还是 DV 算法？

ABR 收到来自同区域其它 ABR 传来的 LSA3 后，它并不运行 SPF 算法。

相反地，它重新生成新的 LSA3（修改通告者以及开销）。然后继续在整个 OSPF 系统内扩散。

此行为依赖中间路由器代替确定到达目的地的全程路由（Full route）的做法是距离矢量协议的行为。

因此，虽然在区域内部 OSPF 协议是一个链路状态协议，但是使用了距离矢量的算法来查找域间路由。

扩展问题 15：详细描述一下 4 类 LSA 的产生过程？

当一台路由器引入外部路由之后，它就成为一台 ASBR 路由器。这时它产生的 1 类 LSA 里 Flag 字段的 E 位会置位，此 1 类 LSA 在区域内泛洪。区域内的 ABR 路由器收到 E 置位的 1 类 LSA 之后，会为这台 ASBR 路由器产生 4 类 LSA，并在 ABR 所连接的其它区域中泛洪。

扩展问题 16:5 类 LSA 有哪几种计算外部路由 metric 值的方式，有什么区别？

External Type 1 和 External Type 2，简称 E1 和 E2。

E2：默认，生成路由时不叠加内部 cost 值，只计算外部 cost 值。

E1：表示累加内部 cost 值，生成路由的 cost 为内部 cost 值+外部 cost 值。

内部 cost:本台路由器到 ASBR 的 cost 值。

外部 cost：ASBR 生成 5 类 LSA 时携带的 cost 值，默认为 1。

扩展问题 17:两种外部路由 metric 计算方式怎么进行优选比较？
通过修改 cost 值可以控制 OSPF 选路

1. 先比较类型，E1 优于 E2
2. 若类型都为 E1。则比较总开销（内部 cost+外部 cost）
3. 若类型都为 E2。先比较外部 cost 值，小的优先。再比较内部 cost 值。
4. 如果外部 cost 和内部 cost 值都一样。就负载均衡。

扩展问题 18:Forwarding Address 的作用？

防止外部路由次优路径。访问外部路由的时候，先转发到 FA 地址。

使用条件：

1. ASBR 与外部网络连接的接口被宣告进 ospf；
2. ASBR 与外部网络连接的接口没有被设置为静默接口；
3. ASBR 与外部网络连接的接口不是 OSPF P2P 或 P2MP 类型的；

扩展问题 20：5 类 LSA 和 7 类 LSA 有什么不同？

1. 泛洪范围，LSA5 是在整个 OSPF 区域内泛洪（除了特殊区域），而 LSA7 只在 NSSA 区域内泛洪；
2. 7 类 LSA 的 FA 地址永不为 0，而 5 类 LSA 的 FA 地址是可能为 0 的；

扩展问题 21：为什么需要在 ABR 上进行 7 类转 5 类的操作？
因为 7 类 LSA 只允许在 NSSA 区域泛洪。为了将 NSSA 区域引入的外部路由发布到其它区域，需要把 Type-7 LSA 转化为 Type-5 LSA 以便在整个 OSPF 网络中通告。

扩展问题 22：如果有多台 ABR 存在，是每台 ABR 都会进行 7 转 5 操作吗？

不会。因为各个路由器 7 转 5 后生成的 5 类 LSA 都是一样，只需让一台路由器进行 7 转 5 即可,达到可以减小其他区域的 LSDB 和路由表的规模。

规则：如果 NSSA 区域有多台 ABR，只有 RouterID 大的 ABR 路由器会进行 7 转 5 操作。

扩展问题 23：7 类转 5 类需要满足什么条件？

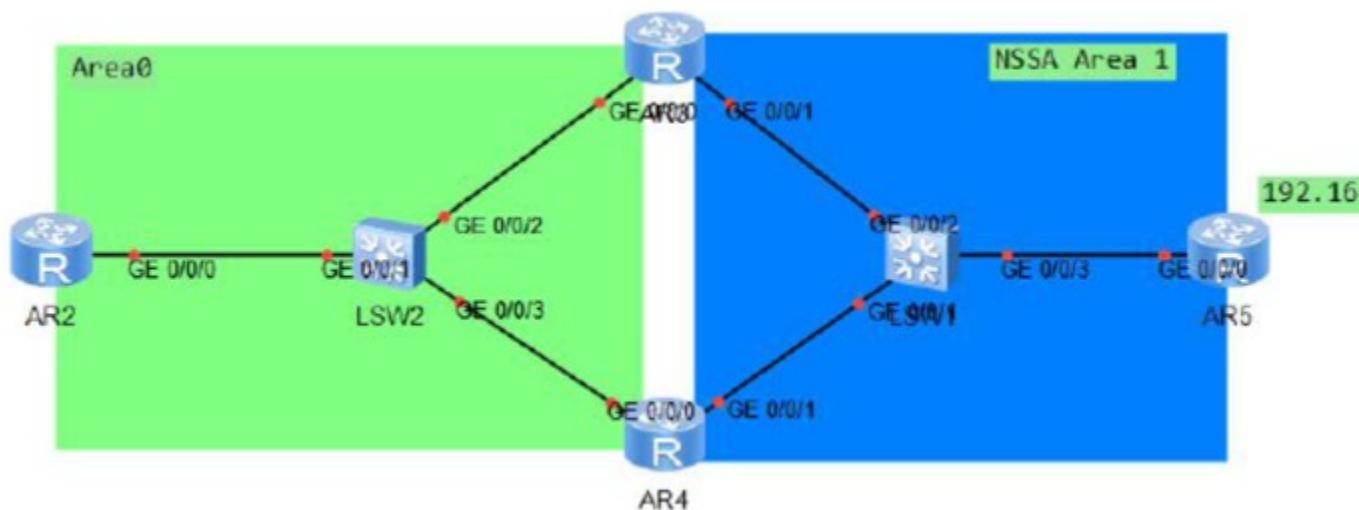
P 置位并且 Forwarding Address 不为 0 的 Type-7 LSA 才能被 ABR 转化为 Type-5LSA。

option 字段：P 用于告知转化路由器该条 7 类 LSA 是否需要转化。

Forwarding Address：用来表示该外部路由在 OSPF 域内的目标地址。

扩展问题 24：option 字段里的 P-bit，什么时候会不置位？

区域边界路由器产生的 7 类 LSA 不会置位 P-bit



当 AR4 引入一条外部路由的时候，会向 Area0 发送一条 LSA 5、向 NSSA 的 Area2 发送一条 LSA7。LSA7 会在 NSSA 中泛洪，AR3 也会收到这条 LSA7。

此时，如果 LSA7 的 P 置位；AR3 就会进行 7 转 5 的操作，将这条外部路由再传到 Area0 中。Area0 就有两条关于引入路由的重复 5 类 LSA。所以 AR4 在生成这条 LSA7 时，P-bit 不置位。

扩展问题 25：7 类 LSA 的 FA 地址是什么，有什么作用？选举规则是怎么样的？

NSSA 区域产生的 LSA7 的 FA 地址一定为非 0，在进行 LSA 7 转 5 也会将 FA 地址复制到 LSA5，主要为了避免次优路径（也可以防止环路，华为不认可）

LSA7 的 FA 地址选举规则，分为标准规则和华为规则：

标准规则：

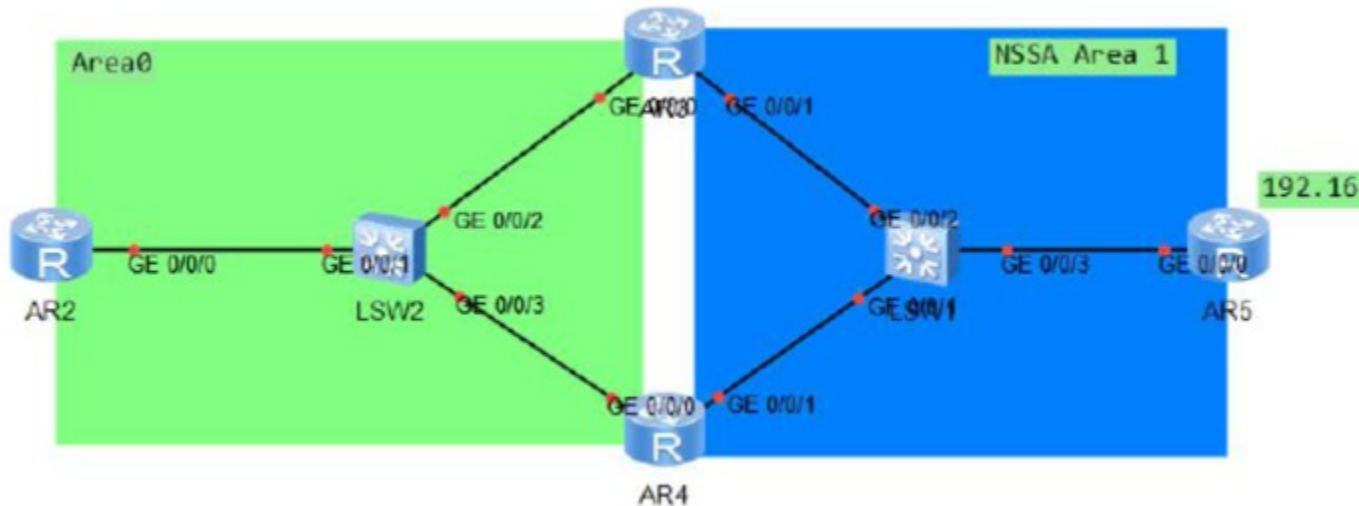
- 1.首先使用 LSA5 的 FA 选举条件，满足 5 类 LSA-FA 选举条件时，使用 5 类的选举方式（可以在 MA 网络中避免次优）
- 2.如果不满足 5 类 LSA 的 FA 地址选举条件，则在 ASBR 上宣告进了 OSPF 域内的 loopback 接口中选择 IP 地址最大的，没有 loopback 接口就选择 IP 地址最大的物理接口（以此来保证 FA 地址一定不为 0）

华为规则：

- 1.第一条和标准的一样，使用 LSA5 的 FA 选举条件，满足 5 类 LSA-FA 选举条件时，使用 5 类的选举方式，在 MA 网络中避免次优路径；
- 2.Forwarding Address 首选第一个 up 的 loopback 接口，没有 loopback 接口就选第一个 up 的物理接口

NSSA 区域的 ABR 不会为 NSSA 区域的 ASBR 产生 LSA4，因为 LSA 7 只会在区域内泛洪，没有产生必要。

扩展问题 26：R3 和 R4 都产生了缺省的 7 类 LSA 泛洪进 NSSA 区域，这时候它们会不会计算对方的缺省 7 类而导致环路的产生？



回答：不会。

OSPF 缺省路由的发布原则：如果 OSPF 路由器已经发布了缺省路由 LSA，那么不再学习其它路由器发布的相同类型缺省路由。即路由器只接收不计算其它路由器发布的相同类型的缺省路由 LSA。

扩展问题 27：NSSA 区域的 ASBR 有几种下发缺省路由的方式？

回答：只有 `nssa default-route-advertise` 一种方式并且要求 ASBR 的路由表中存在外部缺省路由，才会产生 Type7 LSA 的缺省路由。并且此 Type7 LSA 缺省路由不会在 ABR 上转换成 Type5 LSA 缺省路由泛洪到整个 OSPF 域。

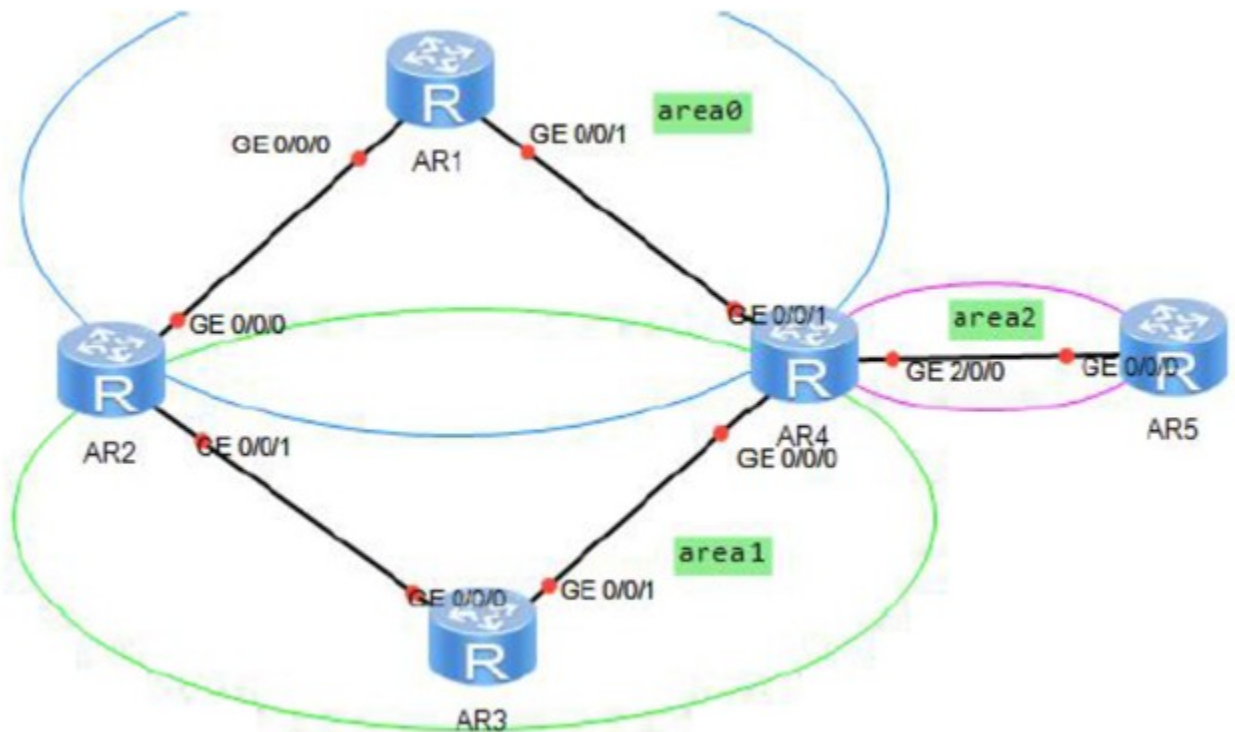
扩展问题 28：OSPF 路由选路的原则，及在什么情况下会负载

选路原则：区域内的>区域间的>区域外的

区域内：1、2LSA；区域间：3LSA；区域外：根据 5、7LSA 的 E1、E2 进行比较；

负载条件：1、cost 一致；2、区域一致

扩展问题 29: R4 与 R5 上分别引入外部路由，问 R2 如何去往这两条外部路由？



R2 通过 R1 访问 R5，R2 通过 R3 访问 R4

R2 通过 area 0 访问 R5 引入的外部路由，因为非骨干区域传来的 4LSA，ABR 不参与计算，R2 会选择 R1 作为他的下一跳

R2 通过 area 1 访问 R4 引入的外部路由。因为 R2 会通过 area0 和 area1 都收到 R4 的 5 类 LSA。R2 上去往 R4 会有两

个下一跳，但两个下一跳属于不同的区域，所以 R2 去往 R4 不能负载。如果 R2 通过两条路的 cost 值相同，R2 则会选择区域号大的作为 R2 的下一跳，R2 会选择 R3 作为下一跳。
(以此来减轻骨干区域的负担)

OSPF 防环、过滤、汇总

OSPF 是如何进行防环的？

(1) 区域内的防环：

LSA1，LSA2：根据 SPF 算法进行防环

(2) 区域间防环：

1. 特殊的区域结构：OSPF 要求所有的非 0 区域必须与骨干区域直接相连，区域间路由需经由骨干区域中转。

解读：OSPF 要求所有的非 0 区域必须与骨干区域直接相连，区域间路由需经由骨干区域中转。这个要求使得区域间的路由传递不能发生在两个非 0 的区域之间，这在很大程度上规避了区域间路由环路的发生，也使得 OSPF 的区域架构在逻辑上形成了一个类似星型的拓扑

2 区域间的水平分割：

a) 从一个非骨干区域学习到的 LSA3 不会再传回该非骨干区域 (因为 1LSA 优于 3LSA，不是根据区域 ID 判断，主要根据 ABR 的 LSDB 判断)

b) 完全意义上的 ABR 从非骨干区域收到的 LSA3 会接收但是不会参与计算也不会传回非骨干区域 (此规则对非完全意义上的 ABR 无效)

(3) 外部路由防环：

LSA4：由于 LSA4 类的产生以及泛洪范围与 LSA3 一致，所以 LSA4 的防环规则与 LSA3 一致。

LSA5、7：当 FA 地址为全零，根据 LSA4 类防环；如果 LSA4 无环，那么 LSA5，7 也无环。当 FA 地址为非全零，根据 LSA1-3 类防环；如果 LSA1-3 无环，那么 LSA5，7 也无环。

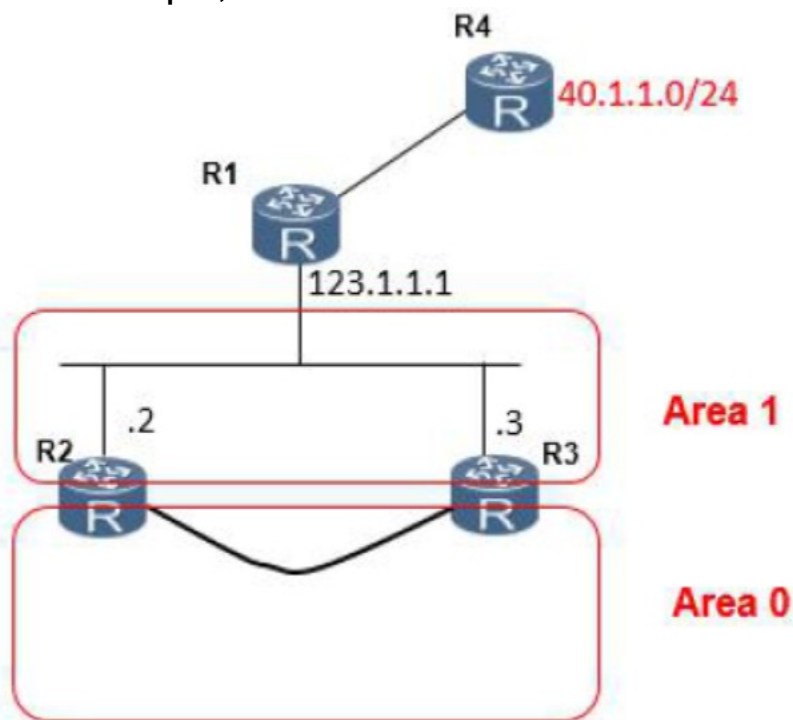
扩展问题 3:场景 1:

R2 上写一条静态路由 (ip route-static 40.1.1.0/24 123.1.1.

1)，重分发到 ospf，分析 R3 到 40.1.1.0/24 的数据怎么走？

R1 上写一条静态路由 (ip route-static 40.1.1.0/24 14.1.1.4) ,

重分发到 ospf，分析 R3 到 40.1.1.0/24 的数据怎么走？



回答：

问题一：R3 经过 R1，R4,访问 40.1.1.0 网络.

问题二：R3 经 R1，R4 访问 40.1.1.0 网络.

主要考察 LSA5 FA 地址何时为 0.0.0.0，何时为非 0.0.0.0；

在场景 1 中，在 R2 上添加静态路由，ip route-static 40.1.1.0/24 123.1.1.1，把静态路由导入到 ospf。

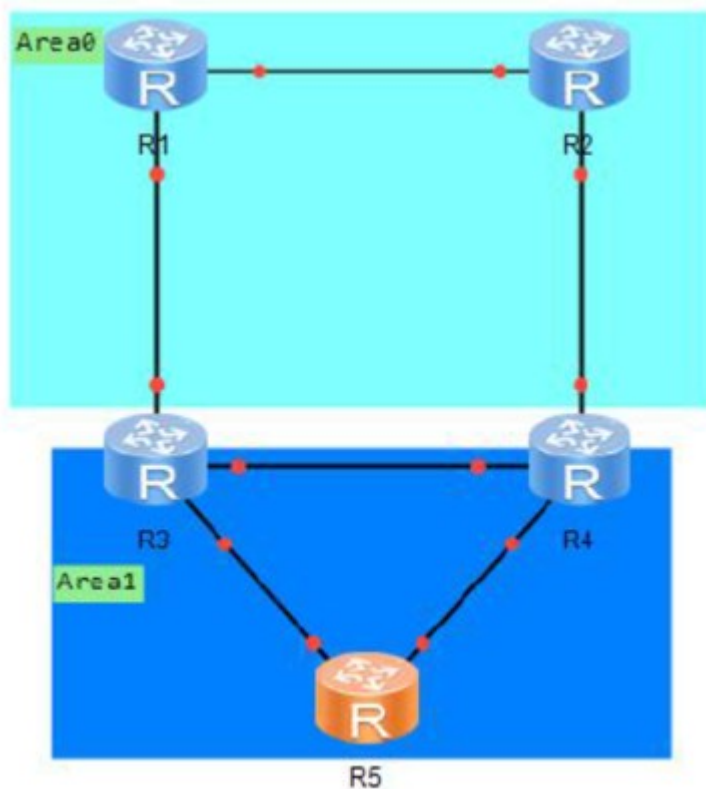
R2 是 ASBR，123.1.1.0/24 网络被发布到 ospf 域中，且出口网络类型是 broadcast，所以其产生的 LSA5 中 FA 是 123.1.1.1。R3 路由表中下一跳为 FA。

在场景 2 中，在 R1 上添加静态路由，ip route-static 40.1.1.0/24 14.1.1.4，并把静态路由导入 ospf。

R1 是 ASBR，14.1.1.0/24 网络没有发布到 ospf 中，所以静态路由的下一跳在 ospf 内部不可达，所以生成 LSA5 的 FA 地址是 0.0.0.0。

R3 的路由表访问 40.1.1.0/24 下一跳是 R1，123.1.1.1。

场景 2：OSPF Area0 (R1、R2 之间链路 down) 被分割后，在 R3 和 R4 上访问对端的 Area0 有什么问题？Area 1 的域内路由器 R5 还能访问到区域 0 的网段吗？



R3 和 R4 没有对端 Area0 的路由，因为它们无法计算对方从 Area1 传递过来的 3 类 LSA。

R5 可以通过 R3 和 R4 通告的 3 类 LSA 访问到 Area0 内的网段。

扩展问题 4:总结一下，在什么情况下 R3 不能访问对端 Area0 内的路由？请详细描述 R3 为什么没到对端 Area0 的路由？

当 R3 在 AREA0 有 Full 的邻接关系后，R3 就不能访问 AREA0 了。

因为 R3 此时是一台完全意义上的 ABR，从非骨干区域收到 3 类 LSA 是接收不计算的。

如何对 LSA 进行过滤？

(1) LSA1，LSA2 过滤：只能在接口下进行过滤：ospf filter-lsa-out all

(2) LSA3 过滤 : (只能在 ABR 上)

1 .接口下进行过滤 :

ospf filter-lsa-out summary

2. 在路由所属的区域下使用 filter 命令过滤 :

filter ip-prefix 172 import/export

ip ip-prefix 172 index 10 deny 172.16.0.0 22

ip ip-prefix 172 index 20 permit 0.0.0.0 0 less-equal 32

3 .汇总不通告的方式过滤 :

abr-summary 172.16.0.0 255.255.252.0 not-advertise

(区域下)

4 .ABR 的进程下过滤 :

filter-policy ip-prefix 172 import/export

(3) LSA4 过滤 :

在接口下使用命令 ospf filter-lsa-out all , 以过滤所有 LSA 的方式来过滤 4 类。

(4) LSA5/LSA7 过滤 : (产生 LSA5/7 的 ASBR 上)

○1 .接口下进行过滤 : ospf filter-lsa-out ase/NSSA

○2 .进程下进行过滤 : filter-policy ip-prefix import/export

○3 .汇总时进行过滤 : asbr-summary 50.1.1.0 255.255.255.0 not-advertise (进程下)

○4. 引入时进行过滤 : import-route 协议+ route-policy

//只有满足路由策略才能生成 5、7 类

LSA

LSA 能不能汇总 , 在哪里汇总 , 为什么 ?

LSA3 只能在 ABR 处 , 且是 LSA1/LSA2 转换为 LSA3 时 , 进行汇总。即最初产生 LSA3 的 ABR 上做汇总 , 其他位置不能汇总

骨干区域路由器配置 Vlink 之后无法进行 ABR 汇总 , 因为可

能会产生环路

LSA5 只能够在产生 LSA5 的 ASBR 处做汇总。如果是在 nssa 区域中，产生 LSA7 的 ASBR，则可以对 LSA7 汇总，也可以在发生 7/5 转换的位置对 LSA5 做汇总。

扩展问题 1:为什么 Area0 不能配置成特殊区域？

特殊区域都会过滤 5 类 LSA，如果 Area0 可以配置成特殊区域，那么将导致 Area0 和某些区域可能不存在外部路由而无法访问外部网络。

扩展问题 2:STUB 区域有哪些 LSA？能不能没有 LSA3？

LSA1、LSA2、LSA3。除了 ABR 下放默认路由的 LSA3，其他的 LSA3 可以没有。

扩展问题 3:total nssa 与 nssa 有什么区别？

total nssa 只是在 ABR 上过滤 LSA3

扩展问题 4:stub 和 nssa 区域有什么区别？

(1) stub 区域中不能存在 ASBR，而 nssa 区域虽然不接收其他区域传的外部路由，但是自己区域内的路由器可以引入外部路由；

(1) hello 报文上的区别，option 字段中置位不同

OSPF Virtual-link

有哪些常见的虚连接使用场景？

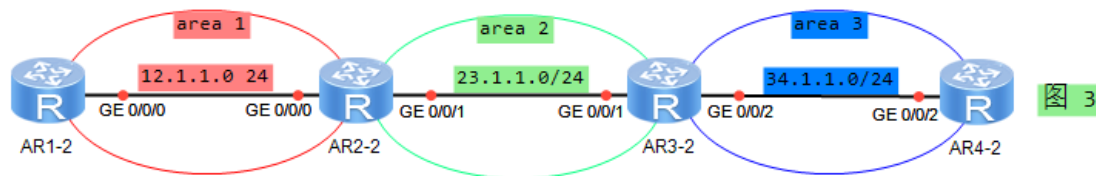
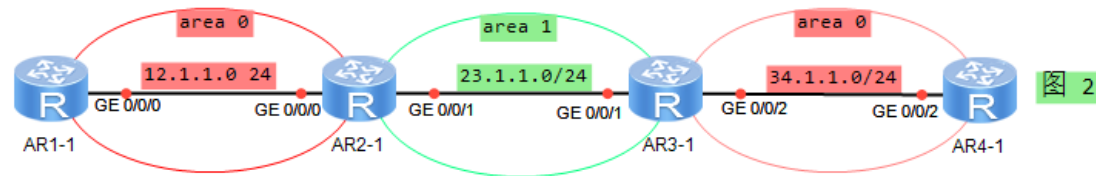
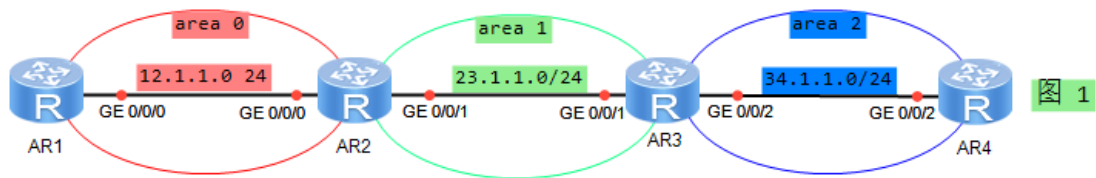
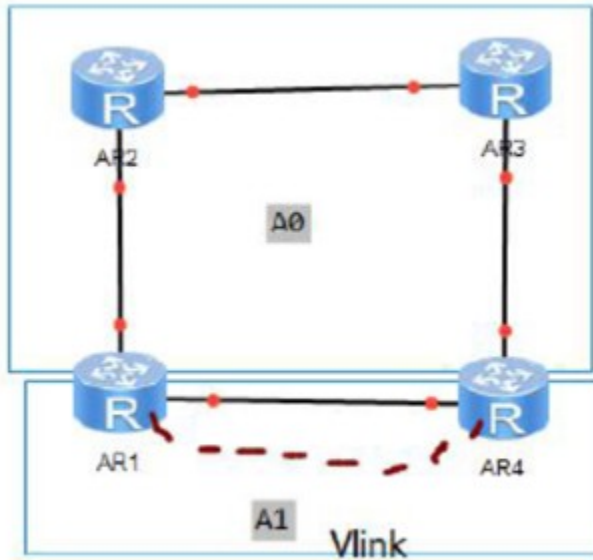


图 1 ， 图 2 ， 可以用虚链路 ， 图 3 无法实现

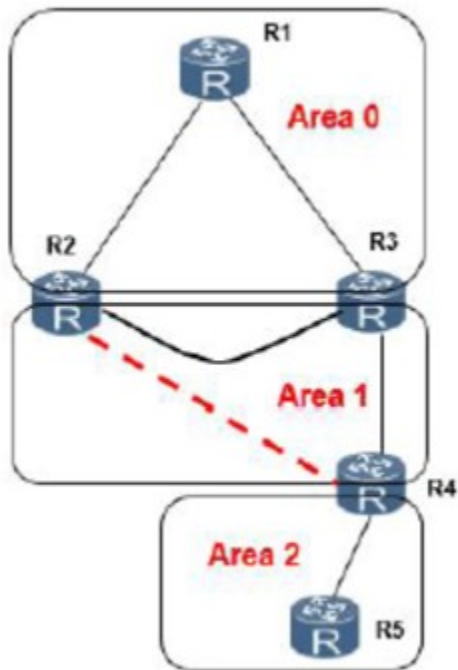
- (1) 把没有与骨干区域直接相连的非骨干区域通过一个非骨干区域 (也叫传输区域) 以逻辑通道的形式连接 ；
 - (2) 通过一个非骨干区域连接连接分割的骨干区域 ； (或者说把两个骨干区域通过一个非骨干区域连接起来)
 - (3) R1 和 R4 间启用 virtual link ， 用来防止 area 0 断开而成为不连续的骨干区域 ， 从而导致被分割的区域无法互相访问 (防止单点故障) 。
- 同时也能解决次优路径问题 ， 例如 R4 访问 R1 时 ， 流量可以走 R4 和 R1 之间的链路。



使用 virtual-link 有什么问题或缺点？

- 1、增加网络的复杂性
- 2、增加非骨干区域路由器的负担
- 3、不能随意修改 Router-ID
- 4、无法针对 Vlink 所在的区域做汇总
- 5、Vlink 配置不当会导致环路

扩展问题 2:画个图讲解一下，虚链路出现环路的场景？



如图所示：在 R2 与 R4 之间建立一条 Vlink。R5 上有一个 loopback 接口，地址为 5.5.5.5/32，宣告进了区域 2。

1. R4 作为区域 2 中的 ABR (R2 与 R4 之间的 Vlink 建立起来后，R4 就成为了一台 ABR)，会产生关于 5.5.5.5 网段的 3 类 LSA，并向区域 1 与区域 0 (Vlink) 泛洪。
2. 此时区域 1 中的关于 5.5.5.5 网段的这条 3 类 LSA 会传递给 R3，R3 在区域 1 中传递给 R2。
3. 区域 0 中的关于 5.5.5.5 网段的这条 3 类 LSA 会通过 Vlink 传递给 R2(Vlink 使用单播传递 LSA)，R2 经区域 0 传递给 R1，R1 经区域 0 传递给 R3(注意 R2 并不能直接把这条 3 类 LSA 传递给 R3，因为 R2 与 R3 之间的链路在区域 1 中)。
4. R3 作为 Area0 和 Area1 的 ABR，会将 Area0 中关于 5.5.5.5 的 3 类 LSA 再传回给 Area1。
5. 此时分析下 R3 的情况，R3 能收到 2 条关于 5.5.5.5 网段的 2 类 LSA：

a)由 R4 产生，在区域 0 中由 R4-R2-R1-R3 传递过来且 Metric 为 1 的

b)由 R4 产生，在区域 1 中由 R4-R3 传递过来且 Metric 为 1 的

6. 从非骨干区域收到的 3 类 LSA，ABR 能接收但不会使用这条 3 类 LSA，即 R3

虽然能收到 2 条关于 5.5.5.5 网段的 3 类 LSA，但它能使用的只有第一条，即由区域 0 传递过来的这条 3 类 LSA。

7. 此时 R3 要访问 5.5.5.5，会先把数据包转发给 R1，R1 转发给 R2，R2 知道去往 5.5.5.5 的数据包是要转发到 R4 的，但 R2 去往 R4 只能经过区域 1 传递给 R3，所以此时产生了 R3-R2-R1-R3 的环路。

总结：在华为设备上此环路的解决办法为修改 Vlink 的邻居建立路由器，在 R3 与 R4 之间配置 Vlink 就可以避免环路。

扩展问题 5：使用 V-link 时，如果每个 abr 上有多个接口，选择的是哪条链路？是怎么选举的？

本地 ABR 查找输区域的 SPF 树，选择一条到达对方 ABR 开销最小的路径。如果有多条路径开销一致，则随机选择一个接口发送。

扩展问题 6：RouterID 冲突或修改会导致不稳定，应该怎么防止？

设计的时候应该提前规划好 router-id，配置的时候提前指定 router-id 再配置 ospf 进程。一般情况下不建议修改 router-id。如果修改,对应的 Vlink 的 router-id 也要修改。

扩展问题 7：Vlink 为什么能够跨路由器建立，可不可以跨区域建立？

Vlink 的 HELLO 包发送是以单播的形式发送，所以可以跨路由建立。

Vlink 邻居不能跨区域建立，因为 Vlink 配置时 peer 指的是对

方的 RouterID，这个 RouterID 必须要在区域内的 SPF 树找到。如果找不到，则无法确定 Hello 包发送的目的地址，所以无法跨区域建立 Vlink。

扩展问题 9：VLINK 建立时，PEER 的时候是对方的 ROUTE ID，那么对方的 ROUTEID 是否需要宣告进 OSPF 中？

VLINK 的 PEER 是 ROUTER-ID，这个 ROUTER-ID 必须要在区域内的 SPF 树找到，在根据这两个 ROUTER-ID 互访的路径 COST 最小的 IP 地址做为 hello 包的源目 IP 地址，而不是以 ROUTER-ID 作为源目 IP 地址。所以不需要宣告进 OSPF 中。

扩展问题 10：做 VLINK 的时候需要注意的问题？

- 1、Vlink 只能配置在非骨干区域中，并且不能配置在特殊区域；
- 2、Vlink 配置时 peer 对方的 router-id；
- 3、Vlink 不能跨区域建立；
- 4、区域 0 中配置了认证，则 Vlink 也要配置认证；

加速收敛

OSPF 对 SPF 算法做了哪些加速收敛的改进？

PRC：当网络上路由发生变化的时候，只对发生变化的路由进行重新计算

I-SPF：除了第一次计算时需要计算全部节点外，每次只计算受到影响的节点，而最后生成的最短路径树 SPT 与原来的算法所计算的结果相同，大大降低了 CPU 的占用率，提高了网络收敛速度

扩展问题 1：一台区域内的路由器，它的 loopback 接口更新

了 IP 地址，会触发哪些更新计算方式？

i-SPF 和 PRC

OSPF 可以做哪些设置加速收敛？

(1) 智能定时器：配置更新 LSA 的时间、接收 LSA 的时间、spf 计算的时间（默认开启）

(2) OSPF 路由的收敛优先级：OSPF 按优先级收敛能够让某些特定的路由优先收敛。

(3) Smart-discover：在 MA 网络中邻居状态，或者 DR、BDR 发生变化时，设备不必等到 Hello 定时器到就可以立刻主动的向邻居发送 Hello 报文

a) 当邻居状态首次到达 2-way 状态，或者由 2-way 及更高状态回到 Init 状态时，立刻主动的向邻居发送 Hello 报文。

b) 该多址网络上的 DR、BDR 的接口状态发生变化时，会主动在该网段发送 Hello 报文，参与到 DR 或者 BDR 的选举中。

(4) 修短接口的 helloTimer、HelloDeadTimer。

(5) 与 BFD 进行联动，提高 ospf 的收敛速度。

(6) 尽量避免 DR 的存在，适当的链路可以通过调整网络类型来减少 DR 竞选时要等待的 wait timer 时间。

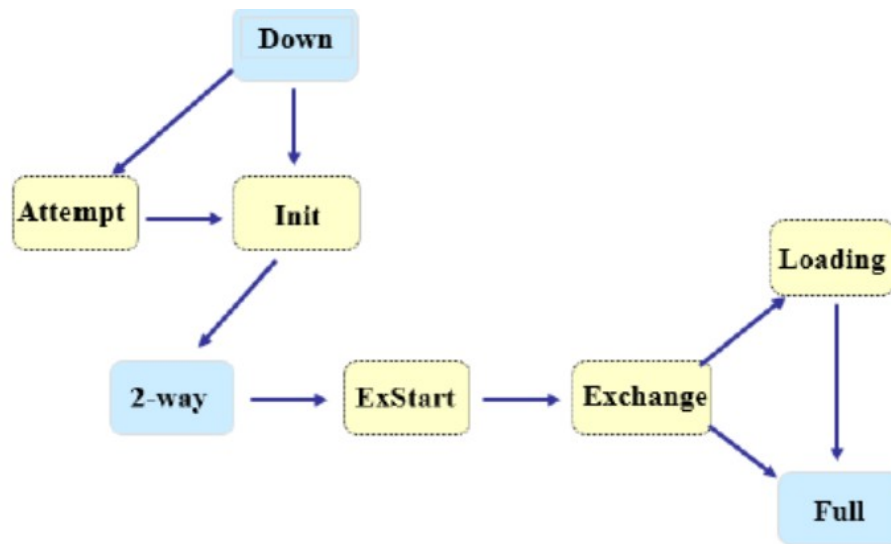
(7) 可以通过配置特殊区域、汇总路由，减少 LSA 的泛洪，以此来降低计算 LSA 的开销。

扩展问题 2：智能定时器具体是怎么实现快速收敛的？

在网络相对稳定、对路由收敛时间要求较高的组网环境中，可以通过智能定时器指定 LSA 的更新、被接收的时间间隔为 0，使得拓扑或者路由的变化可以通过 LSA 发布到网络中，或者立即被感知到，从而加快路由的收敛。

=====

OSPF 形成邻接关系所经历各个阶段有何意义？



1) 失效状态(Down)这是一个邻居会话的初始状态,用来指明在最近一个 Router Dead Interval 的时间内还没有收到来自邻居路由器的 Hello 数据包。除非在 NBMA 网络中的那些邻居路由器,否则,Hello 数据包是不会发送给那些失效的邻居路由器的。在 NBMA 网络环境中,Hello 数据包是每隔 Poll Interval 的时间发送一次的。如果一台邻居路由器从其他更高一些的邻居状态转换到了失效状态,那么路由器将会清空链路状态重传列表、数据库摘要列表和链路状态请求列表。

2) 尝试状态(Attempt)这种状态仅仅适用于 NBMA 网络上的邻居,在 NBMA 网络上邻居路由器是手工配置的。当 NBMA 网络上具有 DR 选取资格的路由器和其邻居路由器相连的接口开始变为有效(Active)时,或者当这台路由器成为 DR 或 BDR 时,这台具有 DR 选取资格的路由器将会把邻居路由器的状态转换到 Attempt 状态。在 Attempt 状态下,路由器将使用 HelloInterval 的时间代替 PollInterval 的时间来作为向邻居发送 Hello 数据包的时间间隔。

3) 初始状态(Init)这一状态表明在最近的 RouterDeadInterval

时间里路由器收到了来自邻居路由器的 Hello 数据包,但是双向通信仍然没有建立。路由器将会在 Hello 数据包的邻居字段中包含这种状态下或更高状态的所有邻居路由器的路由器 ID。

4) 双向通信状态(2-Way)这一状态表明本地路由器已经在来自邻居路由器的 Hello 数据包的邻居字段中看到了它自己的路由器 ID,这也就意味着,一个双向通信的会话已经成功建立了。在多址网络中,邻居路由器必须在这个状态或更高状态时才能有资格被选作该网络上的 DR 或 BDR。如果在 Init 状态下从邻居路由器那里收到一个数据库描述数据包,也可以引起邻居状态直接转换到 2-Way 状态。

5) 信息交换初始状态(Exstart)在这一状态下,本地路由器和它的邻居将建立起主/从关系,并确定数据库描述数据包的序列号,以便为数据库描述数据包的信息交换做准备。在这里,具有最高路由器 ID 的邻居路由器将成为“主”路由器。

6) 信息交换状态(Exchange)在这一状态下,本地路由器将向它的邻居路由器发送可以描述它整个链路状态数据库信息的数据库描述数据包。同时,在这个 Exchange 的状态下,本地路由器也会发送

链路状态请求数据包给它的邻居路由器,用来请求最新的 LSA。

7) 信息加载状态>Loading)在这一状态下,本地路由器将会向它的邻居路由器发送链路状态请求数据包,用来请求最新的 LSA 通告。虽然在 Exchange 状态下已经发现了这些最新的 LSA 通告,但是本地路由器还没有收到这些 LSA 通告。

8) 完全邻接状态(Full)在这一状态下,邻居路由器之间将建立起完全邻接关系,这种邻接关系出现在路由器 LSA 和网络 LSA 中。

ospfv3 新增了哪些 LSA ?

ospf v3 是为 ipv6 开发的.其中添加了 2 中新 LSA: Link-LSA 和 Intra-Area-Prefix-LSA

Router LSA 不再包含地址信息，使能 OSPFv3 的路由器为它所连接的每条链路产生单独的 Link LSA，将当前接口的链路本地地址以及路由器在这条链路上的一系列 IPv6 地址信息向该链路上的所有其它路由器通告。

Router LSA 和 Network LSA 中不再包含路由信息，这两类 LSA 中所携带的路由信息由 Intra Area Prefix LSA 来描述，该类 LSA 用来公告一个或多个 IPv6 地址前缀。

link-lsa 类型是 8,作用有三:

- 1,在链路上通告本地链路地址给其它邻居
- 2,通告本地链路上的邻居关于与自己本地链路相关联的 ipv6 的前缀列表
- 3,通告链路状态的选项集合

Intra-Area-Prefix-LSA 是在 area 范围内通告路由前缀.要通告的链路状态信息是参考 lsa 类型 1,2 而来的.