



G-CNA v2.0课程

讲师：沈老师





路由表组成

路由表作用：数据转发的语句，控制层

```
11.0.0.0/32 is subnetted, 1 subnets
    11.1.1.1 [110/11] via 12.1.1.1, 00:00:19, Ethernet0/0
12.0.0.0/8 is variably subnetted, 2 subnets, 2 masks
    12.1.1.0/24 is directly connected, Ethernet0/0
    12.1.1.2/32 is directly connected, Ethernet0/0
13.0.0.0/24 is subnetted, 1 subnets
    13.1.1.0 [110/20] via 12.1.1.1, 00:00:19, Ethernet0/0
100.0.0.0/8 is variably subnetted, 2 subnets, 2 masks
    100.1.1.0/24 is directly connected Ethernet0/1
    100.1.1.2/32 is directly connected Ethernet0/1
```

路由表三个重要组成：

目标网段+下一跳+出接口



形成路由表的方法

1. 直连路由
2. 静态路由（协议）:Static. IPv6 Static
3. 动态路由协议
 - 内部网关协议（IGP）
 - RIP(v1,v2) RIPng
 - EIGRPv4 EIGRPv6
 - OSPFv2 OSPFv3
 - ISIS
 - 外部网关协议（EGP）
 - BGP MP-BGP



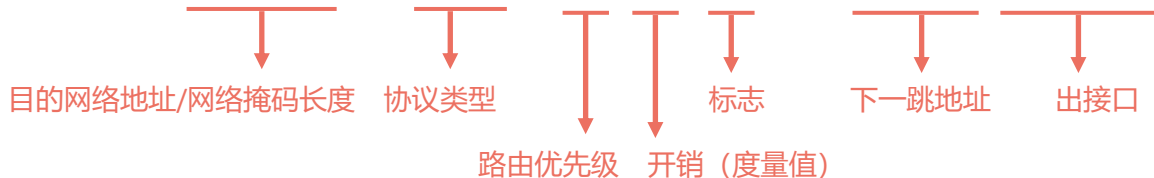
查看IP路由表

```
<Quidway> display ip routing-table  
Route Flags: R - relay, D - download to fib
```

Routing Tables: Public

Destinations : 6 Routes : 6

Destination/Mask	Proto	Pre	Cost	Flags	NextHop	Interface
1.1.1.1/32	Static	60	0	D	0.0.0.0	NULL0
2.2.2.2/32	Static	60	0	D	100.0.0.2	Vlanif100
100.0.0.0/24	Direct	0	0	D	100.0.0.1	Vlanif100
100.0.0.1/32	Direct	0	0	D	127.0.0.1	Vlanif100
127.0.0.0/8	Direct	0	0	D	127.0.0.1	InLoopBack0
127.0.0.1/32	Direct	0	0	D	127.0.0.1	InLoopBack0





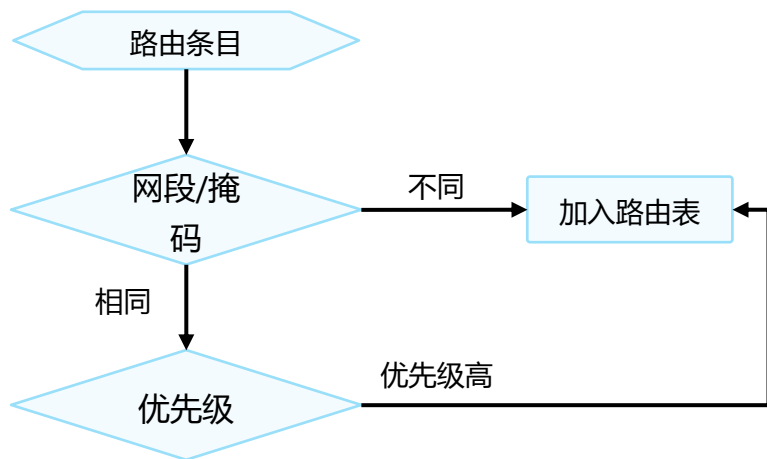
路由表中各个内容的含义

- Destination/Mask: 表示此路由的目的网络地址与网络掩码。将目的地址和子网掩码“逻辑与”后可得到目的主机或路由器所在网段的地址。例如：目的地址为1.1.1.1，掩码为255.255.255.0的主机或路由器所在网段的地址为1.1.1.0。
- Proto (Protocol) : 该路由的协议类型，也即路由器是通过什么协议获知该路由的。
- Pre (Preference) : 表示此路由的路由协议优先级。针对同一目的地，可能存在不同下一跳、出接口等多条路由，这些不同的路由可能是由不同的路由协议发现的，也可以是手工配置的静态路由。优先级最高（数值最小）者将成为当前的最优路由。
- Cost: 路由开销。当到达同一目的地的多条路由具有相同的路由优先级时，路由开销最小的将成为当前的最优路由。
- NextHop: 表示对于本路由器而言，到达该路由指向的目的网络的下一跳地址。该字段指明了数据转发的下一个设备。
- Interface: 表示此路由的出接口。指明数据将从本路由器的哪个接口转发出去。



路由优先级 - 基本概念

优先级比较



- 当路由器从多种不同的途径获知到达同一个目的网段的路由（这些路由的目的网络地址及网络掩码均相同）时，路由器会比较这些路由的优先级，优选优先级值最小的路由。
- 路由来源的优先级值（Preference）越小代表加入路由表的优先级越高。
- 拥有最高优先级的路由将被添加进路由表。

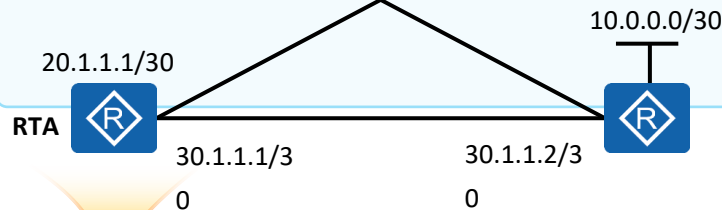


路由优先级 - 比较过程

优先级比较示例

动态路由协议

OSPF



RTA上的路由条目

目的网络/掩码	来源	优先级	下一跳
10.0.0.0/30	静态	60	30.1.1.2
10.0.0.0/30	OSPF	10	20.1.1.2

加入路由表

- RTA通过动态路由协议OSPF和手动配置的方式都发现了到达10.0.0.0/30的路由，此时会比较这两条路由的优先级，优选优先级值最小的路由。
- 每一种路由协议都有相应的优先级。
- OSPF拥有更优的优先级，因此通过OSPF学习到的路由被添加到路由表中。



路由优先级 – 常见默认数值

- 常见路由类型的默认优先级如下：

思科

路由协议	优先级
直连	0
Static	1
EBGP	20
内部EIGRP	90
OSPF	110
IS-IS	115
RIP	120
外部EIGRP	170
IBGP	200
未知	255

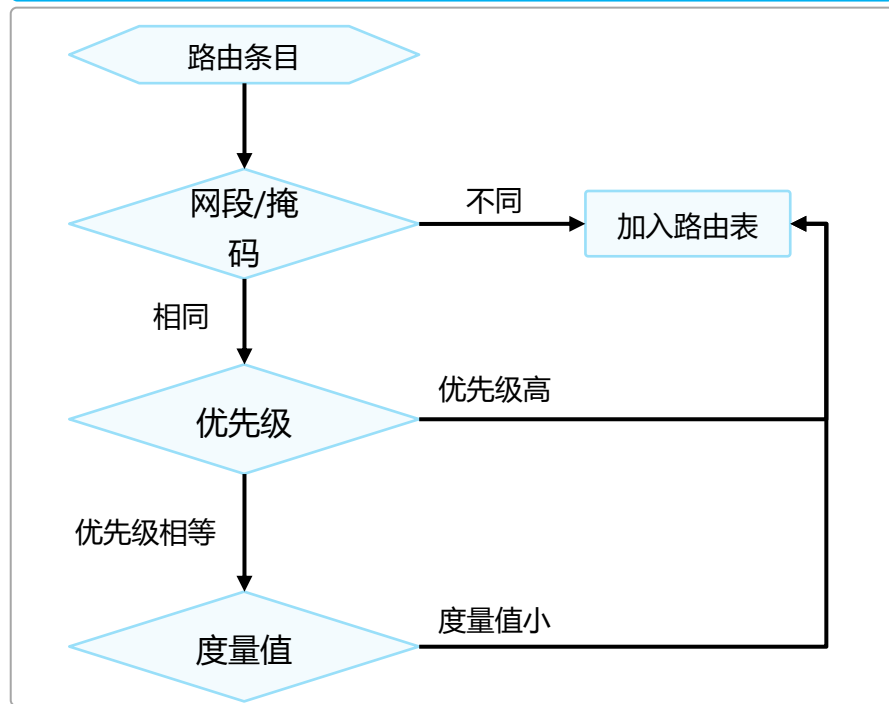
华为

路由协议	外部优先级	内部优先级
Direct	0	0
OSPF	10	10
IS-IS	15	15 (Level-1) 18 (Level-2)
Static	60	60
RIP	100	100
ospf ASE	150	150
IBGP	255	200
EBGP	255	20



度量值 - 基本概念

度量值比较



- 当路由器通过某种路由协议发现了多条到达同一个目的网络的路由时（拥有相同的路由优先级），度量值将作为路由优选的依据之一。
- 路由度量值表示到达这条路由所指目的地址的代价。
- 一些常用的度量值有：跳数、带宽、时延、代价、负载、可靠性等。
- 度量值数值越小越优先，度量值最小路由将会被添加到路由表中。
- 度量值很多时候被称为开销（Cost）。



最长匹配原则

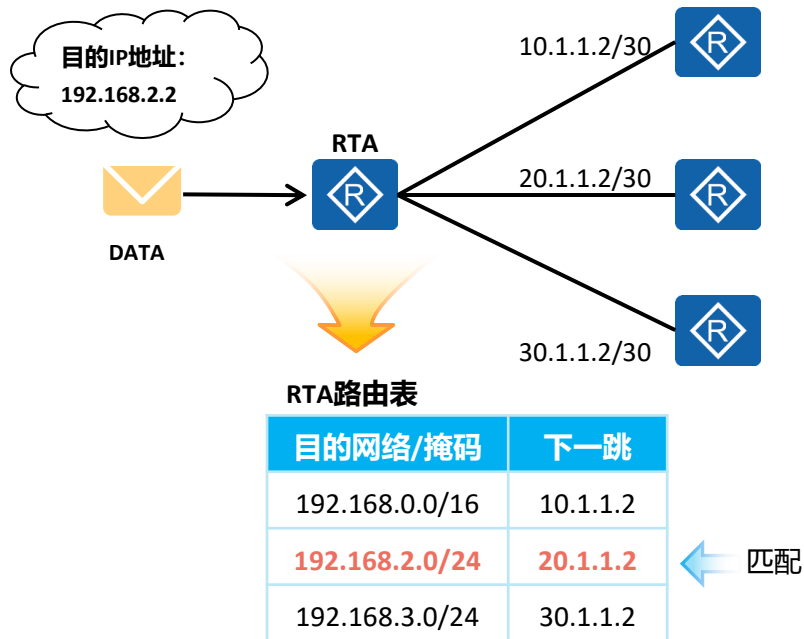
- 当路由器收到一个IP数据包时，会将数据包的目的IP地址与自己本地路由表中的所有路由表项进行逐位（Bit-By-Bit）比对，直到找到匹配度最长的条目，这就是最长前缀匹配机制。

		Bit By Bit 逐位匹配			
数据包目的IP		172.16.2.1			
		172.	16.	00000010	00000001
路由条目1	172.16.1.0 255.255.255.0	172.	16.	00000001	xxxxxx 不匹配
路由条目2	172.16.2.0 255.255.255.0	172.	16.	00000010	xxxxxx 胜利
路由条目3	172.16.0.0 255.255.0.0	172.	16.	xxxxxxx	xxxxxx 不是最长



最长匹配示例 (1)

最长匹配示例

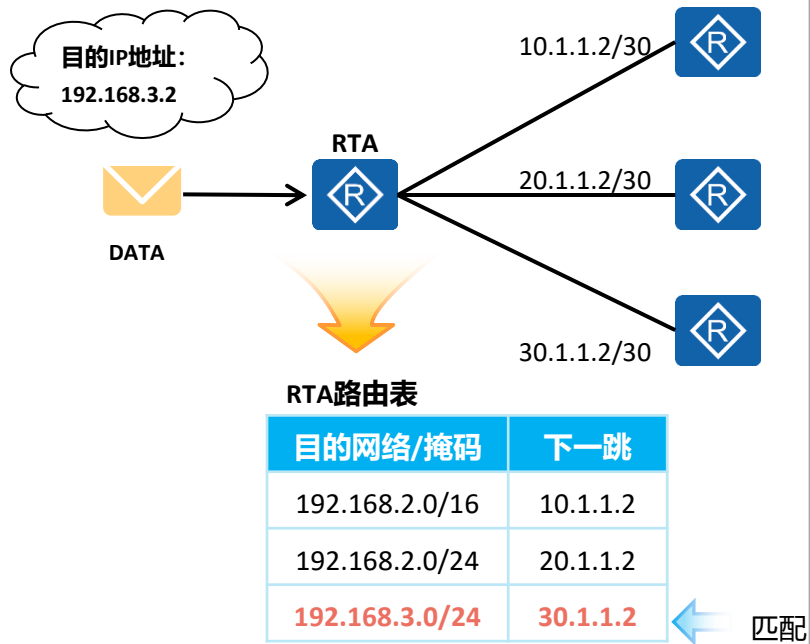


- 根据最长匹配原则进行匹配，能够匹配192.168.2.2的路由存在两条，但是路由的掩码长度中，一个为16 bit，另一个为24 bit，掩码长度为24 bit的路由满足最长匹配原则，因此被选择来指导发往192.168.2.2的报文转发。



最长匹配示例 (2)

最长匹配示例



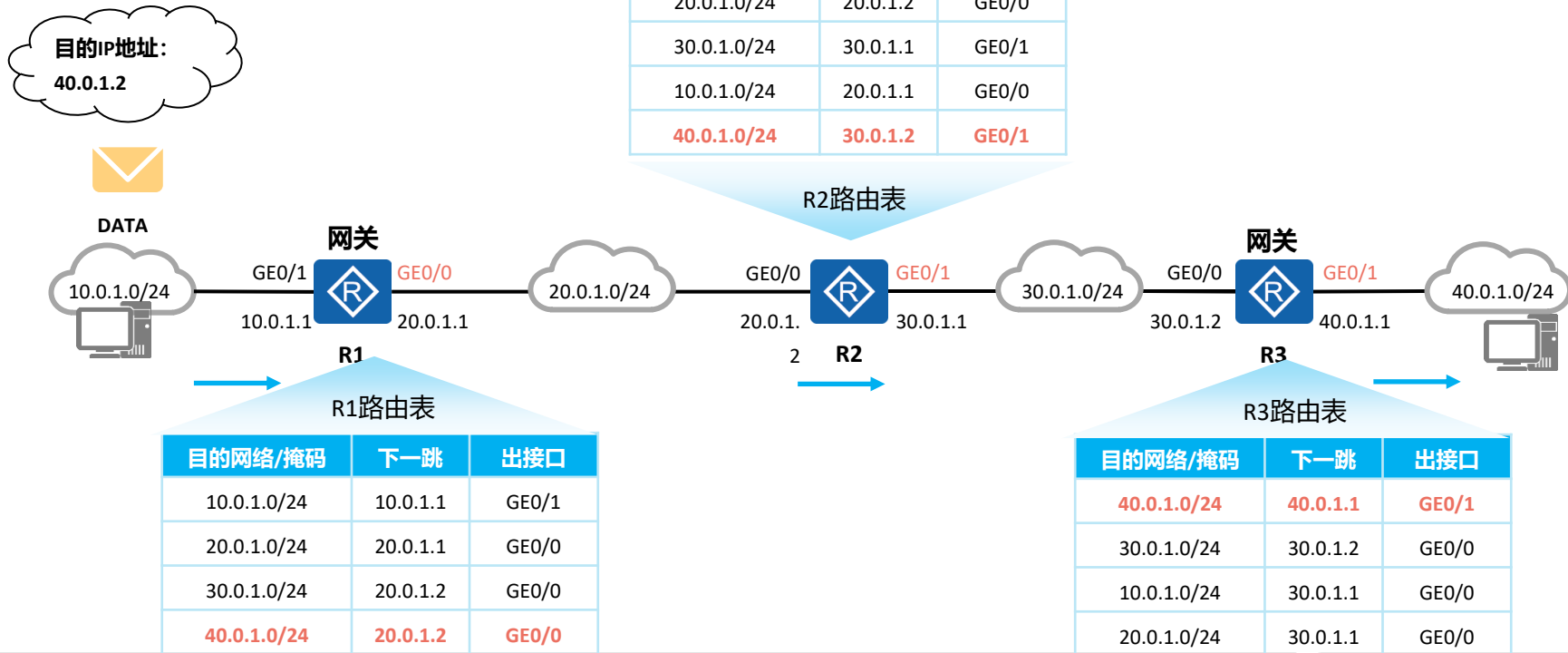
- 根据最长匹配原则匹配，能够匹配到192.168.3.2的路由只有一条，此路由为最终转发依据。



路由转发流程

目的网络/掩码	下一跳	出接口
20.0.1.0/24	20.0.1.2	GE0/0
30.0.1.0/24	30.0.1.1	GE0/1
10.0.1.0/24	20.0.1.1	GE0/0
40.0.1.0/24	30.0.1.2	GE0/1

R2路由表





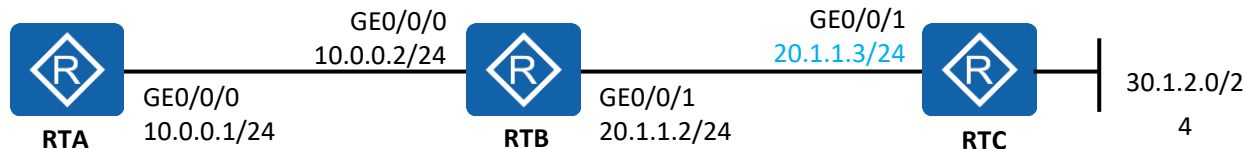
IP路由表小结

- 当路由器从多种不同的途径获知到达同一个目的网段的路由（这些路由的目的网络地址及网络掩码均相同）时，会选择路由优先级值最小的路由；如果这些路由学习自相同的路由协议，则优选度量值最优的。总之，最优的路由加入路由表。
- 当路由器收到一个数据包时，会在自己的路由表中查询数据包的目的IP地址。如果能够找到匹配的路由表项，则依据表项所指示的出接口及下一跳来转发数据；如果没有匹配的表项，则丢弃该数据包。
- 路由器的行为是逐跳的，数据包从源到目的地沿路径每个路由器都必须有关于目标网段的路由，否则就会造成丢包。
- 数据通信往往是双向的，因此要关注流量的往返（往返路由）。



路由递归 (1)

- 路由必须有直连的下一跳才能够指导转发，但是路由生成时下一跳可能不是直连的，因此需要计算出一个直连的下一跳和对应的出接口，这个过程就叫做路由递归。
- 路由递归也被称为路由迭代。

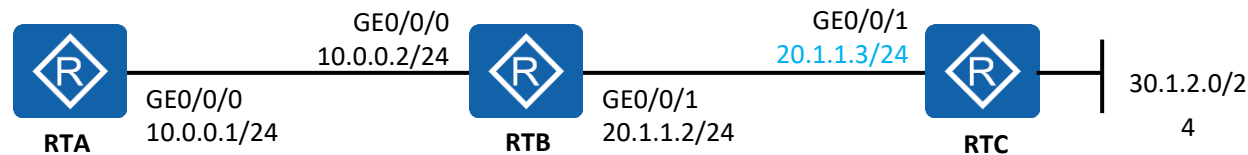


[RTA] ip route-static 30.1.2.0 24 20.1.1.3

去往30.1.2.0/24的路由，下一跳为20.1.1.3，非本地直连网络，如果路由表中没有去往20.1.1.3的路由，该静态路由将不会生效，无法作为有效路由条目，并不会出现在路由表。



路由递归 (2)



[RTA] ip route-static 30.1.2.0 24 20.1.1.3

递归

[RTA] ip route-static 20.1.1.0 24 10.0.0.2

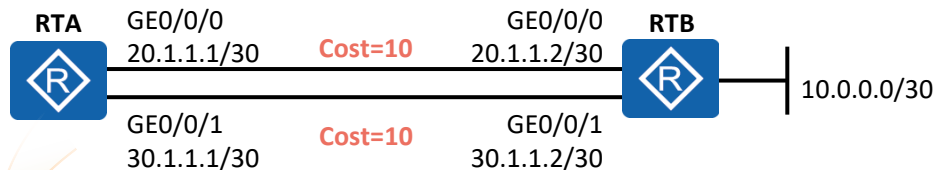
目的网络	下一跳	出接口
30.1.2.0/24	20.1.1.3	GE0/0/0
20.1.1.0/24	10.0.0.2	GE0/0/0

添加一条去往20.1.1.3的路由，下一跳为直连网络内的IP地址10.0.0.2。
去往30.1.2.0/24的路由通过递归查询得到一个直连的下一跳，该路由因此生效。



等价路由

等价路由



RTA路由表

目的网络/掩码	下一跳
10.0.0.0/30	20.1.1.2
	30.1.1.2

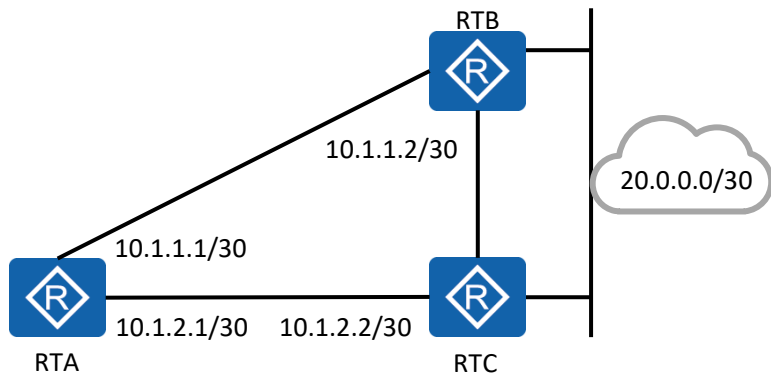
来源相同、开销相同的路由都会被加入路由表，形成的路由为等价路由（两个路由条目指向的目的网段相同，但是具有不同的下一跳地址），路由转发会将流量分布到多条路径上。

- 路由表中存在等价路由之后，前往该目的网段的IP报文路由器会通过所有有效的接口、下一跳转发，这种转发行为被称为负载分担。



浮动路由 – 基本概念

浮动路由



RTA上配置浮动路由

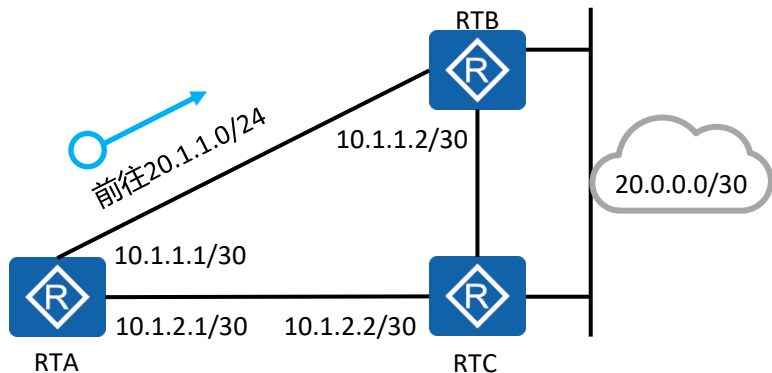
```
[RTA] ip route-static 20.0.0.0 30 10.1.1.2  
[RTA] ip route-static 20.0.0.0 30 10.1.2.2 preference 70
```

- 静态路由支持配置时手动指定优先级，可以通过配置目的地址/掩码相同、优先级不同、下一跳不同的静态路由，实现转发路径的备份。
- 浮动路由是主用路由的备份，保证链路故障时提供备份路由。主用路由下一跳可达时该备份路由不会出现在路由表。



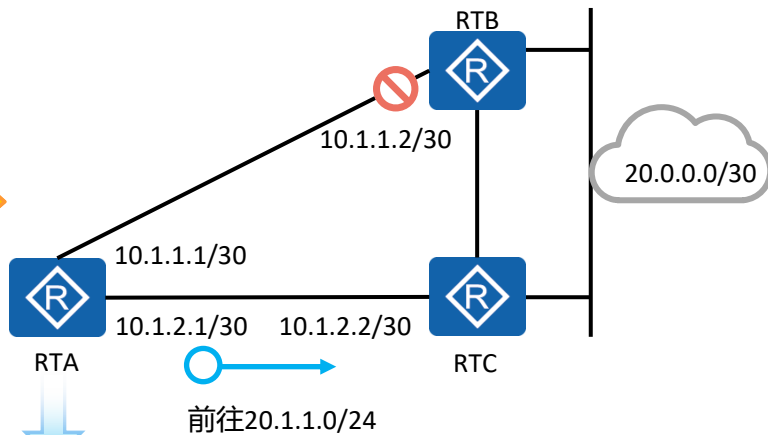
浮动路由 - 示例

浮动路由切换



主链路正常时RTA上路由表

目的网络	下一跳	优先级
20.0.0.0	10.1.1.2	60



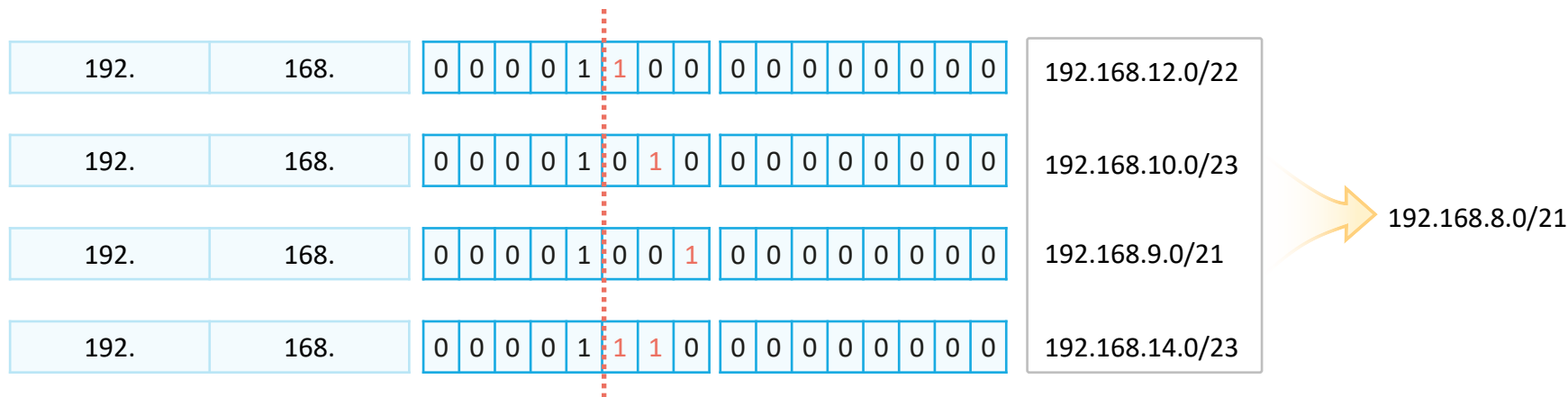
主链路故障时RTA上路由表

目的网络	下一跳	优先级
20.0.0.0	10.1.2.2	70



CIDR

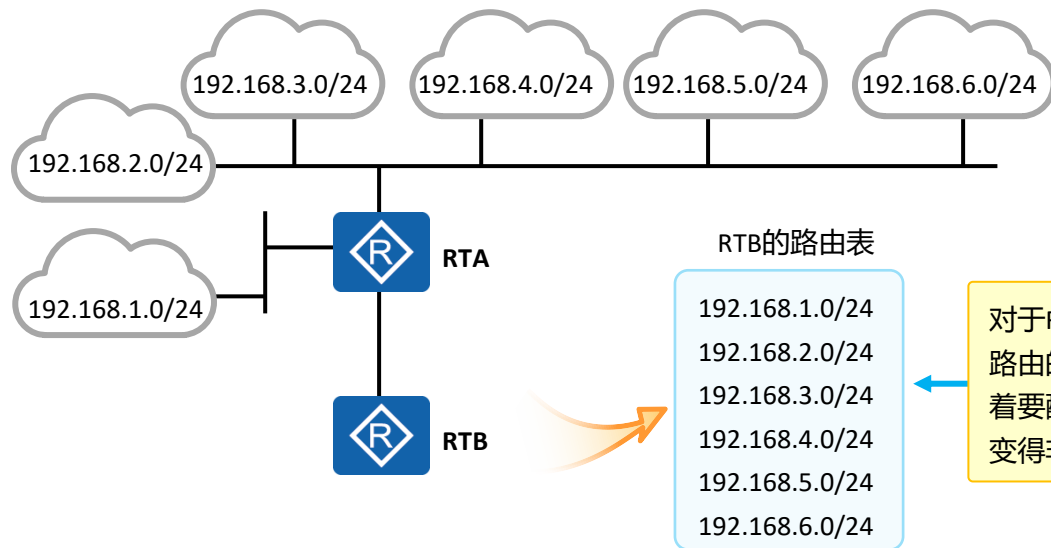
- CIDR (classless inter-domain routing, 无类别域间路由) 采用IP地址加掩码长度来标识网络和子网, 而不是按照传统A、B、C等类型对网络地址进行划分。
- CIDR容许任意长度的掩码长度, 将IP地址看成连续的地址空间, 可以使用任意长度的前缀分配, 多个连续的前缀可以聚合成一个网络, 该特性可以有效减少路由表条目数量。





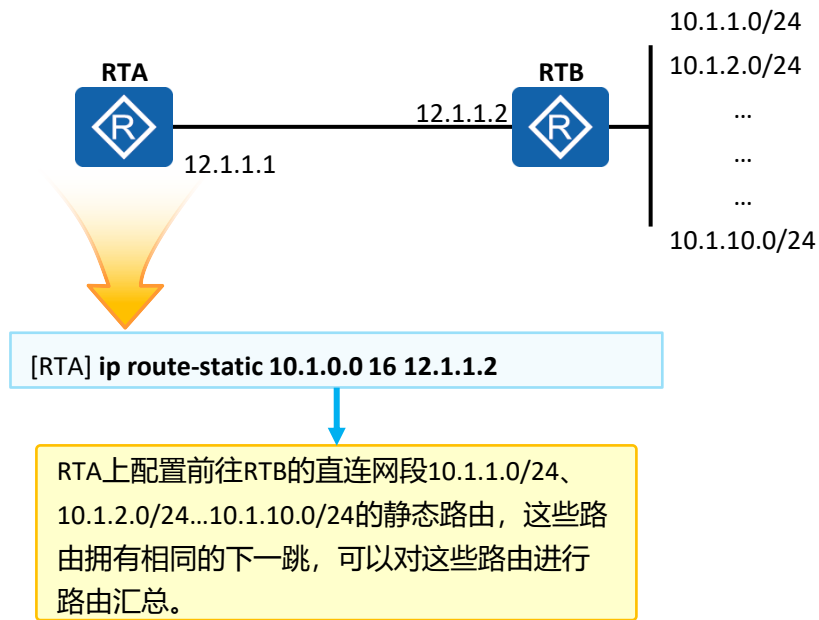
路由汇总需求

- 子网划分、VLSM解决了地址空间浪费的问题，但同时也带了新的问题：路由表中的路由条目数量增加。
- 为减少路由条目数量可以使用路由汇总。





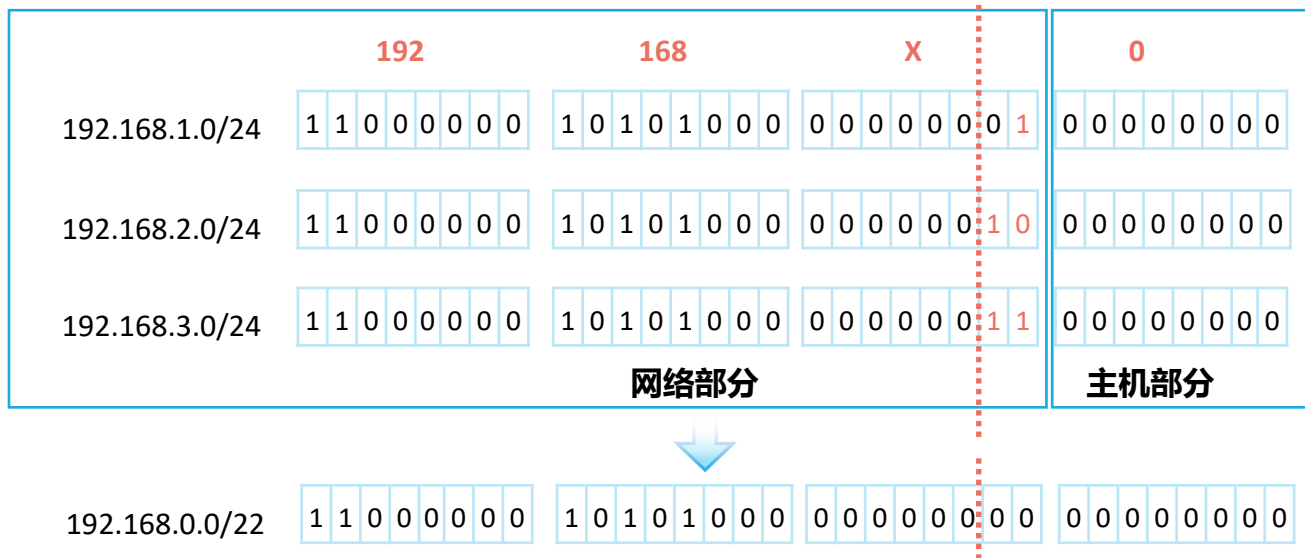
路由汇总简介



- 路由汇总将一组具有相同前缀的路由汇聚成一条路由，从而达到减小路由表规模以及优化设备资源利用率的目的。
- 路由汇总采用了CIDR的思想：将相同前缀的地址聚合成一个。
- 我们把汇聚之前的这组路由称为精细路由或明细路由，把汇聚之后的这条路由称为汇总路由或聚合路由。



汇总计算



- 基于一系列连续的、有规律的IP网段，如果需计算相应的汇总路由，且确保得出的汇总路由刚好“囊括”上述IP网段，则需保证汇总路由的掩码长度尽可能长。
- 诀窍在于：将明细路由的目的网络地址都换算成二进制，然后排列起来，找出所有目的网络地址中“相同的比特位”。



汇总引发的问题 (1)

路由汇总带来的环路问题

[RTB] ip route-static 0.0.0.0 0 12.1.1.2

[RTA] ip route-static 10.1.0.0 16 12.1.1.1

10.1.1.0/24

10.1.2.0/24

...

...

...

10.1.10.0/24



RTB

1

12.1.1.1

RTB收到前往10.1.20.0/24的流量，流量匹配默认路由，转发给RTA

12.1.1.2



RTA

2

Internet

RTA上对路由进行了汇总，因此该流量匹配汇总路由10.1.0.0/16，被转发给RTB

环路

3



汇总引发的问题 (2)

路由汇总带来的环路问题 – 解决方案

```
[RTB] ip route-static 0.0.0.0 0 12.1.1.2
```

```
[RTA] ip route-static 10.1.0.0 16 12.1.1.1
```

10.1.1.0/24

10.1.2.0/24

...

...

...

10.1.10.0/24

RTB



12.1.1.1

12.1.1.2

RTA



Internet

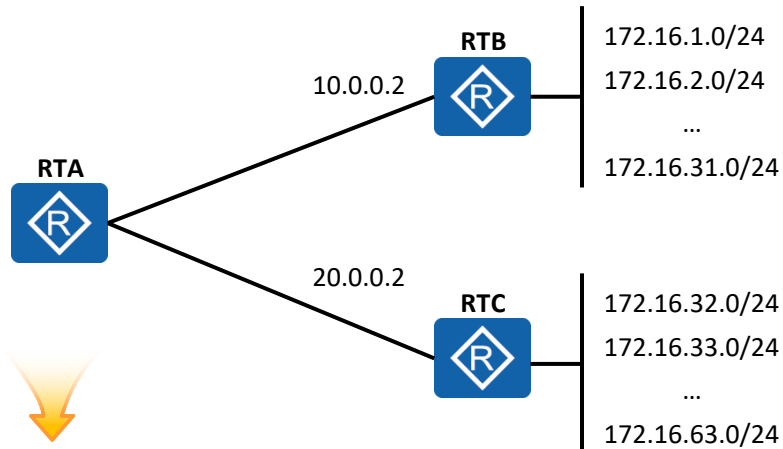
```
[RTB] ip route-static 10.1.0.0 16 0 NULL0
```

- 在RTB上增加一条指向Null0的路由，即可解决上述问题。因此，在部署路由汇总的时候要格外注意，要规避环路问题。



精确汇总 (1)

精确进行路由汇总



```
[RTA] ip route-static 172.16.0 16 10.0.0.2
```

- 为了让RTA能够到达RTB上的172.16.1.0/24-172.16.31.0/24网段，配置了一条静态的汇总路由，这条网段虽然优化了网络配置，但是汇总的范围太广，将RTC上的网段也包括在内，导致前往RTC上网段的流量到达RTA之后会被发往RTB，造成数据包的丢失，这种路由为不精确的路由。为此配置汇总路由时要尽量精确，刚好包括所有明细路由。



精确汇总 (2)

精确进行路由汇总

	10	1	0	0
10.1.1.0/24	0 0 0 0 1 0 1 0	0 0 0 0 0 0 0 1	0 0 0 0 0 0 0 1	0 0 0 0 0 0 0 0
10.1.2.0/24	0 0 0 0 1 0 1 0	0 0 0 0 0 0 0 1	0 0 0 0 0 0 1 0	0 0 0 0 0 0 0 0
10.1.3.0/24	0 0 0 0 1 0 1 0	0 0 0 0 0 0 0 1	0 0 0 0 0 0 1 1	0 0 0 0 0 0 0 0

/22

```
ip route-static 10.1.1.0 24 12.1.1.2  
ip route-static 10.1.2.0 24 12.1.1.2  
ip route-static 10.1.3.0 24 12.1.1.2
```



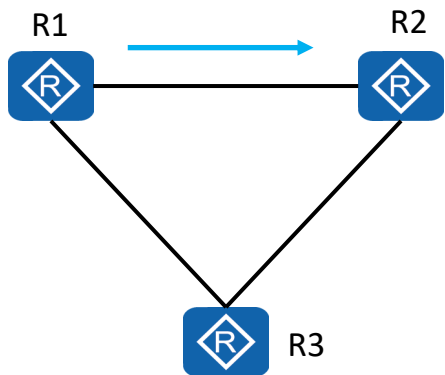
```
ip route-static 10.1.1.0 22 12.1.1.2
```

精确计算汇总后的网络号、掩码，避免汇总后掩码过小。



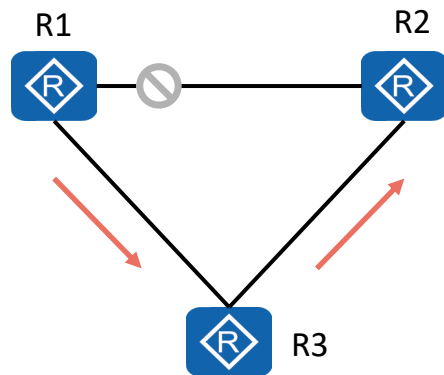
为什么需要动态路由协议？

- 静态路由是由工程师手动配置和维护的路由条目，命令行简单明确，适用于小型或稳定的网络。静态路由有以下问题：
 - 无法适应规模较大的网络：随着设备数量增加，配置量急剧增加。
 - 无法动态响应网络变化：网络发生变化，无法自动收敛网络，需要工程师手动修改。



R1-R2静态路由

链路故障



手动配置R1-R3-R2静态路由



动态路由分类

按工作区域分类

IGP (Interior Gateway Protocols, 内部网关协议)

EIGRP

OSPF

IS-IS

EGP (Exterior Gateway Protocols, 外部网关协议)

BGP

按工作机制及算法分类

(Distance Vector Routing Protocols, 距离矢量路由协议)

EIGRP

(Link-State Routing Protocols, 链路状态路由协议)

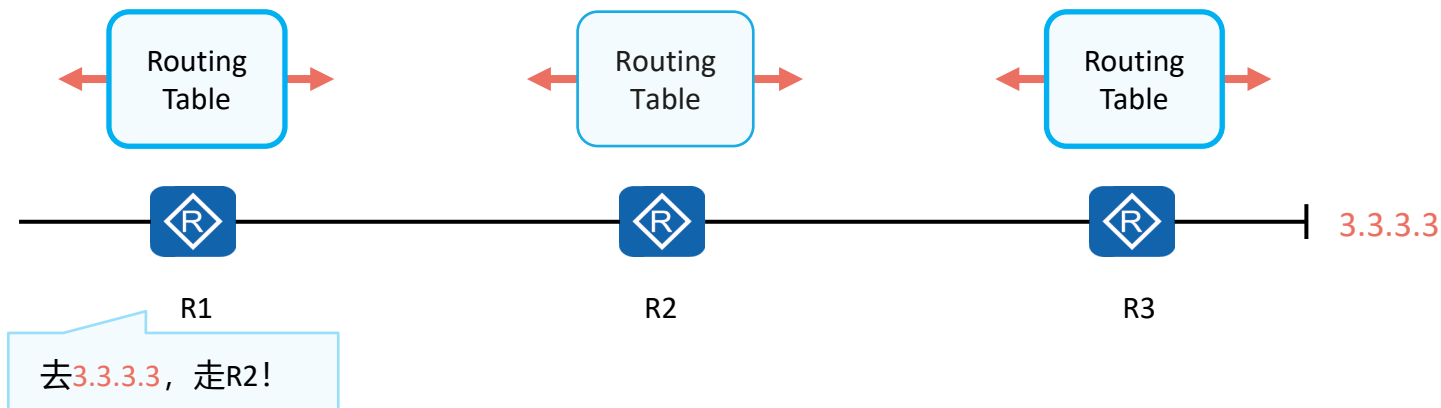
OSPF

IS-IS



距离矢量路由协议

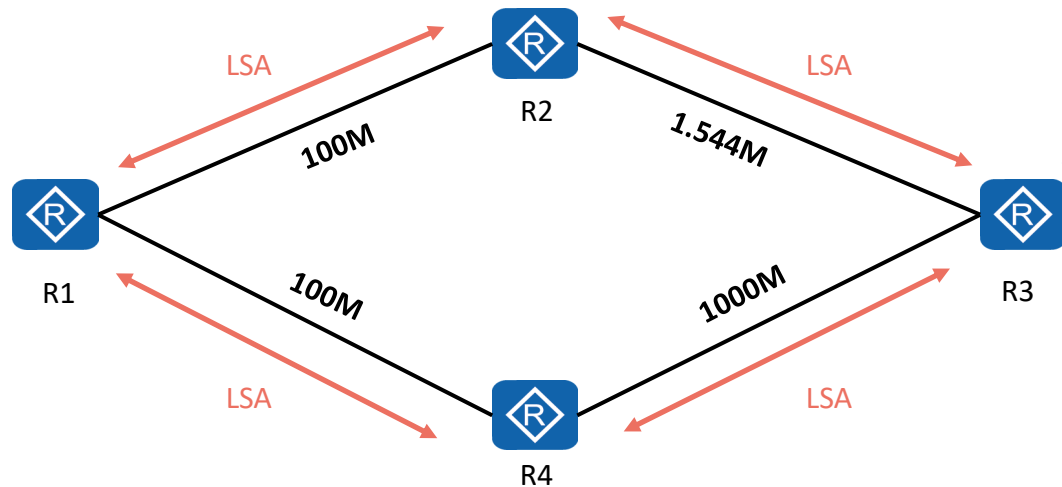
- 运行距离矢量路由协议的路由器周期性的泛洪自己的路由表。通过路由的交互，每台路由器都从相邻的路由器学习到路由，并且加载进自己的路由表中。
- 对于网络中的所有路由器而言，路由器并不清楚网络的拓扑，只是简单的知道要去往某个目的方向在哪里，距离有多远。这即是距离矢量算法的本质。





链路状态路由协议 - LSA泛洪

- 与距离矢量路由协议不同，链路状态路由协议通告的是链路状态而不是路由表。运行链路状态路由协议的路由器之间首先会建立一个协议的邻居关系，然后彼此之间开始交互LSA（Link State Advertisement，链路状态通告）。

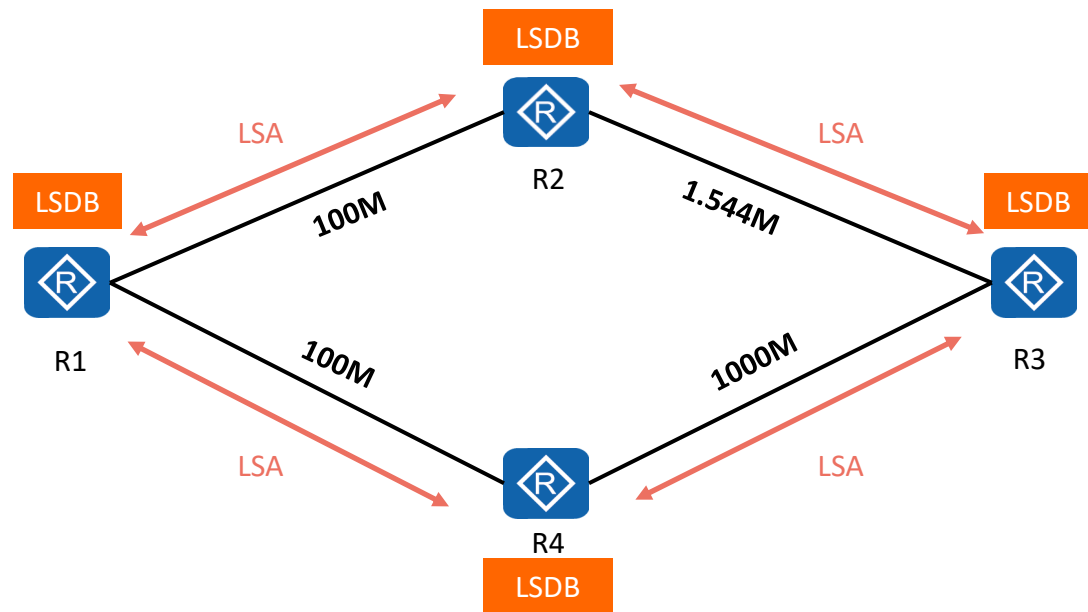


- 不再通告路由信息，而是LSA。
- LSA描述了路由器接口的状态信息，例如接口的开销、连接的对象等。



链路状态路由协议 - LSDB组建

- 每台路由器都会产生LSAs，路由器将接收到的LSAs放入自己的LSDB（Link State DataBase，链路状态数据库）。路由器通过LSDB，掌握了全网的拓扑。

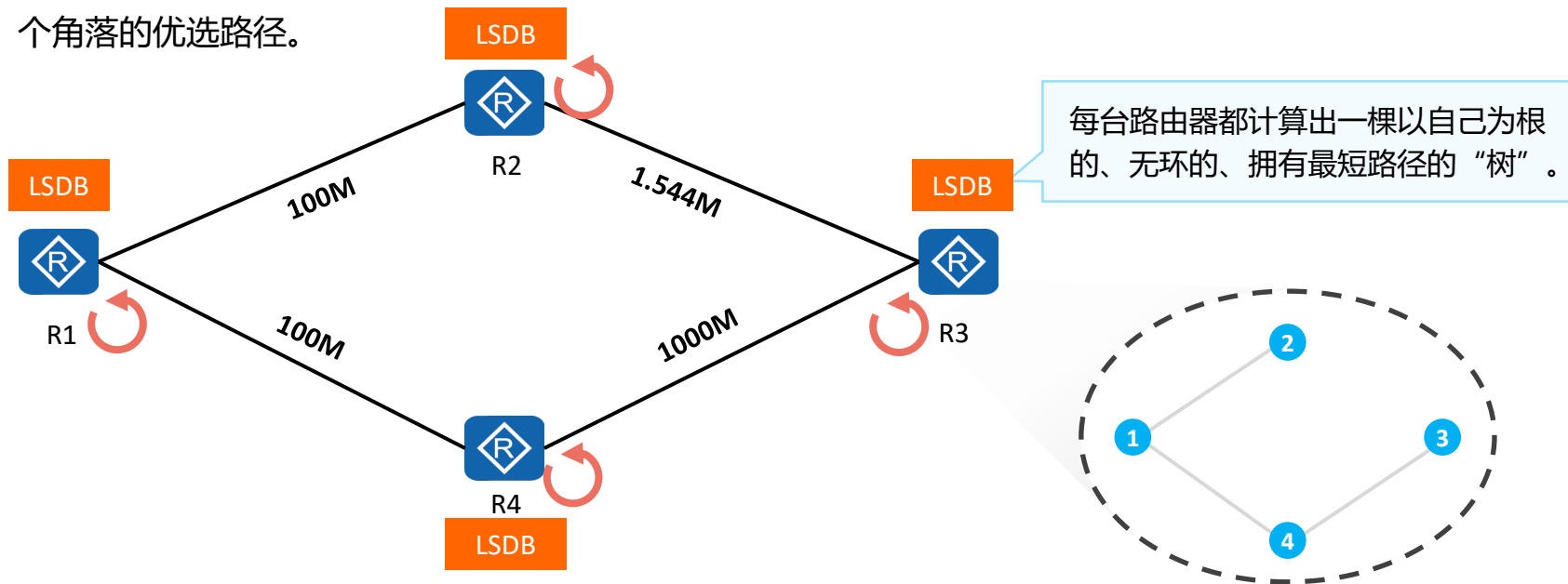


- 路由器将LSA存放在LSDB中
- LSDB汇总了网络中路由器对于自己接口的描述
- LSDB包含全网拓扑的描述



链路状态路由协议 - SPF计算

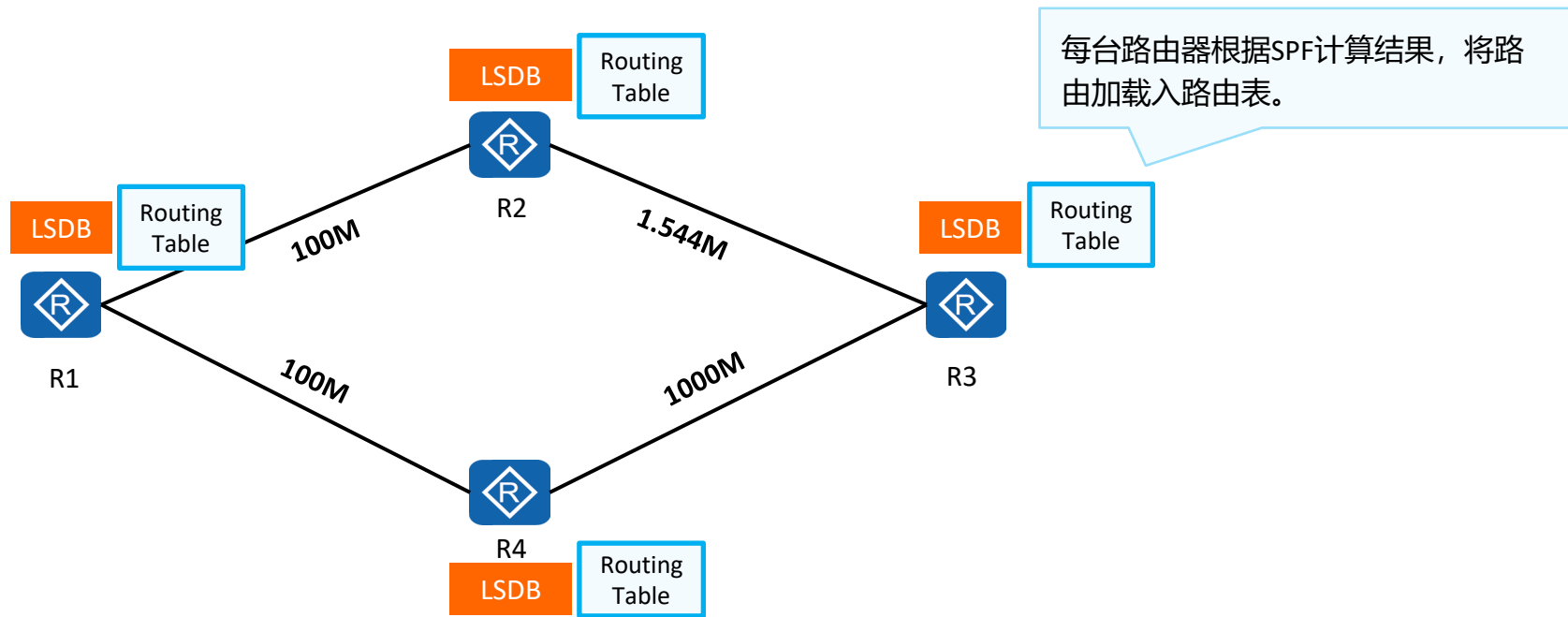
- 每台路由器基于LSDB，使用SPF（Shortest Path First，最短路径优先）算法进行计算。每台路由器都计算出一棵以自己为根的、无环的、拥有最短路径的“树”。有了这棵“树”，路由器就已经知道了到达网络各个角落的优选路径。





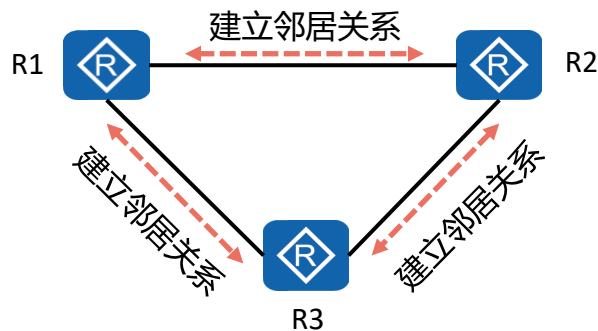
链路状态路由协议 - 路由表生成

- 最后，路由器将计算出来的优选路径，加载进自己的路由表（Routing Table）。



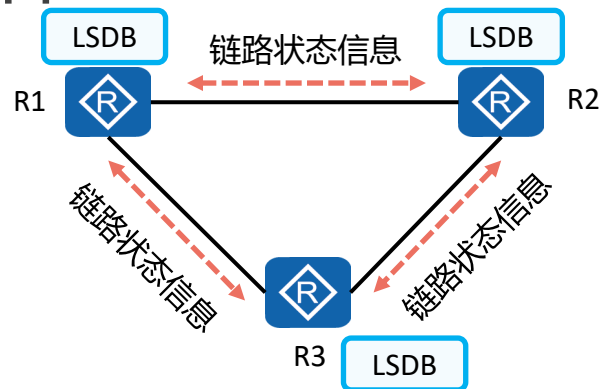


链路状态路由协议总结



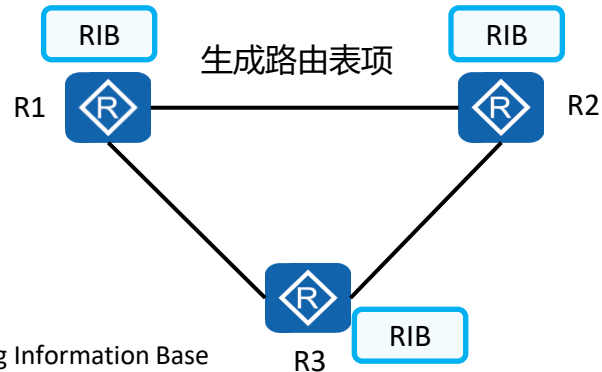
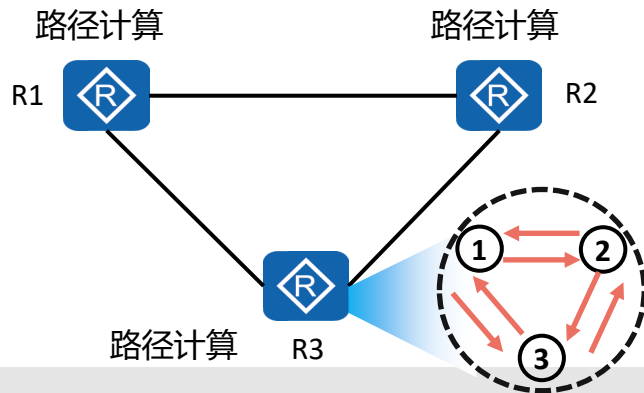
1

2



3

4



RIB: Routing Information Base

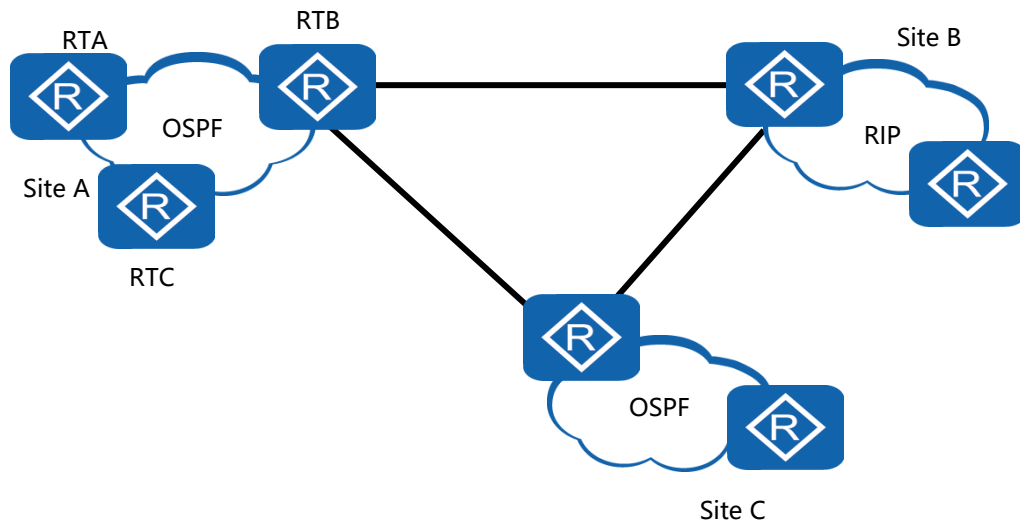


OSPF简介

- OSPF是典型的链路状态路由协议，是目前业内使用非常广泛的IGP协议之一。
- 目前针对IPv4协议使用的是OSPF Version 2（RFC2328）；针对IPv6协议使用OSPF Version 3（RFC2740）。
- 运行OSPF路由器之间交互的是LS（Link State，链路状态）信息，而不是直接交互路由。LS信息是OSPF能够正常进行拓扑及路由计算的关键信息。
- OSPF路由器将网络中的LS信息收集起来，存储在LSDB中。路由器都清楚区域内的网络拓扑结构，这有助于路由器计算无环路径。
- 每台OSPF路由器都采用SPF算法计算达到目的地的最短路径。路由器依据这些路径形成路由加载到路由表中。
- OSPF支持VLSM（Variable Length Subnet Mask，可变长子网掩码），支持手工路由汇总。
- 多区域的设计使得OSPF能够支持更大规模的网络。



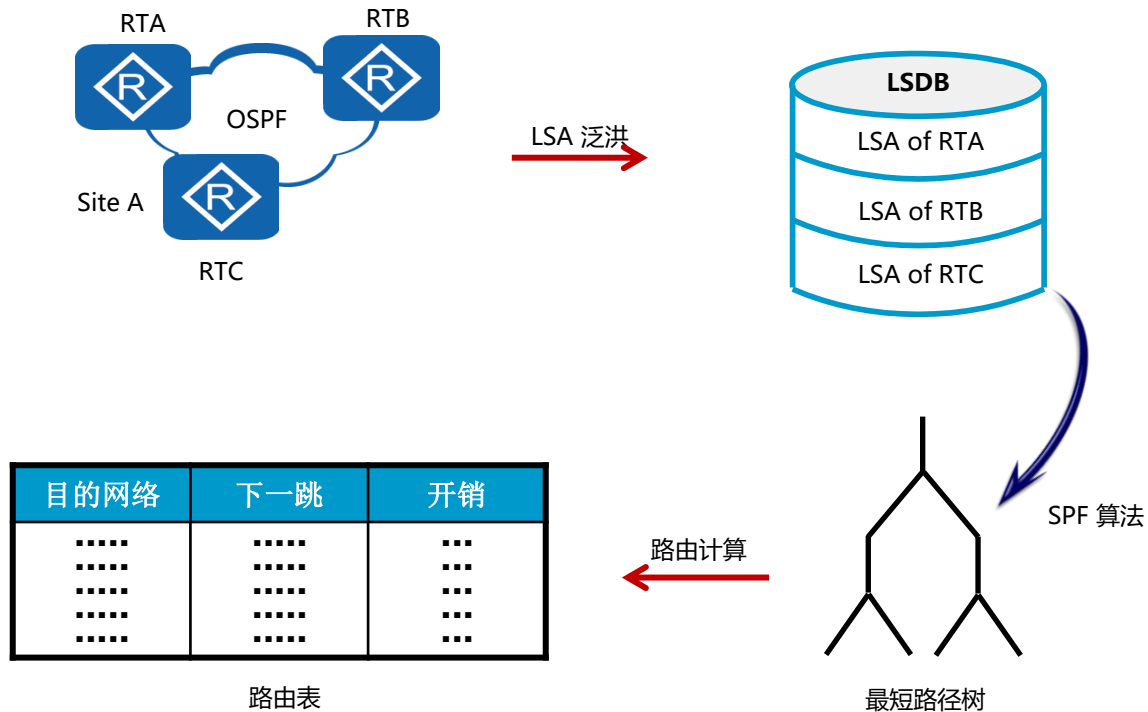
开放式最短路径优先（OSPF）



- 无环路
- 收敛快
- 扩展性好
- 支持认证



OSPF原理介绍





OSPF报文

IP Header

OSPF Protocol Packet

- OSPF报文封装在IP报文中，协议号为89。
- OSPF报文类型有5种：
 - Hello 报文
 - DD（Database Description）报文
 - LSR（Link State Request）报文
 - LSU（Link State Update）报文
 - LSACK（Link State Acknowledgment）报文



OSPF三大表项 - 邻居表

- OSPF有三张重要的表项，OSPF邻居表、LSDB表和OSPF路由表。对于OSPF的邻居表，需要了解：
 - OSPF在传递链路状态信息之前，需先建立OSPF邻居关系。
 - OSPF的邻居关系通过交互Hello报文建立。
 - OSPF邻居表显示了OSPF路由器之间的邻居状态，使用display ospf peer查看。

[R1]display ospf peer

Router ID:1.1.1.1



R1

GE1/0/0

10.1.1.1/30

Router ID:2.2.2.2



R2

GE1/0/0

10.1.1.2/30

<R1> display ospf peer

OSPF Process 1 with Router ID 1.1.1.1

Neighbors

Area 0.0.0.0 interface 10.1.1.1(GigabitEthernet1/0/0)'s neighbors

Router ID: 2.2.2.2 Address: 10.1.1.2 GR State: Normal

State: Full Mode:Nbr is Master Priority: 1

DR: 10.1.1.1 BDR: 10.1.1.2 MTU: 0

Dead timer due in 35 sec

Retrans timer interval: 5

Neighbor is up for 00:00:05

Authentication Sequence: [0]



OSPF三大表项 - LSDB表

- 对于OSPF的LSDB表，需要了解：
 - LSDB会保存自己产生的及从邻居收到的LSA信息，本例中R1的LSDB包含了三条LSA。
 - Type标识LSA的类型， AdvRouter标识发送LSA的路由器。
 - 使用命令行display ospf lsdb查看LSDB表。

[R1]display ospf lsdb

Router ID:1.1.1.1



R1

GE1/0/0

10.1.1.1/30

Router ID:2.2.2.2



R2

GE1/0/0

10.1.1.2/30

<R1> display ospf lsdb

OSPF Process 1 with Router ID 1.1.1.1

Link State Database

Area: 0.0.0.0

Type	LinkState ID	AdvRouter	Age	Len	Sequence	Metric
Router	2.2.2.2	2.2.2.2	98	36	8000000B	1
Router	1.1.1.1	1.1.1.1	92	36	80000005	1
Network	10.1.1.2	2.2.2.2	98	32	80000004	0



OSPF三大表项 - OSPF路由表

- 对于OSPF的路由表，需要了解：
 - OSPF路由表和路由器路由表是两张不同的表项。本例中OSPF路由表有三条路由。
 - OSPF路由表包含Destination、Cost和NextHop等指导转发的信息。
 - 使用命令display ospf routing查看OSPF路由表。

[R1]display ospf routing

Router ID:1.1.1.1



R1

GE1/0/0

10.1.1.1/30

Router ID:2.2.2.2



R2

GE1/0/0

10.1.1.2/30

<R1> display ospf routing

OSPF Process 1 with Router ID 1.1.1.1

Routing Tables

Routing for Network

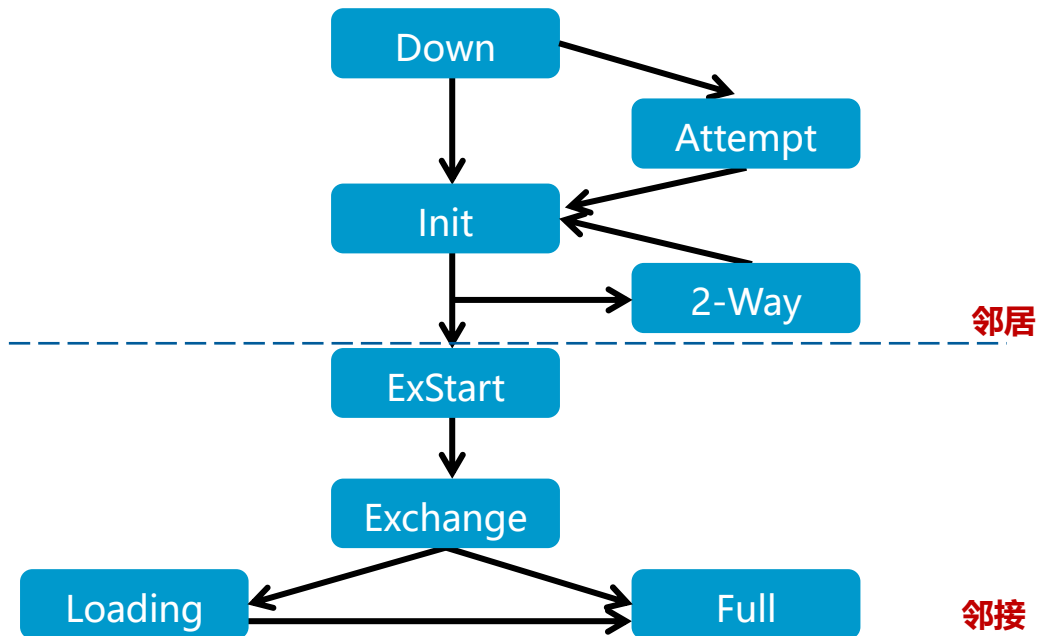
Destination	Cost	Type	NextHop	AdvRouter	Area
1.1.1.1/32	0	stub	1.1.1.1	1.1.1.1	0.0.0.0
10.1.1.0/20	1	Transit	10.1.1.1	1.1.1.1	0.0.0.0
2.2.2.2/32	1	stub	10.1.1.2	2.2.2.2	0.0.0.0

Total Nets: 3

Intra Area: 3 Inter Area: 0 ASE: 0 NSSA: 0



邻居状态机



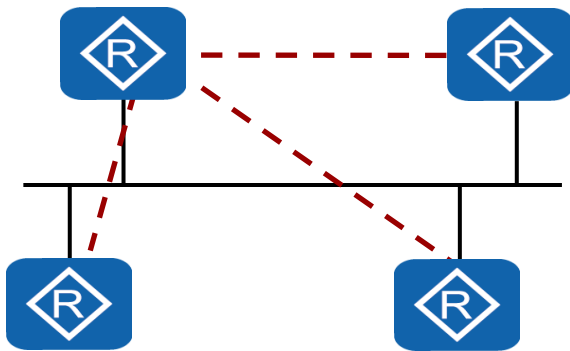


Router ID、邻居和邻接

我有三个邻居

RTA (Router ID=1.1.1.1)

RTB (Router ID=2.2.2.2)



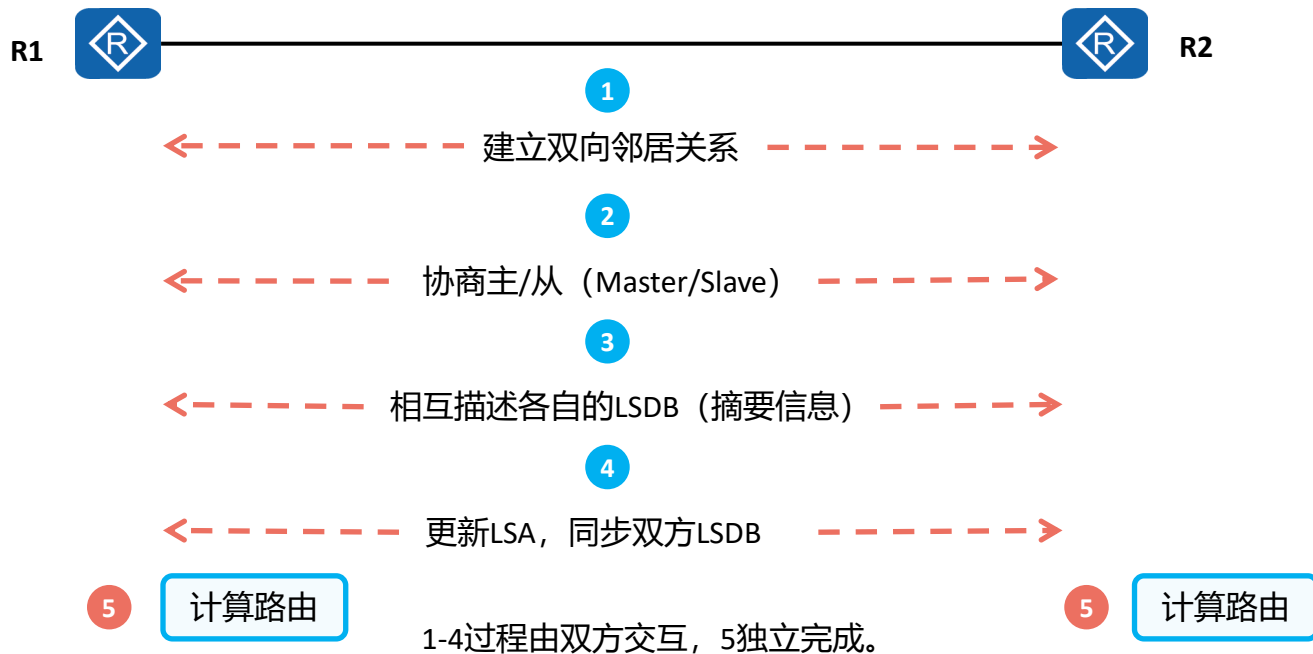
RTC (Router ID=3.3.3.3)

RTD (Router ID=4.4.4.4)



初识OSPF邻接关系建立过程

- OSPF完成邻接关系的建立有四个步骤，建立邻居关系、协商主/从、交互LSDB信息，同步LSDB。





OSPF邻接关系建立流程 - 1

Router ID:1.1.1.1

Router ID:2.2.2.2

R1



R2



Hello报文
我是1.1.1.1, 我还不知道链路上有谁

Init

发现R1 (1.1.1.1) 了, 将它添加到我的邻居表。邻居表中R1状态为Init。

Hello报文
我是2.2.2.2, 我发现了邻居1.1.1.1

2-way

发现R2 (2.2.2.2) 了, 将它添加到我的邻居表。由于R2发现我了, 所以邻居表中R2的状态为2-Way。

Hello报文
我是1.1.1.1, 我发现了邻居2.2.2.2

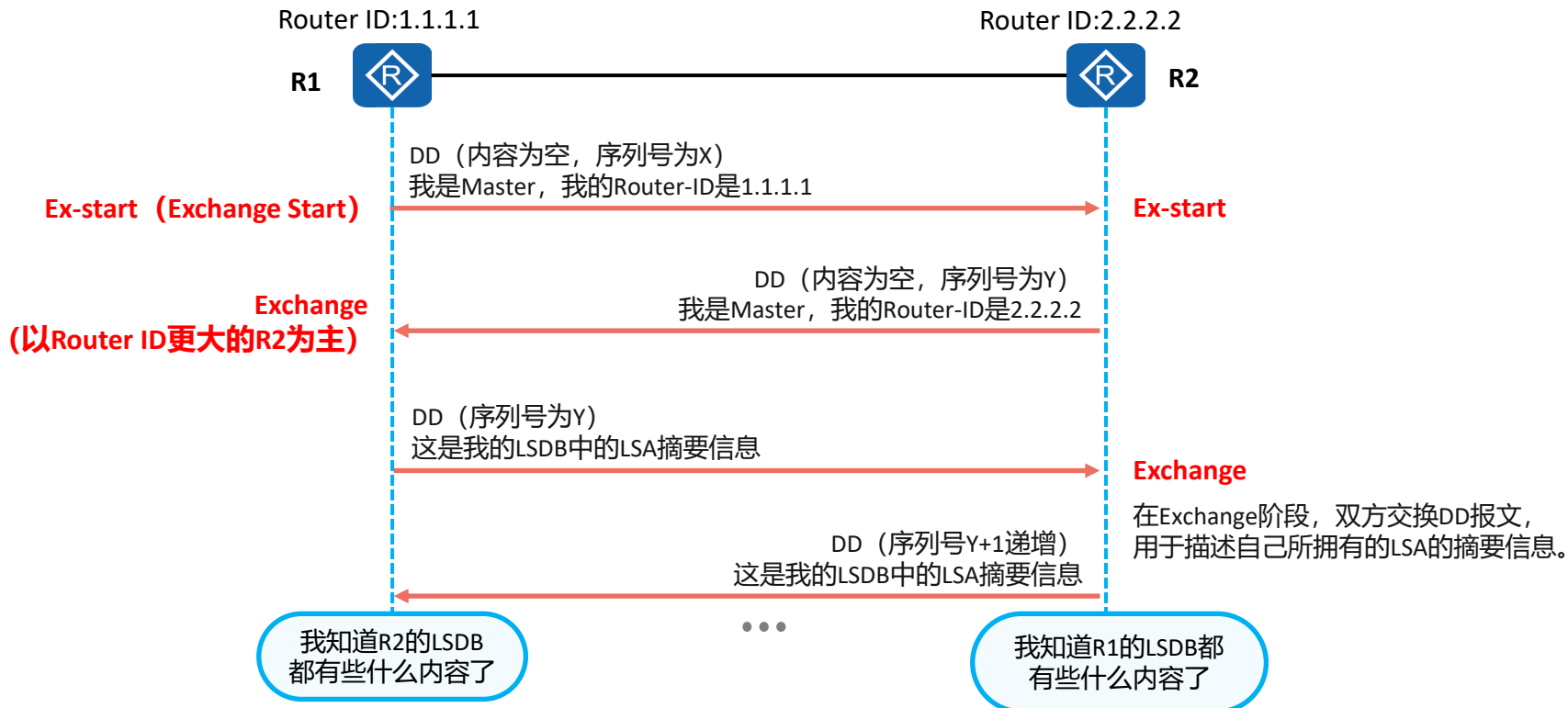
2-way

R1发现我了, 我在邻居表中将1.1.1.1的状态切换到2-Way。

我们是邻居了



OSPF邻接关系建立流程 - 2&3





OSPF邻接关系建立流程 - 4

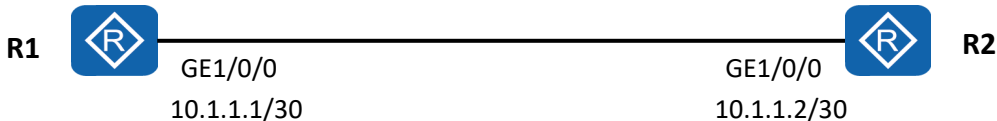




OSPF邻居表回顾

Router ID:1.1.1.1

Router ID:2.2.2.2



```
<R1> display ospf peer
OSPF Process 1 with Router ID 1.1.1.1
Neighbors
Area 0.0.0.0 interface 10.1.1.1(GigabitEthernet1/0/0)'s neighbors
Router ID: 2.2.2.2 Address: 10.1.1.2 GR State: Normal
State: Full Mode:Nbr is Master Priority: 1
DR: 10.1.1.1 BDR: 10.1.1.2 MTU: 0
Dead timer due in 35 sec
Retrans timer interval: 5
Neighbor is up for 00:00:05
Authentication Sequence: [ 0 ]
```

邻居的Router-ID为
2.2.2.2

邻居的状态为Full

R1在GE1/0/0接口上,
在Area0中发现了邻居

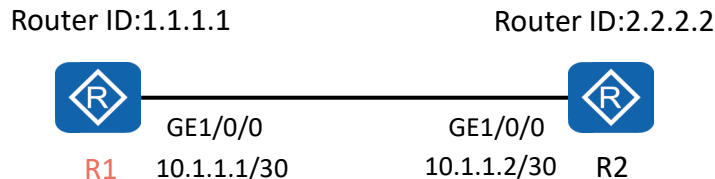
邻居2.2.2.2是Master

思考: 邻居表中的
DR/BDR是什么?



OSPF网络类型简介

- 在学习DR和BDR的概念之前，需要首先了解OSPF的网络类型。
- OSPF网络类型是一个非常重要的接口变量，这个变量将影响OSPF在接口上的操作，例如采用什么方式发送OSPF协议报文，以及是否需要选举DR、BDR等。
- 接口默认的OSPF网络类型取决于接口所使用的数据链路层封装。
- 如图所示，OSPF的有四种网络类型，Broadcast、NBMA、P2MP和P2P。



[R1-GigabitEthernet1/0/0] ospf network-type ?

broadcast	Specify OSPF broadcast network
nbma	Specify OSPF NBMA network
p2mp	Specify OSPF point-to-multipoint network
p2p	Specify OSPF point-to-point network



OSPF网络类型 (1)

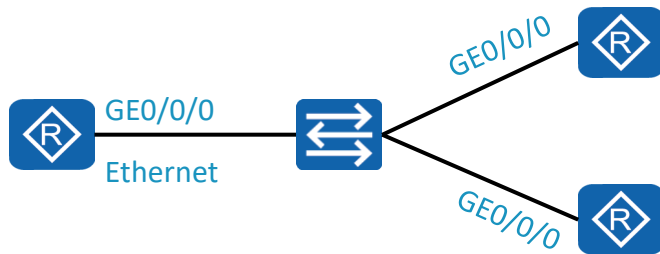
- 一般情况下，链路两端的OSPF接口网络类型必须一致，否则双方无法建立邻居关系。
- OSPF网络类型可以在接口下通过命令手动修改以适应不同网络场景，例如可以将BMA网络类型修改为P2P。

P2P (Point-to-Point, 点对点)



- P2P指的是在一段链路上只能连接两台网络设备的环境。
- 典型的例子是PPP链路。当接口采用PPP封装时，OSPF在该接口上采用的缺省网络类型为P2P。

BMA (Broadcast Multiple Access, 广播式多路访问)



- BMA也被称为Broadcast，指的是一个允许多台设备接入的、支持广播的环境。
- 典型的例子是Ethernet（以太网）。当接口采用Ethernet封装时，OSPF在该接口上采用的缺省网络类型为BMA。



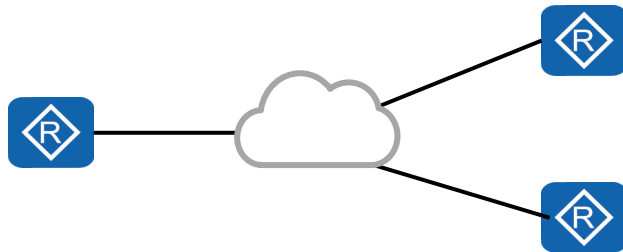
OSPF网络类型 (2)

NBMA (Non-Broadcast Multiple Access, 非广播式多路访问)



- NBMA指的是一个允许多台网络设备接入且不支持广播的环境。
- 典型的例子是帧中继 (Frame-Relay) 网络。

P2MP (Point to Multi-Point, 点到多点)

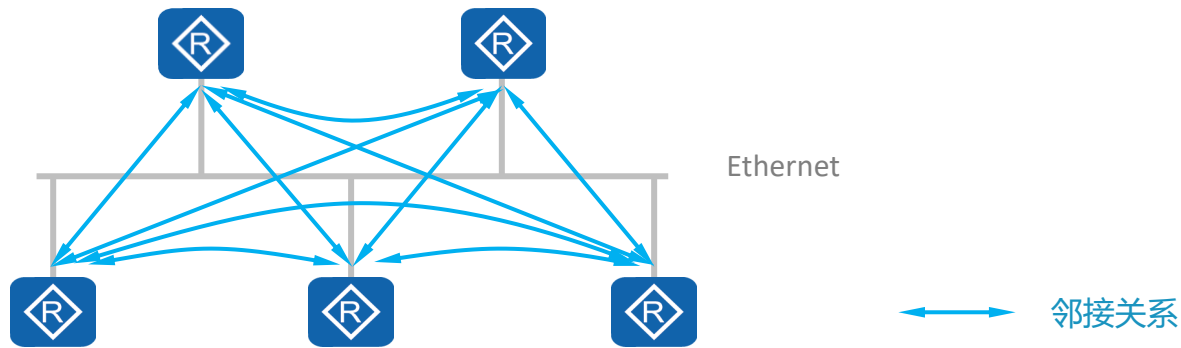


- P2MP相当于将多条P2P链路的一端进行捆绑得到的网络。
- 没有一种链路层协议会被缺省的认为是P2MP网络类型。该类型必须由其他网络类型手动更改。
- 常用做法是将非全连通的NBMA改为点到多点的网络。



DR与BDR的背景

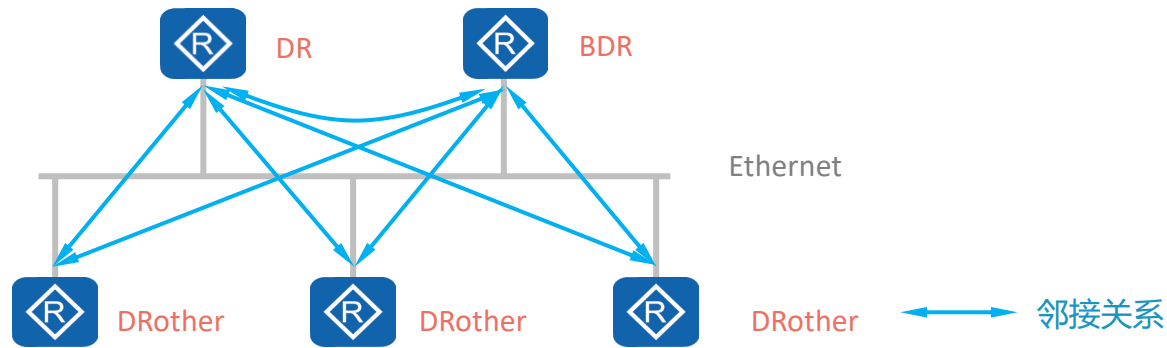
- MA (Multi-Access) 多路访问网络有两种类型：广播型多路访问网络 (BMA) 及非广播型多路访问网络 (NBMA)。以太网 (Ethernet) 是一种典型的广播型多路访问网络。
- 在MA网络中，如果每台OSPF路由器都与其他的所有路由器建立OSPF邻接关系，便会导致网络中存在过多的OSPF邻接关系，增加设备负担，也增加了网络中泛洪的OSPF报文数量。
- 当拓扑出现变更，网络中的LSA泛洪可能会造成带宽的浪费和设备资源的损耗。





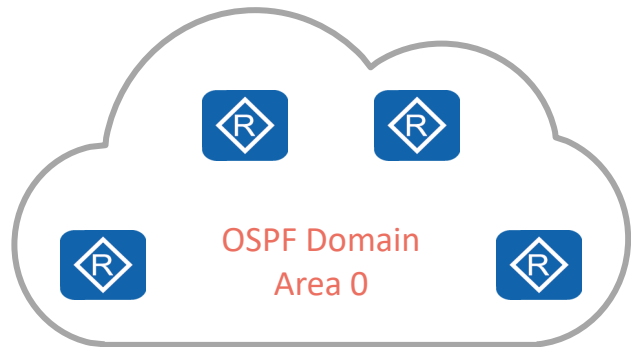
DR与BDR

- 为优化MA网络中OSPF邻接关系，OSPF指定了三种OSPF路由器身份，DR（Designated Router，指定路由器）、BDR（Backup Designated Router，备用指定路由器）和DRoher路由器。
- 只允许DR、BDR与其他OSPF路由器建立邻接关系。DRoher之间不会建立全毗邻的OSPF邻接关系，双方停滞在2-way状态。
- BDR会监控DR的状态，并在当前DR发生故障时接替其角色。





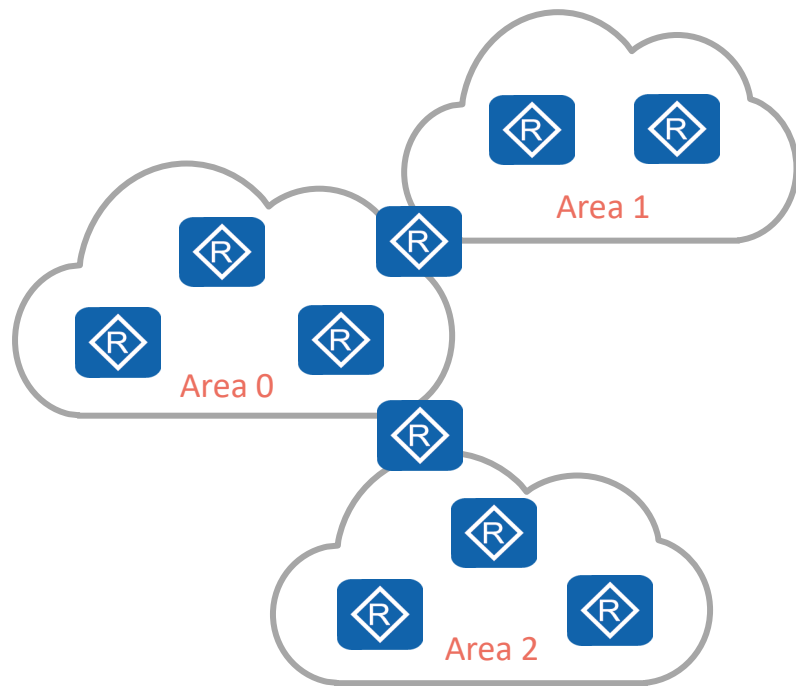
OSPF域与单区域



- OSPF域 (Domain)：一系列使用相同策略的连续OSPF网络设备所构成的网络。
- OSPF路由器在同一个区域 (Area) 内网络中泛洪LSA。为了确保每台路由器都拥有对网络拓扑的一致认知，LSDB需要在区域内进行同步。
- 如果OSPF域仅有一个区域，随着网络规模越来越大，OSPF路由器的数量越来越多，这将导致诸多问题：
 - LSDB越来越庞大，同时导致OSPF路由表规模增加。路由器资源消耗多，设备性能下降，影响数据转发。
 - 基于庞大的LSDB进行路由计算变得困难。
 - 当网络拓扑变更时，LSA全域泛洪和全网SPF重计算带来巨大负担。



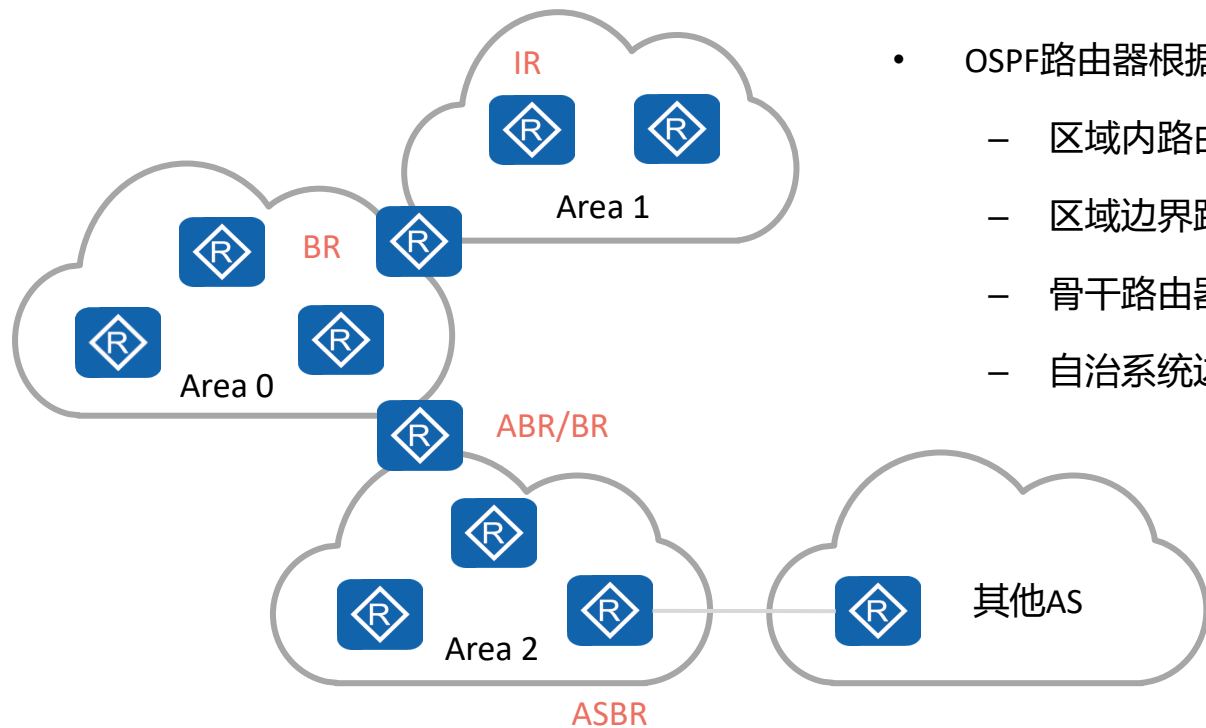
OSPF多区域



- OSPF引入区域 (Area) 的概念, 将一个OSPF域划分成多个区域, 可以使OSPF支撑更大规模组网。
- OSPF多区域的设计减小了LSA泛洪的范围, 有效的把拓扑变化的影响控制在区域内, 达到网络优化的目的。
- 在区域边界可以做路由汇总, 减小了路由表规模。
- 多区域提高了网络扩展性, 有利于组建大规模的网络。



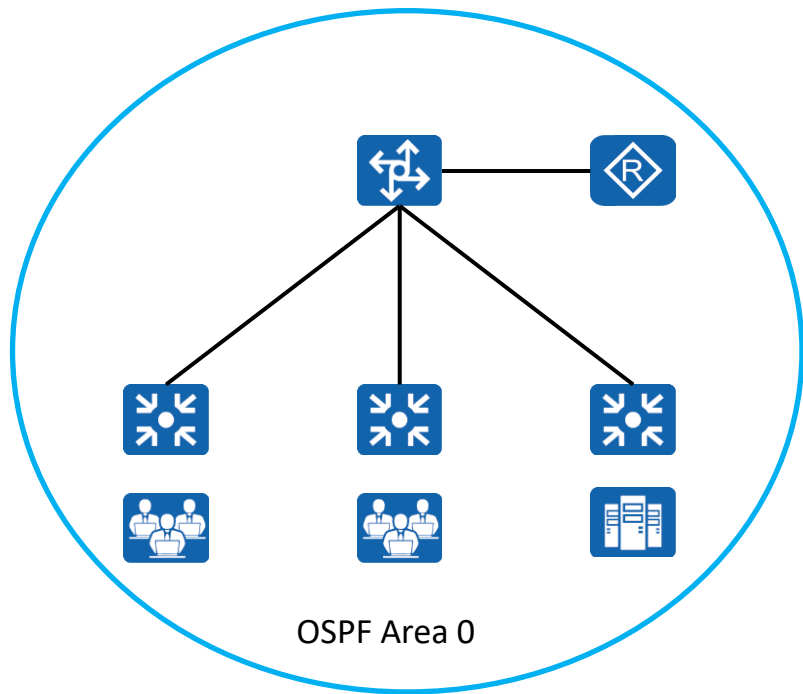
OSPF路由器类型



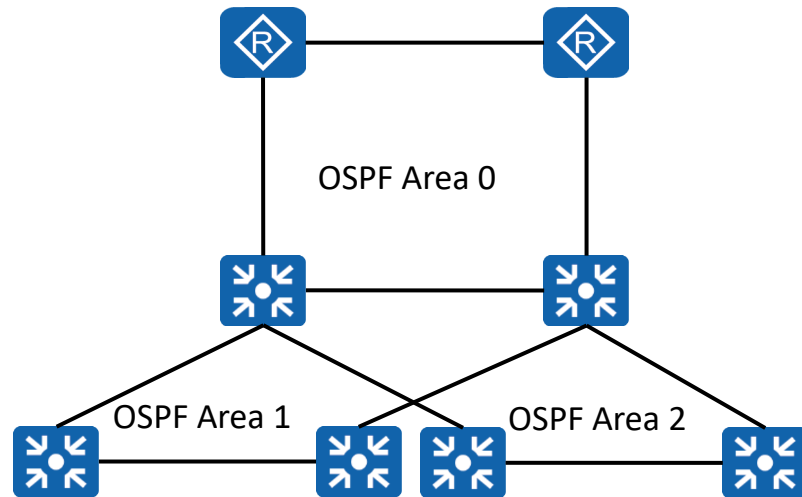
- OSPF路由器根据其位置或功能不同，有这样几种类型：
 - 区域内路由器 (Internal Router)
 - 区域边界路由器ABR (Area Border Router)
 - 骨干路由器 (Backbone Router)
 - 自治系统边界路由器ASBR (AS Boundary Router)



OSPF单区域&多区域典型组网



中小型企业网（单区域）



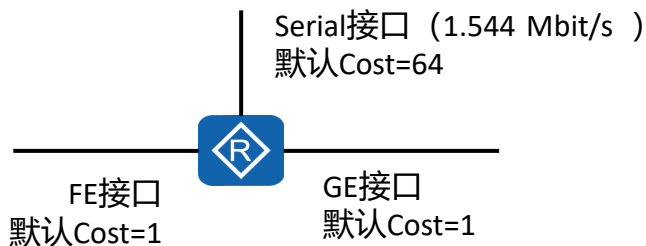
大型企业网（多区域）



OSPF的基础术语：度量值

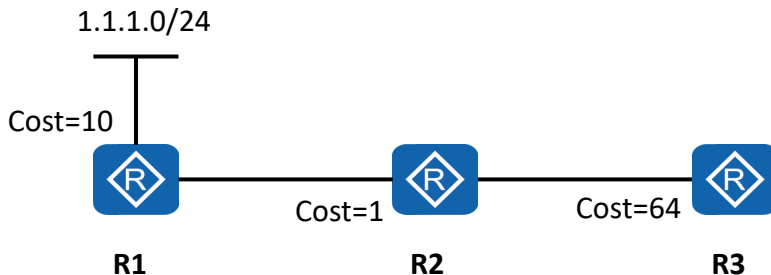
- OSPF使用Cost（开销）作为路由的度量值。每一个激活了OSPF的接口都会维护一个接口Cost值，缺省时 $\text{接口Cost值} = \frac{100 \text{ Mbit/s}}{\text{接口带宽}}$ 。其中100 Mbit/s为OSPF指定的缺省参考值，该值是可配置的。
- 笼统地说，一条OSPF路由的Cost值可以理解为是从目的网段到本路由器沿途所有入接口的Cost值累加。

OSPF接口Cost值



- OSPF不同接口因其带宽不同，有不同的Cost。

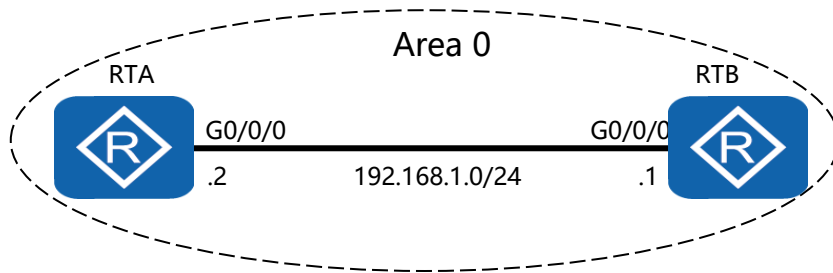
OSPF路径累计Cost值



- 在R3的路由表中，到达1.1.1.0/24的OSPF路由的Cost值=10+1+64，即75。



OSPF配置



```
[RTA]ospf router-id 1.1.1.1
[RTA-ospf-1]area 0
[RTA-ospf-1-area-0.0.0.0]network 192.168.1.0 0.0.0.255
```



配置验证

```
[RTA]display ospf peer
```

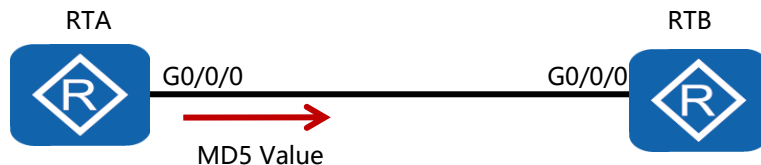
```
OSPF Process 1 with Router ID 1.1.1.1  
Neighbors
```

```
Area 0.0.0.0 interface 192.168.1.2(GigabitEthernet0/0/0)'s  
neighbors
```

```
Router ID: 2.2.2.2      Address: 192.168.1.1  
State: Full  Mode:Nbr is  Slave  Priority: 1  
DR: 192.168.1.2  BDR: 192.168.1.1  MTU: 0  
Dead timer due in 40  sec  
Retrans timer interval: 5  
Neighbor is up for 00:00:31  
Authentication Sequence: [ 0 ]
```



OSPF认证



```
[RTA]interface GigabitEthernet0/0/0
```

```
[RTA-GigabitEthernet0/0/0]ospf authentication-mode md5 1 huawei
```

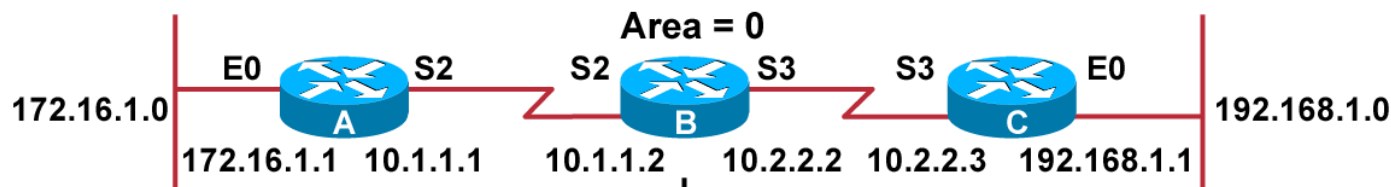


配置验证

```
<RTA>terminal debugging
<RTA>debugging ospf packet
Aug 19 2013 08:10:06.850.2+00:00 R2 RM/6/RMDEBUG: Source Address:
192.168.1.2
Aug 19 2013 08:10:06.850.3+00:00 R2 RM/6/RMDEBUG: Destination
Address: 224.0.0.5
.....
Aug 19 2013 08:10:06.850.6+00:00 R2 RM/6/RMDEBUG: Area: 0.0.0.0,
Chksum: 0
Aug 19 2013 08:10:06.850.7+00:00 R2 RM/6/RMDEBUG: AuType: 02
Aug 19 2013 08:10:06.850.8+00:00 R2 RM/6/RMDEBUG: Key(ascii): * *
* * * * *
```




OSPF 配制举例



```
router ospf 100
network 10.1.1.2 0.0.0.0 area 0
network 10.2.2.2 0.0.0.0 area 0
```

ICND20GR_242



查看 OSPF 配置

```
Router#show ip protocols
```

- 查看启用的路由协议

```
Router#show ip route
```

- 查看路由表

```
Router#show ip ospf interface
```

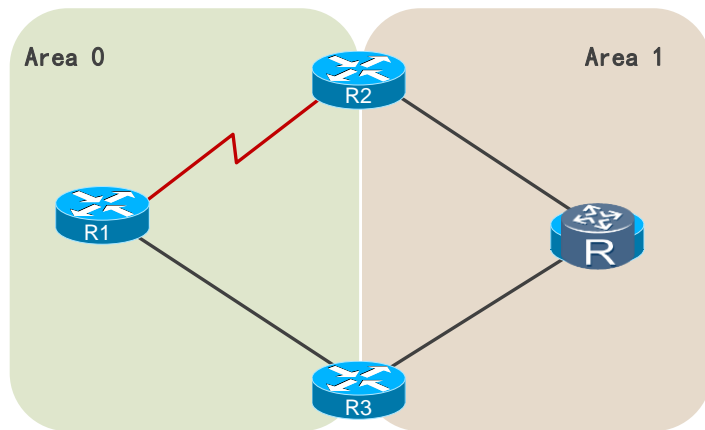
- 查看区域号和与此相关的信息

```
Router#show ip ospf neighbor
```

- 查看在每一个接口上的邻居信息



课堂实验七



实验需求：

1. 四台路由器都创建一个环回口
2. 全网运行OSPF，自治系统号90，router-id手动指定loopback0
3. 实现R1的环回口 ping R4的环回口
4. 并且tracetrouter 路径是两条等价负载均衡路径
5. 讨论ospf区域是路由器概念还是接口概念？



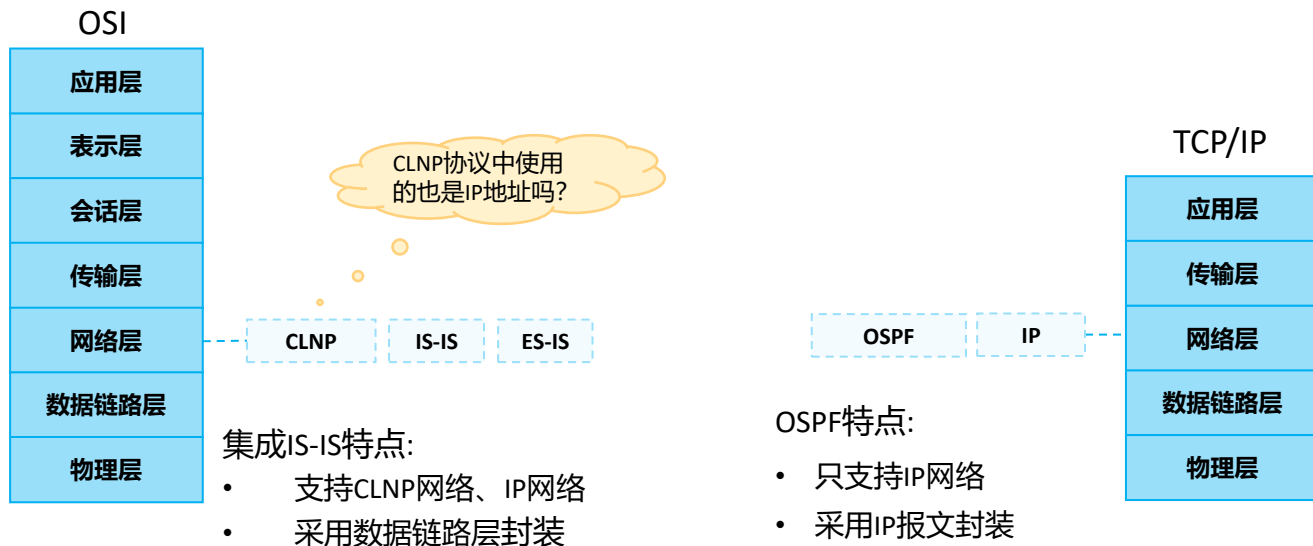
ISIS简介

- 和OSPF一样，IS-IS也是一种基于链路状态并使用最短路径优先算法进行路由计算的一种IGP协议。IS-IS最初是国际化标准组织ISO为它的无连接网络协议CLNP设计的一种动态路由协议。
- 为了提供对IP的路由支持，IETF在RFC1195中对IS-IS进行了扩充和修改，使它能够同时应用在TCP/IP和OSI环境中，修订后的IS-IS协议被称为集成化的IS-IS。由于IS-IS的简便性及扩展性强的特点，目前在大型ISP的网络中被广泛地部署。



IS-IS概述

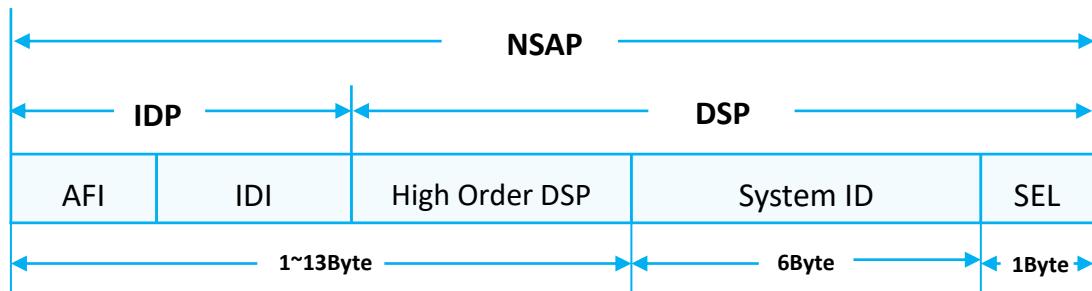
IS-IS是ISO定义的OSI协议栈中的CLNS（ConnectionLess Network Service，无连接网络服务）的一部分。





NSAP

- NSAP (Network Service Access Point, 网络服务访问点) 是OSI协议栈中用于定位资源的地址, 主要用于提供网络层和上层应用之间的接口。NSAP包括IDP及DSP, 如下图所示:



- IDP (Initial Domain Part) 相当于IP地址中的主网络号。它是由ISO规定, 并由AFI (Authority and Format Identifier) 与IDI (Initial Domain Identifier) 两部分组成。AFI表示地址分配机构和地址格式, IDI用来标识域。
- DSP (Domain Specific Part) 相当于IP地址中的子网号和主机地址。它由High Order DSP、System ID和SEL三个部分组成。High Order DSP用来分割区域, System ID用来区分主机, SEL (NSAP Selector) 用来指示服务类型。



NET

- NET (Network Entity Title, 网络实体名称) 是OSI协议栈中设备的网络层信息, 主要用于路由计算, 由区域地址 (Area ID) 和System ID组成, 可以看作是特殊的NSAP (SEL为00的NSAP)。
- NET的长度与NSAP的相同, 最长为20Byte, 最短为8Byte。
- 在IP网络中运行IS-IS时, 只需配置NET, 根据NET地址设备可以获取到Area ID以及System ID。

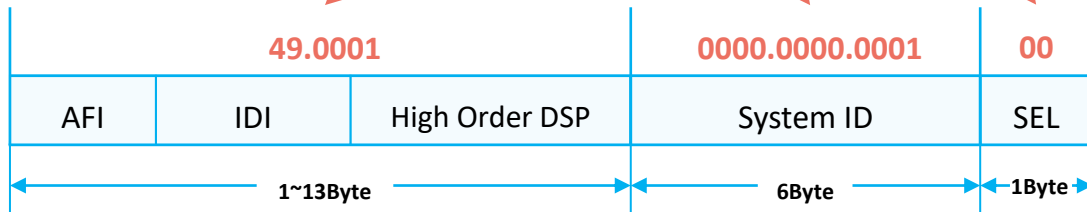
NET示例: 49.0001.0000.0000.0001.00

Area ID System ID SEL

49.0001

0000.0000.0001

00

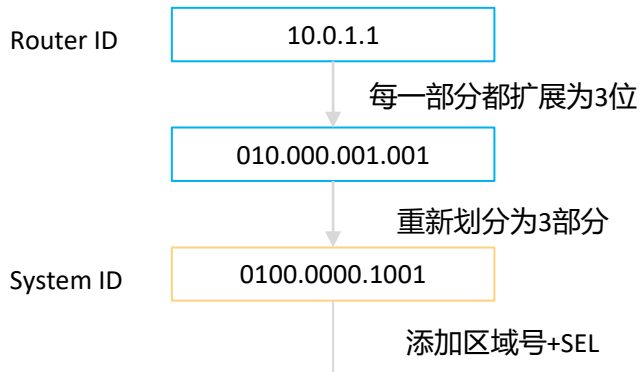


Area ID跟OSPF的不太一样, 拓扑结构也不同吗?



NET的配置举例

- 每台运行IS-IS的网络设备至少需拥有一个NET，当然，一台设备也可以同时配置多个NET，但是这些NET的System ID必须相同。
- 在华为的网络设备上，System ID的长度总是固定的6Byte。在一个IS-IS路由域中，设备的System ID必须唯一，为了便于管理，一般根据Router ID配置System ID。



例如一台设备的区域号为49.0001，Router ID 为 10.0.1.1，则其对应的NET地址为：

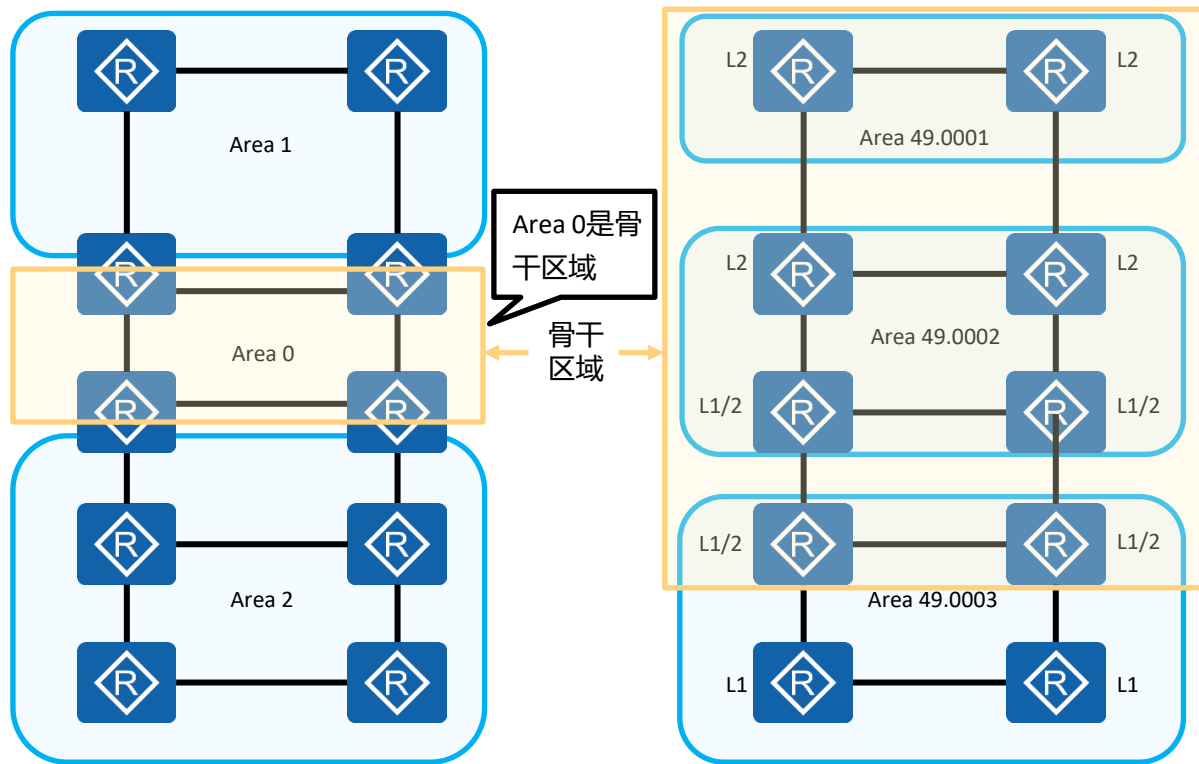
49.0001.0100.0000.1001.00。

49.0001.0100.0000.1001.00

NET



IS-IS和OSPF区域划分的区别



- IS-IS在自治系统内采用骨干区域与非骨干区域两级的分层结构：
 - Level-1路由器部署在非骨干区域。
 - Level-2路由器和Level-1-2路由器部署在骨干区域。
- 每一个非骨干区域都通过Level-1-2路由器与骨干区域相连。

如图所示，整个骨干区域不仅包括Area 49.0002中的所有路由器，还包括其它区域的Level 2和Level-1-2路由器。

L1: Level-1

L1/2: Level-1-2

L2: Level-2

OSPF

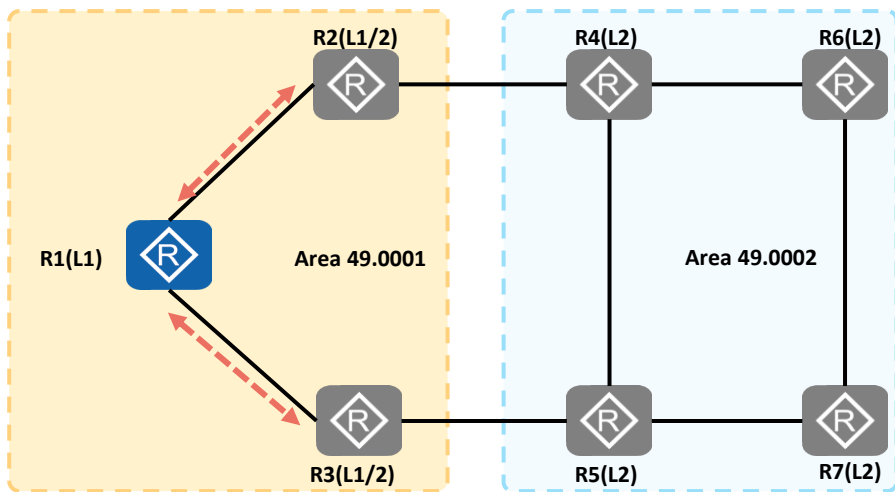
IS-IS



IS-IS路由器的分类（1）

Level-1路由器

- Level-1路由器（例如图中的R1）是一种IS-IS区域内部路由器，它只与属于同一区域的Level-1和Level-1-2路由器形成邻接关系，这种邻接关系称为Level-1邻接关系。Level-1路由器无法与Level-2路由器建立邻接关系。
- Level-1路由器只负责维护Level-1的链路状态数据库LSDB，该LSDB只包含本区域的路由信息。值得一提的是，Level-1路由器必须通过Level-1-2路由器接入IS-IS骨干区域从而访问其他区域。



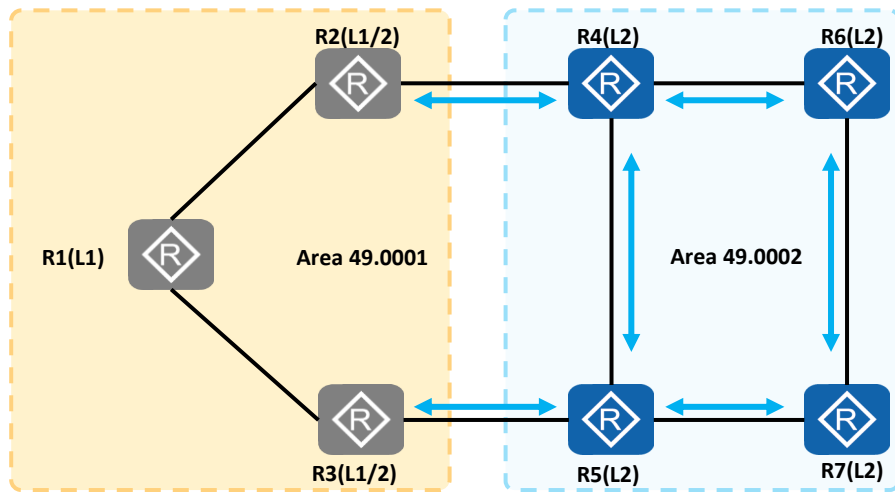
Level-1邻接关系



IS-IS路由器的分类 (2)

Level-2路由器

- Level-2路由器（例如图中的R4、R5、R6、R7）是IS-IS骨干路由器，它可以与同一或者不同区域的Level-2路由器或者Level-1-2路由器形成邻接关系。Level-2路由器维护一个Level-2的LSDB，该LSDB包含整个IS-IS域的所有路由信息。
- 所有Level-2级别（即形成Level-2邻接关系）的路由器组成路由域的骨干网，负责在不同区域间通信。路由域中Level-2级别的路由器必须是物理连续的，以保证骨干网的连续性。

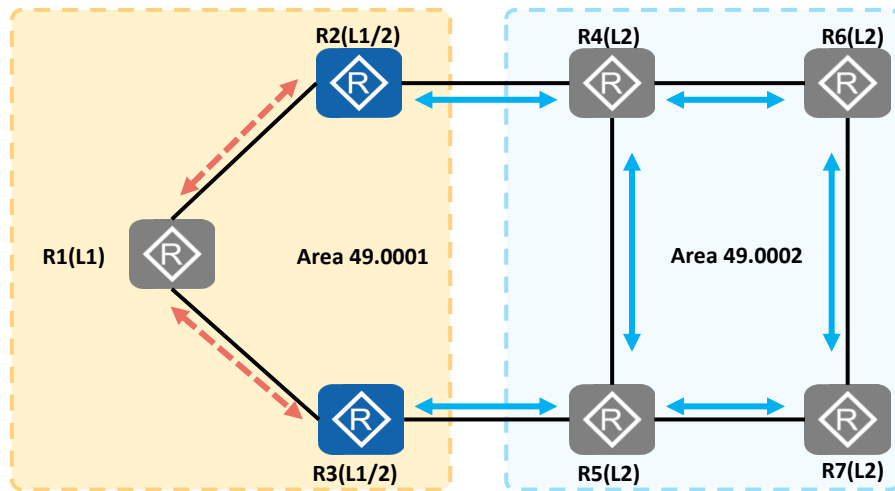




IS-IS路由器的分类 (3)

Level-1-2路由器

- Level-1-2路由器与OSPF中的ABR非常相似，它也是IS-IS骨干网络的组成部分。
- Level-1-2路由器维护两个LSDB，Level-1的LSDB用于区域内路由，Level-2的LSDB用于区域间路由。
- 同时属于Level-1和Level-2的路由器称为Level-1-2路由器（例如图中的R2和R3），它可以与同一区域的Level-1和Level-1-2路由器形成Level-1邻接关系，也可以与其他区域的Level-2和Level-1-2路由器形成Level-2的邻接关系。



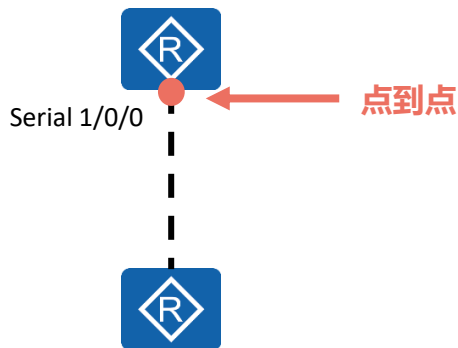
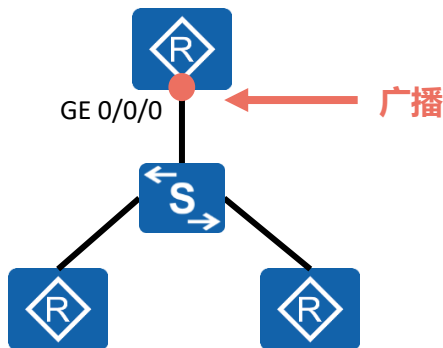
---> Level-1邻接关系
=> Level-2邻接关系



IS-IS支持的网络类型

IS-IS会自动根据接口的数据链路层封装决定该接口的缺省网络类型，IS-IS支持两种类型的网络：

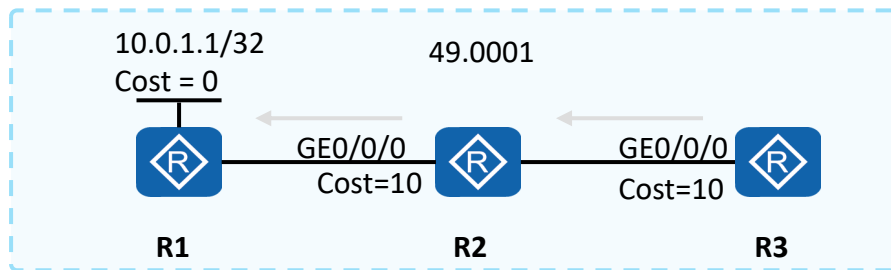
- 广播（Broadcast）：如Ethernet。
- 点到点（P2P）：如PPP、HDLC等。





IS-IS开销值

- IS-IS使用Cost（开销）作为路由度量值，Cost值越小，则路径越优。IS-IS链路的Cost与设备的接口有关，与OSPF类似，每一个激活了IS-IS的接口都会维护接口Cost。然而与OSPF不同的是，IS-IS接口的Cost在缺省情况下并不与接口带宽相关（在实际部署时，IS-IS也支持根据带宽调整Cost值），无论接口带宽多大，缺省时Cost为10。
- 一条IS-IS路径的Cost等于本路由器到达目标网段沿途的所有链路的Cost总和。
- IS-IS有三种方式来确定接口的开销，按照优先级由高到低分别是：
 - 接口开销：为单个接口设置开销。
 - 全局开销：为所有接口设置开销。
 - 自动计算开销：根据接口带宽自动计算开销。



如图所示，全网运行IS-IS，则R3到达10.0.1.1/32的Cost值为20（10+10+0）。



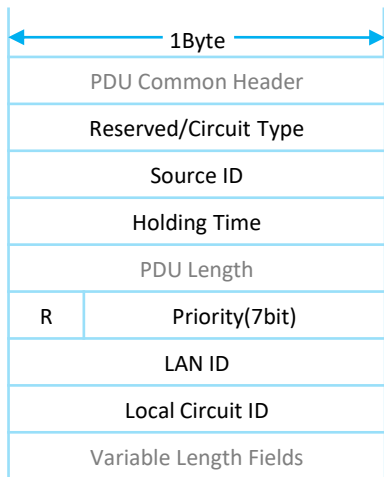
IS-IS邻接关系建立原则

- IS-IS按如下原则建立邻接关系：
 - 只有同一层次的相邻路由器才有可能成为邻接。
 - 对于Level-1路由器来说，Area ID必须一致。
 - 链路两端IS-IS接口的网络类型必须一致。
 - 链路两端IS-IS接口的地址必须处于同一网段（默认情况下）。
- 由于IS-IS是直接运行在数据链路层上的协议，并且最早设计是给CLNP使用的，IS-IS邻接关系的形成与IP地址无关。但在实际的部署中，在IP网络上运行IS-IS时，需要检查对方的IP地址的。如果接口配置了从IP，那么只要双方有某个IP（主IP或者从IP）在同一网段，就能建立邻接，不一定要主IP相同。



IIH

IIH报文用于建立和维持邻接关系，广播网络中的Level-1 IS-IS路由器使用Level-1 LAN IIH；广播网络中的Level-2 IS-IS路由器使用Level-2 LAN IIH；点到点网络中则使用P2P IIH。



Reserved/Circuit Type: 表示路由器的类型 (01表示L1, 10表示L2, 11表示L1/L2)。

Source ID : 发出Hello报文的路由器的System ID。

Holding Time : 保持时间。在此时间内如果没有收到邻接发来的Hello报文，则中止已建立的邻接关系。

Priority : 选举DIS的优先级，取值范围为0 ~ 127。数值越大，优先级越高。该字段只在广播网中的Hello消息(LAN IIH消息)携带；点到点网络的Hello消息(P2P IIH消息)没有此字段，也没有此字段之前的R保留位。

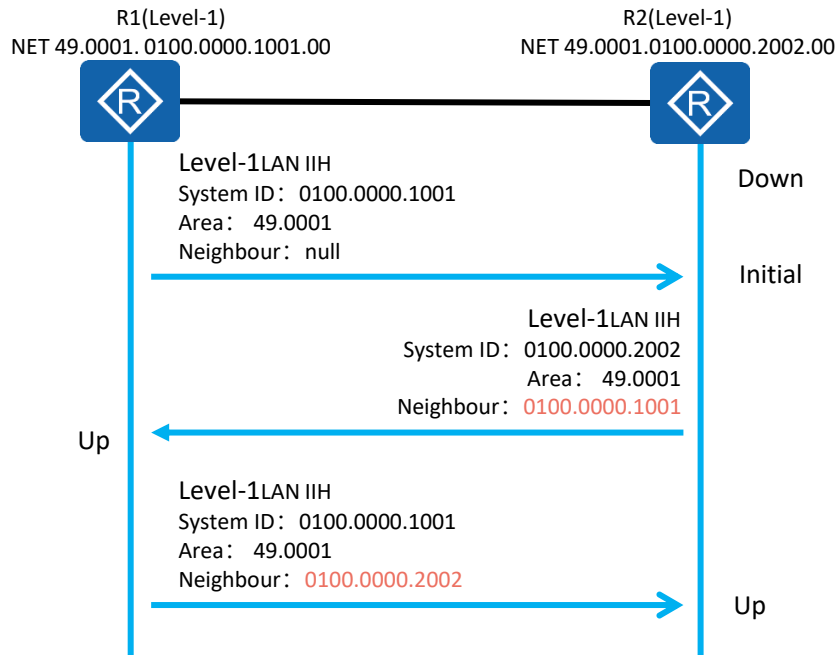
LAN ID : 包括DIS的System ID和伪节点ID。该字段只在广播网中的Hello消息(LAN IIH消息)携带；点到点网络的Hello消息(P2P IIH消息)没有此字段。

Local Circuit ID : 本地链路ID。该字段只在点到点网络的Hello消息(P2P IIH消息)携带；广播网中的Hello消息(LAN IIH消息)没有此字段。



广播网络中的邻接关系建立过程

两台运行IS-IS的路由器在交互协议报文实现路由功能之前必须首先建立邻接关系。在不同类型的网络上，IS-IS的邻接建立方式并不相同。在广播网络中，使用三次握手建立邻接关系。



R1及R2通过千兆以太网接口互联，这两台直连的Level-1路由器建立邻接关系的过程如下：

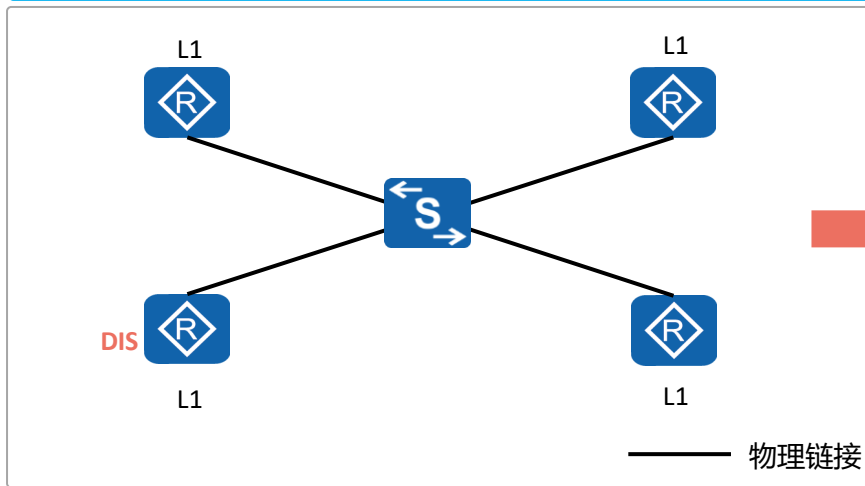
1. 在Down状态下，R1组播发送Level-1 LAN IIH，此报文中邻接列表为空。
2. R2收到此报文后，将邻接状态标识为Initial。然后，R2再向R1回复Level-1 LAN IIH，此报文中标识R1为R2的邻接。
3. R1收到此报文后，将自己与R2的邻接状态标识为Up。然后R1再向R2发送一个标识R2为R1邻接的Level-1 LAN IIH。
4. R2收到此报文后，将自己与R1的邻接状态标识为Up。这样，两个路由器成功建立了邻接关系。
5. 广播网络中需要选举DIS，在邻接关系建立后，路由器会等待两个Hello报文间隔，再进行DIS的选举。



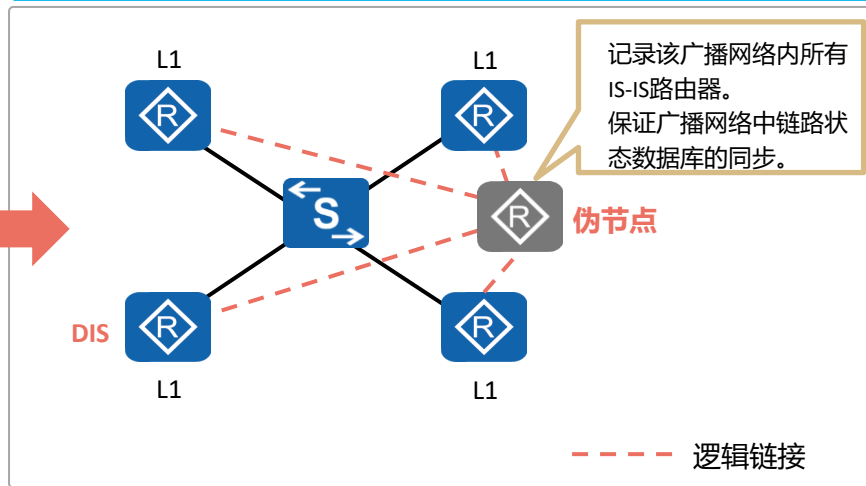
DIS与伪节点

- 在广播网络中，IS-IS需要在所有的路由器中选举一个路由器作为DIS（Designated Intermediate System）。
- DIS用来创建和更新伪节点（Pseudonodes），并负责生成伪节点的LSP，用来描述这个网络上有哪些网络设备。伪节点是用来模拟广播网络的一个虚拟节点，并非真实的路由器。在IS-IS中，伪节点用DIS的System ID和Circuit ID（非0值）标识。

物理拓扑



逻辑拓扑





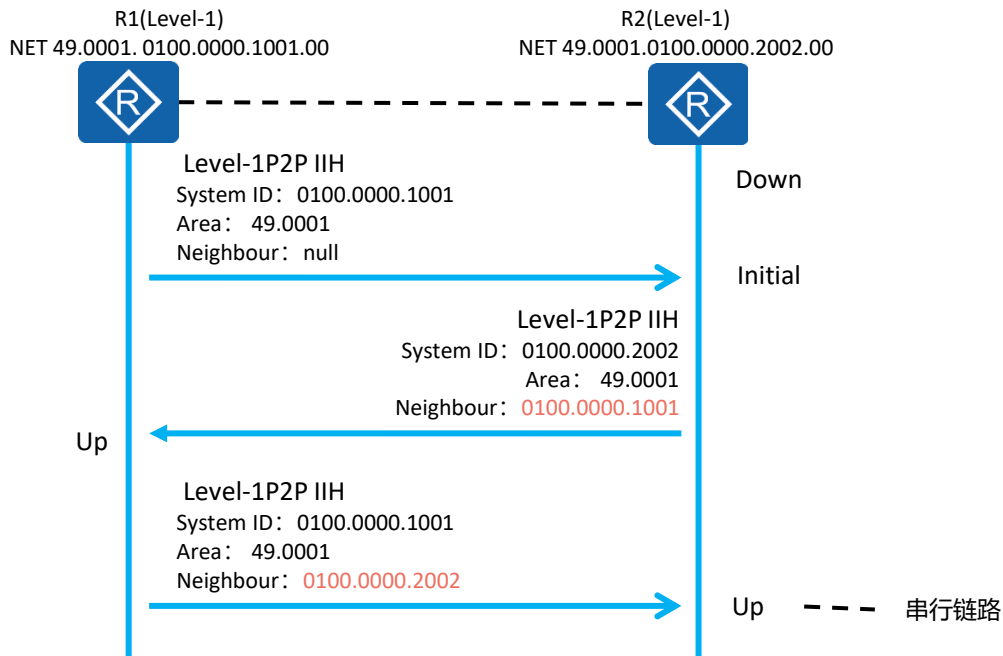
IS-IS中的DIS与OSPF中的DR

- Level-1和Level-2的DIS是分别选举的，用户可以为不同级别的DIS选举设置不同的优先级。
- DIS的选举规则如下：
 - DIS优先级数值最大的被选为DIS。
 - 如果优先级数值最大的路由器有多台，则其中MAC地址最大的路由器会成为DIS。
- DIS发送Hello PDU的时间间隔是普通路由器的1/3，这样可以确保DIS出现故障时能够被更快速地被发现。
- IS-IS中DIS与OSPF协议中DR（Designated Router）的区别：
 - 在IS-IS广播网中，优先级为0的路由器也参与DIS的选举，而在OSPF中优先级为0的路由器则不参与DR的选举。
 - 在IS-IS广播网中，当有新的路由器加入，并符合成为DIS的条件时，这个路由器会被选中成为新的DIS，原有的伪节点被删除。此更改会引起一组新的LSP泛洪。而在OSPF中，当一台新路由器加入后，即使它的DR优先级值最大，也不会立即成为该网段中的DR。
 - 在IS-IS广播网中，同一网段上的同一级别的路由器之间都会形成邻接关系，包括所有的非DIS路由器之间也会形成邻接关系。而在OSPF中，路由器只与DR和BDR建立邻接关系。



点到点网络中的邻接关系建立过程

- 点到点网络中，邻接关系的建立使用两次握手方式：只要路由器收到对端发来的Hello报文，就单方面宣布邻接为Up状态，建立邻接关系。
- 两次握手机制存在明显的缺陷，**华为设备在点到点网络中使用IS-IS时，默认使用三次握手建立邻接关系**。此方式通过三次发送P2P IIH最终建立起邻接关系。





LSP

- IS-IS链路状态报文LSP用于交换链路状态信息。LSP分为两种：Level-1 LSP和Level-2 LSP。Level-1 LSP由Level-1路由器传送，Level-2 LSP由Level-2路由器传送，Level-1-2路由器则可传送以上两种LSP。
- 两类LSP有相同的报文格式。

1Byte			
PDU Common Header			
PDU Length			
Remaining Lifetime			
Holding Time			
LSP ID			
Sequence Number			
Checksum			
P	ATT(4bit)	OL	IS Type
Variable Length Fields			

Remaining Lifetime : LSP的生存时间，以秒为单位。

LSP ID:由三部分组成，System ID、伪节点ID和LSP分片后的编号。

Sequence Number: LSP的序列号。在路由器启动时所发送的第一个LSP报文中的序列号为1，以后当需要生成新的LSP时，新LSP的序列号在前一个LSP序列号的基础上加1。更高的序列号意味着更新的LSP。

Checksum : LSP的校验和。

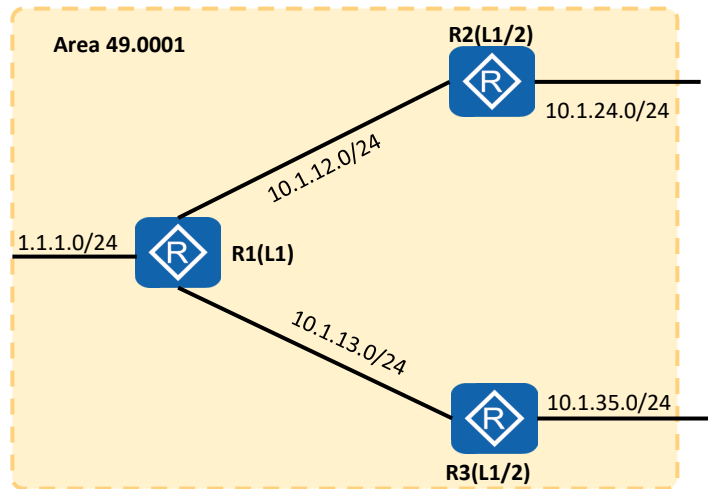
ATT (Attachment) : 由Level-1-2路由器产生，用来指明始发路由器是否与其它区域相连。虽然此标志位也存在于Level-1和Level-2的LSP中，但实际上此字段只和Level-1-2路由器始发的L1 LSP有关。

OL (LSDB Overload, 1bit) : 过载标志位。设置了过载标志位的LSP虽然还会在网络中扩散，但是在计算通过超载路由器的路由时不会被采用。即对路由器设置过载位后，其它路由器在进行SPF计算时不会考虑这台路由器。当路由器内存不足时，系统自动在发送的LSP报文中设置过载标志位。

IS Type(2bit): 生成LSP的路由器的类型。用来指明是Level-1还是Level-2路由器（01表示Level-1，11表示Level-2）。



IS-IS的LSDB



R1的System-ID 0100.0000.1001
R2的System-ID 0100.0000.2002
R3的System-ID 0100.0000.3003

<R1> display isis lsdb

Database information for ISIS(1)

Level-1 Link State Database

LSPID	Seq Num	Checksum	Holdtime	Length	ATT/P/OL
-------	---------	----------	----------	--------	----------

0100.0000.1001.00-00*	0x00000005	0x13a8	1187	97	0/0/0
-----------------------	------------	--------	------	----	-------

0100.0000.1001.01-00*	0x00000001	0xda2e	1185	55	0/0/0
-----------------------	------------	--------	------	----	-------

0100.0000.2002.00-00	0x00000004	0x94e9	1188	86	1/0/0
----------------------	------------	--------	------	----	-------

Total LSP(s): 5

*(In TLV)-Leaking Route, *(By LSPID)-Self LSP, +-Self LSP(Extended),

ATT-Attached, P-Partition, OL-Overload

0100.0000.1001

00

00

*

System-ID

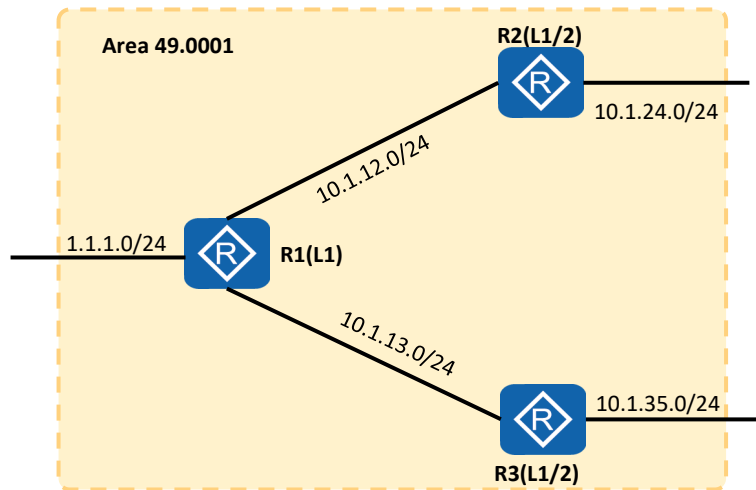
伪节点标识

分片号

自己产生的LSP



查看非伪节点的LSP



```
<R1> display isis lsdb 0100.0000.1001.00-00 verbose
```

Database information for ISIS(1)

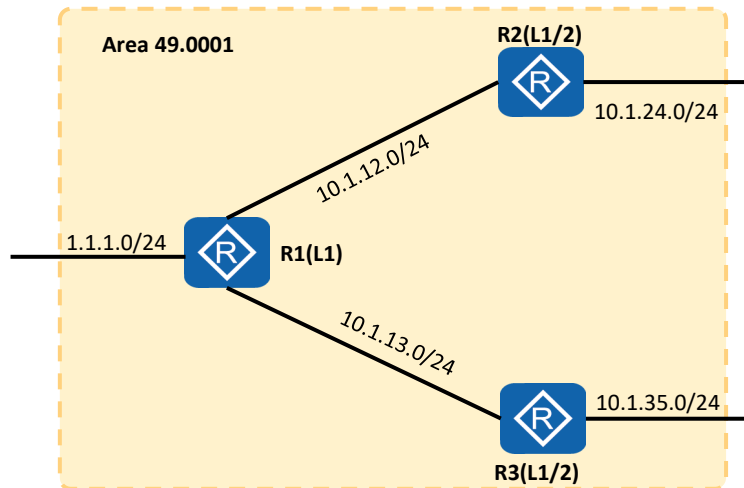
Level-1 Link State Database

LSPID	Seq Num	Checksum	Holdtime	Length	ATT/P/OL

0100.0000.1001.00-00*	0x0000000e	0x9a75	1072	113	0/0/0
SOURCE	0100.0000.1001.00				
NLPID	IPV4				
AREA ADDR	49.0123				
INTF ADDR	10.1.12.1 //描述接口信息				
INTF ADDR	10.1.13.1				
NBR ID	0100.0000.1001.01			COST: 10	//描述邻接关系
NBR ID	0100.0000.3003.01			COST: 10	
IP-Internal	10.1.12.0 255.255.255.0			COST: 10	
IP-Internal	10.1.13.0 255.255.255.0			COST: 10	//描述路由信息
IP-Internal	1.1.1.0 255.255.255.0			COST: 10	
Total LSP(s): 1					



查看伪节点LSP



```
<R1> display isis lsdb 0100.0000.1001.01-00 verbose
```

Database information for ISIS(1)

Level-1 Link State Database

LSPID	Seq Num	Checksum	Holdtime	Length	ATT/P/OL
-------	---------	----------	----------	--------	----------

0100.0000.1001.01-00*	0x00000009	0xca36	431	55	0/0/0
-----------------------	------------	--------	-----	----	-------

SOURCE	0100.0000.1001.01
--------	-------------------

NLPID	IPV4
-------	------

NBR ID	0100.0000.1001.00 COST: 0	//描述该广播网络中所有IS-IS路由器
--------	---------------------------	----------------------

NBR ID	0100.0000.2002.00 COST: 0	//伪节点到达其他路由器的Cost值为0
--------	---------------------------	----------------------

Total LSP(s): 1

*(In TLV)-Leaking Route, *(By LSPID)-Self LSP, +-Self LSP(Extended),

ATT-Attached, P-Partition, OL-Overload

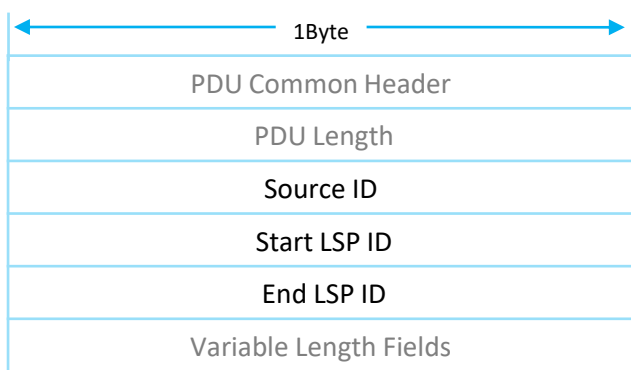
在伪节点LSP中，只包含邻接信息而不包含路由信息。



CSNP

CSNP包含该设备LSDB中所有的LSP摘要，路由器通过交互 CSNP来判断是否需要同步LSDB。

- 在广播网络上，CSNP由DIS定期发送（缺省的发送周期为10秒）。
- 在点到点网络上，CSNP只在第一次建立邻接关系时发送。



Source ID：发出CSNP报文的路由器的System ID。

Start LSP：CSNP报文中第一个LSP的ID值。

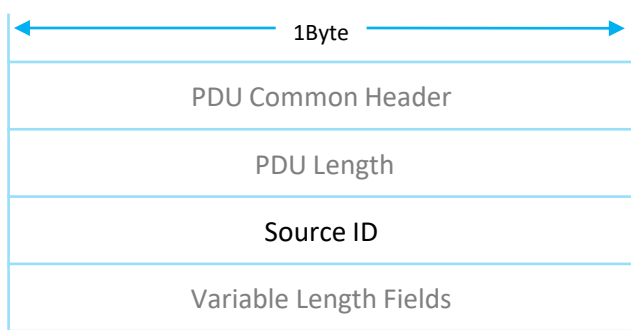
End LSP ID：CSNP报文中最后一个LSP的ID值。



PSNP

PSNP只包含部分LSP的摘要信息（与CSNP不同）：

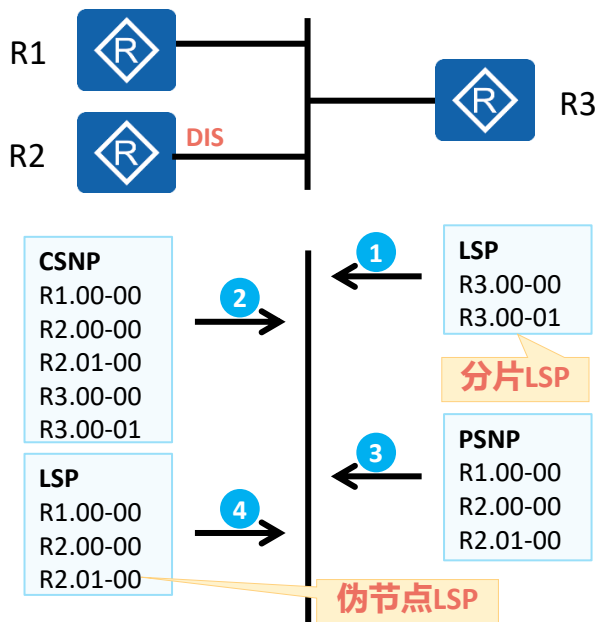
- 当发现LSDB不同步时，PSNP来请求邻居发送新的LSP。
- 在点到的网络中，当收到LSP时，使用PSNP对收到的LSP进行确认。



Source ID：发出PSNP报文的路由器的System ID。



广播网络中LSP的同步过程



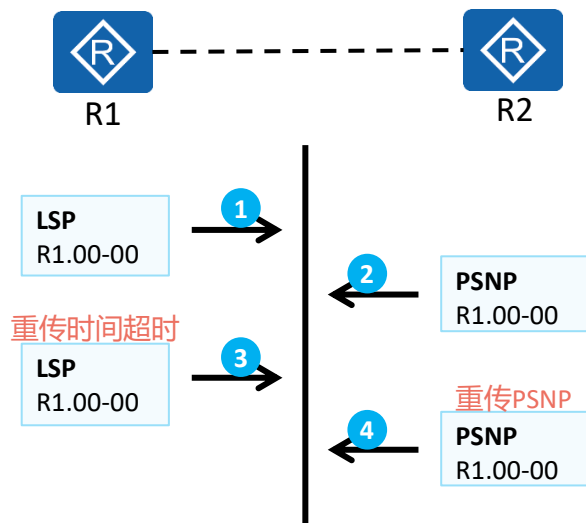
广播网络CSNP报文只由DIS组播发送。

广播网络中新加入路由器与DIS同步LSDB数据库的过程：

1. 新加入的路由器R3首先发送IIH报文，与该广播域中的路由器建立邻接关系。建立邻接关系之后，R3等待LSP刷新定时器超时，然后将自己的LSP发往组播地址（Level-1: 01-80-C2-00-00-14；Level-2: 01-80-C2-00-00-15）。这样网络上所有的邻接都将收到该LSP。
2. 该网段中的DIS会把收到R3的LSP加入到LSDB中，并等待CSNP报文定时器超时并发送CSNP报文。
3. R3收到DIS发来的CSNP报文，对比自己的LSDB数据库，然后向DIS发送PSNP报文请求自己没有的LSP。
4. DIS收到该PSNP报文请求后向R3发送对应的LSP进行LSDB的同步。



点到点网络中LSP的同步过程

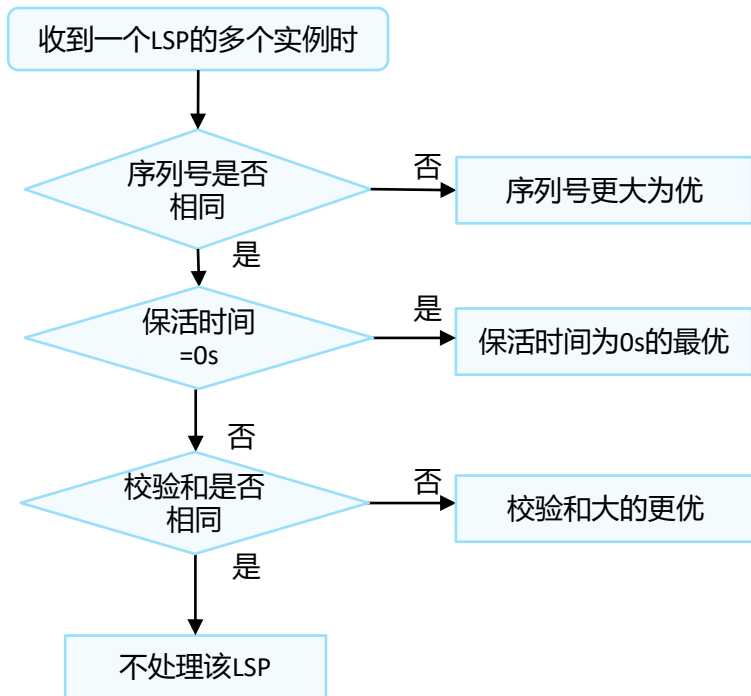


点到点网络上LSDB数据库的同步过程:

- R1先与R2建立邻接关系。
- 建立邻接关系之后，R1与R2会先发送CSNP给对端设备。如果对端的LSDB与CSNP没有同步，则发送PSNP请求索取相应的LSP。
- 假设R2向R1索取相应的LSP。
 1. R1发送R2请求的LSP的同时启动LSP重传定时器，并等待R2发送的PSNP作为收到LSP的确认。
 2. 如果在接口LSP重传定时器超时后，R1没有收到R2发送的PSNP报文作为应答。
 3. 则R1重新发送该LSP。
 4. R2收到LSP后，发送PSNP进行确认。



LSP的处理机制



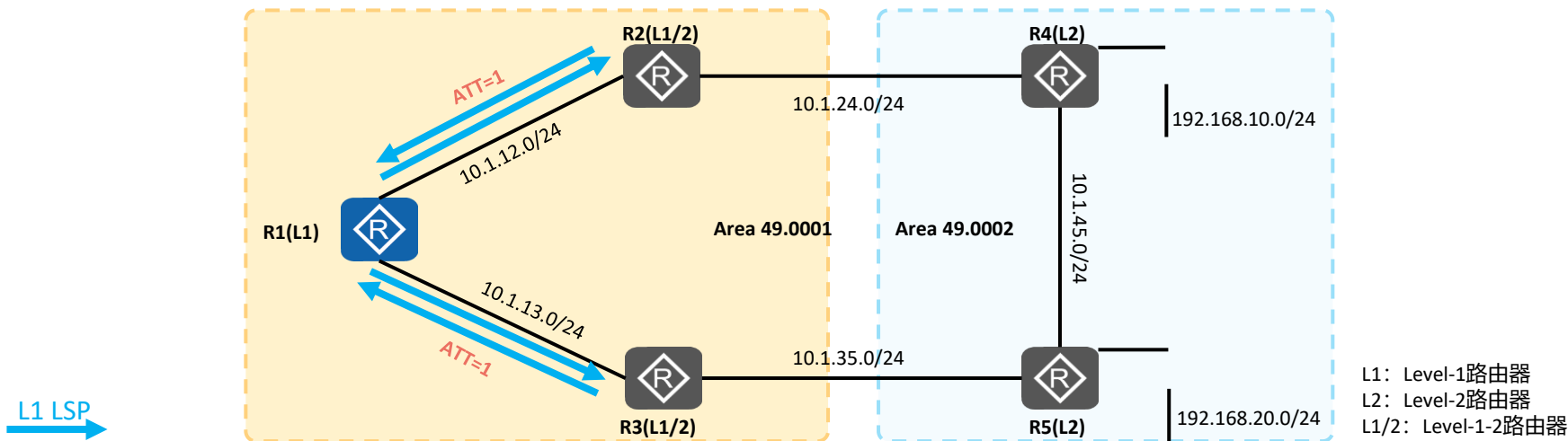
IS-IS通过交互LSP实现链路状态数据库同步，路由器收到LSP后，按照以下原则处理：

- 若收到的LSP比本地LSP的更优，或者本地没有收到的LSP：
 - 在广播网络中：将其加入数据库，并组播发送新的LSP。
 - 在点到点网络中：将其加入数据库，并发送PSNP报文来确认收到此LSP，之后将这新的LSP发送给除了发送该LSP的邻居以外的邻居。
- 若收到的LSP和本地LSP无法比较出优劣，则不处理该LSP。



Level-1路由器的路由计算

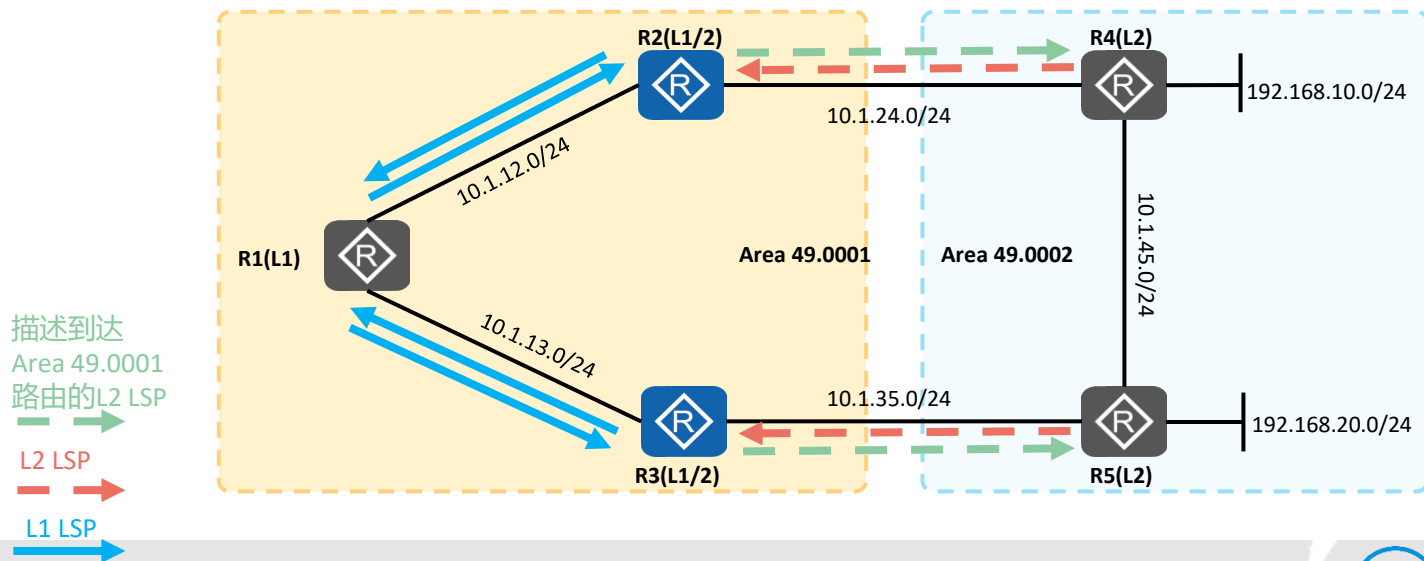
- R1是Level-1路由器，只维护Level-1 LSDB，该LSDB中包含同属一个区域的R2及R3以及R1自己产生的Level-1 LSP。
- R1根据LSDB中的Level-1 LSP计算出Area 49.0001内的拓扑，以及到达区域内各个网段的路由信息。
- R2及R3作为Area 49.0001内的Level-1-2路由器，会在它们向该区域下发的Level-1 LSP中设置ATT标志位，用于向区域内的Level-1路由器宣布可以通过自己到达其他区域。R1作为Level-1路由器，会根据该ATT标志位，计算出指向R2或R3的默认路由。





Level-1-2路由器的路由计算

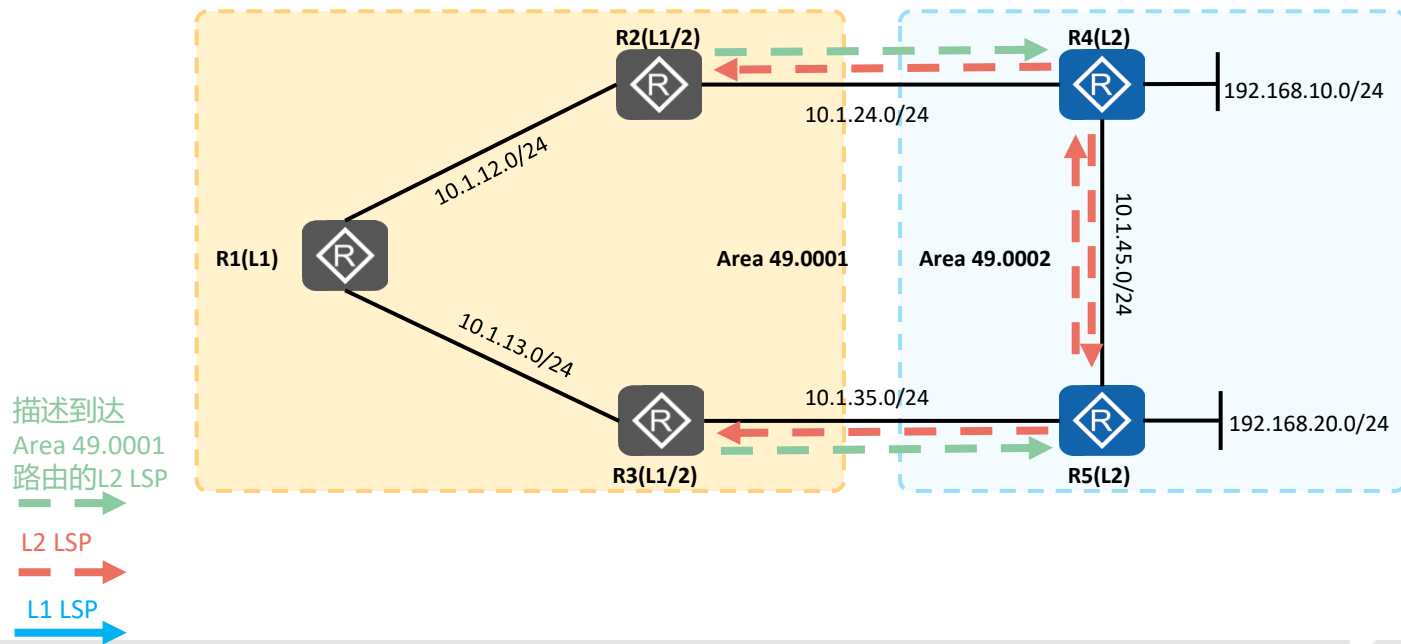
- R2及R3都维护Level-1 LSDB，它们能够通过这些LSDB中的LSP计算出Area 49.0001的路由。
- R2及R3都维护Level-2 LSDB，它们能够通过这些LSDB中的LSP计算出Area 49.0002的路由。
- R2及R3将到达Area 49.0001的路由以Level-2 LSP的形式发送到Area 49.0002。





Level-2路由器的路由计算

R4及R5作为Level-2路由器，只会维护Level-2 LSDB，它们能够根据该LSDB计算出到达全网各个网段的路由。





IS-IS协议的基本配置 (1)

1. 创建IS-IS进程，进入IS-IS进程

```
[Huawei] isis [process-id]
```

参数process-id用来指定一个IS-IS进程。如果不指定参数process-id，则系统默认的进程为1。

2. 配置网络实体名称（NET）

```
[Huawei-isis-1] network-entity net
```

通常情况下，一个IS-IS进程下配置一个NET即可。当区域需要重新划分时，例如将多个区域合并，或者将一个区域划分为多个区域，这种情况下配置多个NET可以在重新配置时仍然能够保证路由的正确性。由于一个IS-IS进程中区域地址最多可配置3个，所以NET最多也只能配3个。在配置多个NET时，必须保证它们的System ID都相同。

3. 配置全局Level级别

```
[Huawei-isis-1] is-level { level-1 | level-1-2 | level-2 }
```

缺省情况下，设备的Level级别为level-1-2。

在网络运行过程中，改变IS-IS设备的级别可能会导致IS-IS进程重启并可能会造成IS-IS邻居断连，建议用户在配置IS-IS时即完成设备级别的配置。



IS-IS协议的基本配置 (2)

4. 进入接口视图

```
[Huawei]interface interface-type interface-number
```

参数interface-type为接口类型，参数interface-number为接口编号。

5. 在接口上使能IS-IS协议

```
[Huawei-GigabitEthernet0/0/1] isis enable [ process-id ]
```

配置该命令后，IS-IS将通过该接口建立邻居、扩散LSP报文。

6. 配置接口Level级别

```
[Huawei-GigabitEthernet0/0/1] isis circuit-level [ level-1 | level-1-2 | level-2 ]
```

缺省情况下，接口的Level级别为level-1-2。

两台Level-1-2设备建立邻居关系时，缺省情况下，会分别建立Level-1和Level-2邻居关系。如果只希望建立Level-1或者Level-2的邻居关系，可以通过修改接口的Level级别实现。



IS-IS协议的基本配置 (3)

7. 设置接口的网络类型为P2P

```
[Huawei-GigabitEthernet0/0/1]isis circuit-type p2p
```

缺省情况下，接口网络类型根据物理接口决定。

使用该命令将广播网接口模拟成P2P接口时，接口发送Hello报文的间隔时间、宣告邻居失效前IS-IS没有收到的邻居Hello报文数目、点到点链路上LSP报文的重新间隔时间以及IS-IS各种认证均恢复为缺省配置，而DIS优先级、DIS名称、广播网络上发送CSNP报文的间隔时间等配置均失效。

8. 恢复接口的缺省网络类型

```
[Huawei-GigabitEthernet0/0/1] undo isis circuit-type
```

使用该命令恢复接口的缺省网络类型时，接口发送Hello报文的间隔时间、宣告邻居失效前IS-IS没有收到的邻居Hello报文数目、点到点链路上LSP报文的重新间隔时间、IS-IS各种认证、DIS优先级和广播网络上发送CSNP报文的间隔时间均恢复为缺省配置。

9. 修改接口的DIS优先级

```
[Huawei-GigabitEthernet0/0/1] isis dis-priority priority [ level-1 | level-2 ]
```

缺省情况下，IS-IS接口DIS优先级为64。

该命令用来指定挑选对应层次DIS（Designated Intermediate System）时接口的优先级。



IS-IS与OSPF差异性

差异性	IS-IS	OSPF
网络类型	少	多
开销方式	复杂	简便
区域类型	少	多
路由报文类型	简单	多样
路由收敛速度	很快	快
扩展性	强	一般
路由负载能力	超强	强

我该怎么选？还是按照自己的实际需要来吧！



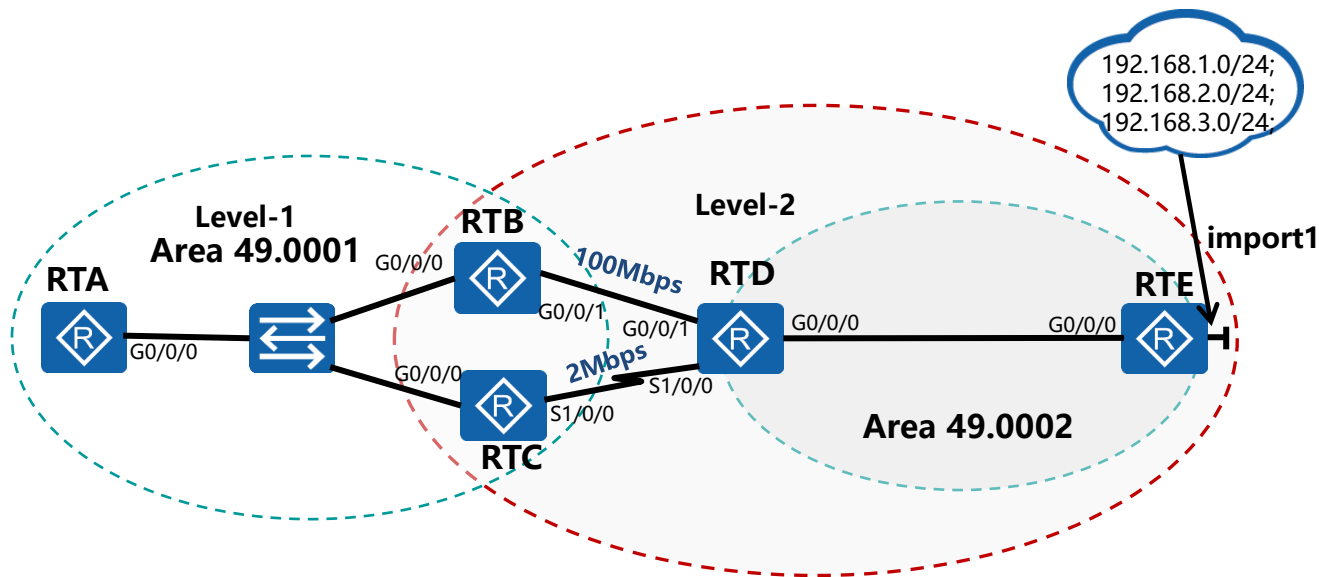


术语对照表

缩略语	OSI术语	IETF术语
IS	Intermediate System	Router
ES	End System	Host
DIS	Designated Intermediate System	OSPF中的DR
SysID	System ID	OSPF中的Router ID
LSP	Link State PDU	OSPF中的LSA
IIH	IS-IS Hello PDU	OSPF中的Hello报文
PSNP	Partial Sequence Number PDU	OSPF中的LSR或LSAck报文
CSNP	Complete Sequence Number PDU	OSPF中的DD报文



IS-IS路由配置需求



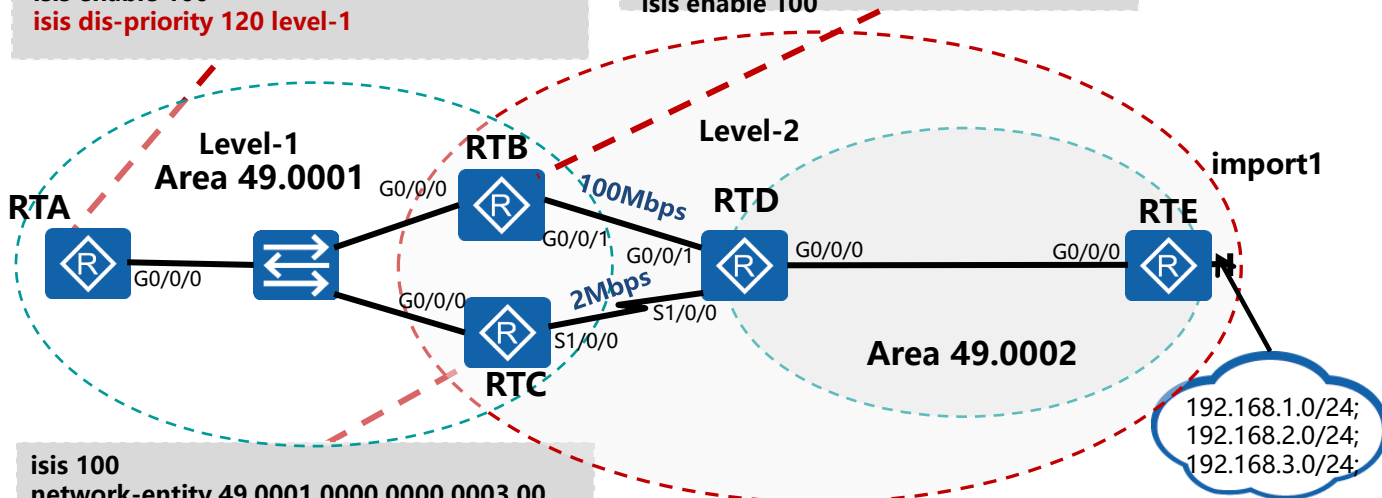
- 如图所示，客户网络所有路由器路由协议要求启用IS-IS，使全网路由可达。全部IS-IS进程号统一为100，其中RTA在Area49.0001区域为DIS，RTD与RTE之间要求采用P2P网络类型，RTE引入直连链路192.168.X.X，要求RTA访问Area49.0002走最优路径。
- 根据上述描述，进行正确配置，使网络路由达到客户需求。



IS-IS路由配置实现 (1)

```
isis 100
network-entity 49.0001.0000.0000.0001.00
is-level level-1
#
interface GigabitEthernet0/0/0
isis enable 100
isis dis-priority 120 level-1
```

```
isis 100
network-entity
49.0001.0000.0000.0002.00
#
interface GigabitEthernet0/0/0
isis enable 100
```



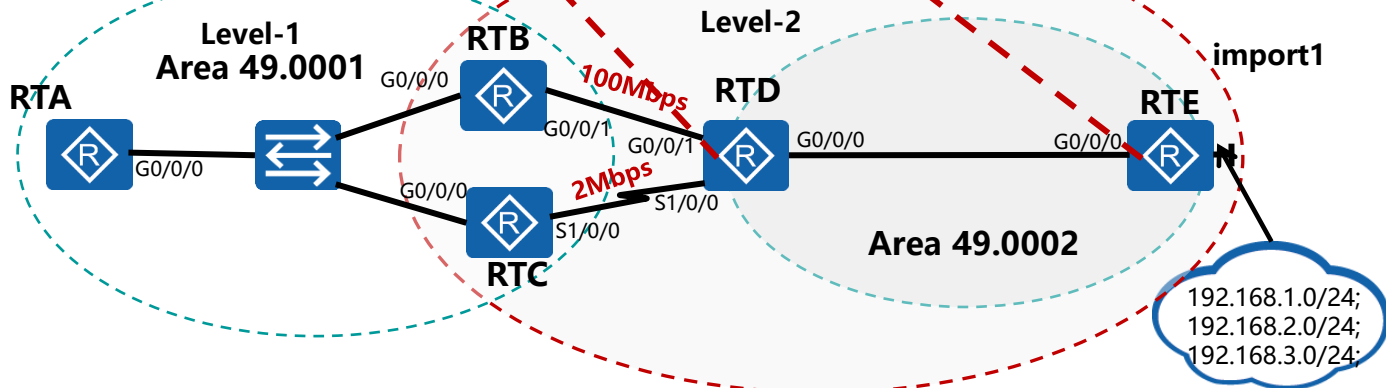
```
isis 100
network-entity 49.0001.0000.0000.0003.00
#
interface GigabitEthernet0/0/0
isis enable 100
```



IS-IS路由配置实现 (2)

```
isis 100
network-entity
49.0002.0000.0000.0004.00
is-level level-2
#
interface GigabitEthernet0/0/0
isis enable 100
isis circuit-type P2P
```

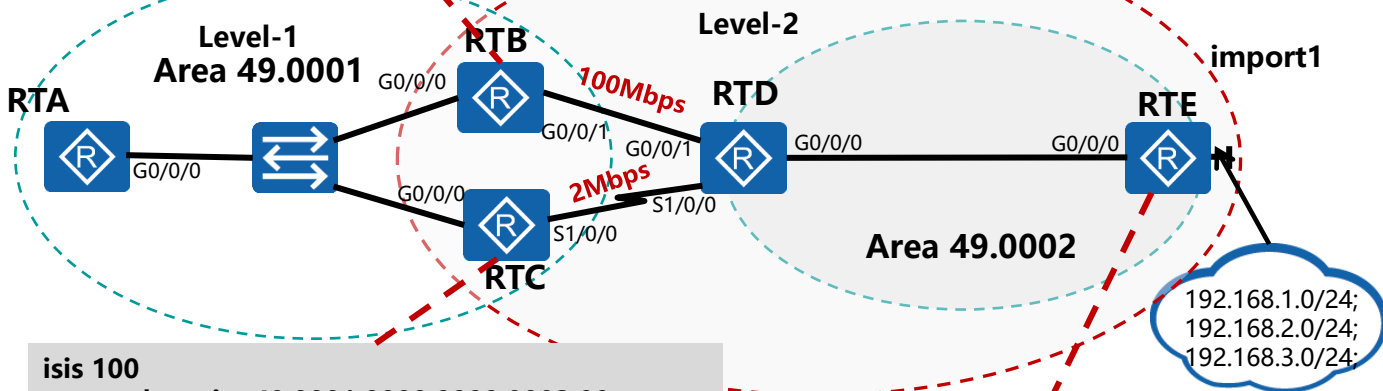
```
isis 100
network-entity
49.0002.0000.0000.0005.00
is-level level-2
#
interface GigabitEthernet0/0/0
isis enable 100
isis circuit-type P2P
```





IS-IS路由配置实现 (3)

```
isis 100
network-entity 49.0001.0000.0000.0002.00
import-route isis level-2 into level-1
#
interface GigabitEthernet0/0/1
isis enable 100
```



```
isis 100
network-entity 49.0001.0000.0000.0003.00
import-route isis level-2 into level-1
#
interface Serial 1/0/0
isis enable 100
```

```
isis 100
import-route direct
```

THANK YOU

Ping 通您的梦想 ~

腾讯课堂交流群：17942636

ADD：苏州市干将东路666号和基广场401-402； Tel：0512-8188 8288；

课程咨询QQ：2853771087 ； 官网 :www.51glab.com