

MPLS LDP 原理与配置

动态 LSP 通过 LDP 协议实现对 FEC 的分类、标签的分配及 LSP 的建立和维护等操作。

动态 LSP 的特点：

组网配置简单，易于管理和维护；

支持基于路由动态建立 LSP，网络拓扑发生变化时，能及时反映网络状况。

LDP 协议

LDP 的 hello 时间为 5s，holdtime 15s，发送的组播地址为 224.0.0.2，

端口号为 UDP 646 发现邻居，其他为 TCP 646

57	42.078000	2.2.2.2	1.1.1.1	LDP	Keep Alive Message
58	42.250000	1.1.1.1	2.2.2.2	TCP	ldp > 63848 [ACK] Seq=157 Ack=175
59	42.250000	1.1.1.1	2.2.2.2	LDP	Keep Alive Message
60	42.297000	2.2.2.2	1.1.1.1	TCP	63848 > ldp [ACK] Seq=175 Ack=175
61	42.390000	192.168.12.2	224.0.0.5	OSPF	Hello Packet
62	42.859000	192.168.12.1	224.0.0.2	LDP	Hello Message
63	45.203000	192.168.12.2	224.0.0.2	LDP	Hello Message
64	47.906000	192.168.12.1	224.0.0.2	LDP	Hello Message
65	50.203000	192.168.12.2	224.0.0.2	LDP	Hello Message

```
9 6.563000 192.168.12.2 224.0.0.2 LDP Hello Message
Frame 9: 76 bytes on wire (608 bits), 76 bytes captured (608 bits)
Ethernet II, Src: HuaweiTe_a4:59:82 (54:89:98:a4:59:82), Dst: IPv4mcast_00:00:02 (01:
Internet Protocol, Src: 192.168.12.2 (192.168.12.2), Dst: 224.0.0.2 (224.0.0.2)
User Datagram Protocol, Src Port: ldp (646) Dst Port: ldp (646)
Label Distribution Protocol
  Version: 1
  PDU Length: 30
  LSR ID: 2.2.2.2 (2.2.2.2)
  Label Space ID: 0
  Hello Message
    0... .. = U bit: Unknown bit not set
    Message Type: Hello Message (0x100)
    Message Length: 20
    Message ID: 0x0000015c
    Common Hello Parameters TLV
      00.. .... = TLV Unknown bits: Known TLV, do not Forward (0x00)
      TLV Type: Common Hello Parameters TLV (0x400)
      TLV Length: 4
      Hold Time: 15
      0... .. = Targeted Hello: Link Hello
      .0.. .... = Hello Requested: Source does not request periodic hellos
      ..00 0000 0000 0000 = Reserved: 0x0000
    IPv4 Transport Address TLV
      00.. .... = TLV Unknown bits: Known TLV, do not Forward (0x00)
      TLV Type: IPv4 Transport Address TLV (0x401)
      TLV Length: 4
      IPv4 Transport Address: 2.2.2.2 (2.2.2.2)
```

LDP 四类消息：

发现（Discovery）消息：用于通告和维护网络中邻居的存在，如 Hello 消息。

会话（Session）消息：用于建立、维护和终止 LDP 对等体之间的会话，如 Initialization 消息、Keepalive 消息。

通告（Advertisement）消息：用于创建、改变和删除 FEC 的标签映射，如 Address 消息、Label Mapping 消息。

通知（Notification）消息：用于提供建议性的消息和差错通知。

LDP 5 种会话状态：

Non-Existent，Initialized，OpenRec，OpenSent，Operational

Non-Existent 状态：该状态为 LDP Session 最初的状态，在此状态双方发送 HELLO 消息，选举主动方，在收到 TCP 连接建立成功事件的触发后变为 Initialized 状态。

Initialized 状态：该状态下分为主动方和被动方两种情况，主动方将主动发送 Initialization 消息，转向 OpenSent 状态，等待回应的 Initialization 消息；被动方在此状态等待主动方发给自己的 Initialization 消息，如果收到的 Initialization 消息的参数可以接受，则发送 Initialization 和 KeepAlive 转向 OpenRec 状态。主动方和被动方在此状态下收到任何非 Initialization 消息或等待超时，都会转向 Non-Existent 状态。

OpenSent 状态：此状态为主动方发送 Initialization 消息后的状态，在此状态等待被动方回答 Initialization 消息和 KeepAlive 消息，如果收到的 Initialization 消息中的参数可以接受则转向 OpenRec 状态，如果参数不能接受或 Initialization 消息超时则断开 TCP 连接转向 Non-Existent 状态。

OpenRec 状态：在此状态不管主动方还是被动方都是发出 KeepAlive 后的状态，在等待对方回应 KeepAlive，只要收到 KeepAlive 消息就转向 Operational 状态；如果收到其它消息或 KeepAlive 超时则转向 Non-Existent 状态。

Operational 状态：该状态是 LDP Session 成功建立的标志。在此状态下可以发送和接收所有其它的 LDP 消息。在此状态如果 KeepAlive 超时或收到致命错误的 Notification 消息（Shutdown 消息）或者自己主动发送 Shutdown 消息主动结束会话，都会转向 Non-Existent 状态。

标签发布方式：DU（Downstream Unsolicited，下游自主方式）

Label Advertisement Mode

标签的分配控制方式：有序标签分配控制方式

标签

的保持方式：自由标签保持方式

Label Distribution Mode：Ordered

Lab

el Retention Mode : Liberal

标签的发布方式：华为采用 DU 下游自主

DU (Downstream Unsolicited , 下游自主方式) :

对于一个到达同一目地址报文的分组 , LSR 无需从上游获得标签请求消息即可进行标签分配与分发。

DoD (Downstream on Demand , 下游按需方式) :

对于一个到达同一目的地址报文的分组 , LSR 获得标签请求消息之后才进行标签分配与分发。

标签的分配控制方式：华为采用 Ordered 有序标签分配控制

Independent (独立标签分配控制方式) : 本地 LSR 可以自主地分配一个标签绑定到某个 IP 分组 , 并通告给上游 LSR , 而无需等待下游的标签。

Ordered (有序标签分配控制方式) : 只有当该 LSR 已经具有此 IP 分组的下一跳的标签 , 或者该 LSR 就是该 IP 分组的出节点时 , 该 LSR 才可以向上游发送此 IP 分组的标签。

标签的保持方式：华为采用 Liberal 自由标签保持

Liberal (自由标签保持方式) : 对于从邻居 LSR 收到的标签映射 , 无论邻居 LSR 是不是自己的下一跳都保留。

Conservative (保守标签保持方式) : 对于从邻居 LSR 收到的标签映射 , 只有当邻居 LSR 是自己的下一跳时才保留。

0-15 保留标签的含义

表1 特殊标签

标签值	含义	描述
0	IPv4 Explicit NULL Label	表示该标签必须被弹出（即标签被剥掉），且报文的转发必须基于IPv4。如果出节点分配给倒数第二跳节点的标签值为0，则倒数第二跳LSR需要将值为0的标签正常压入报文标签值顶部，转发给最后一跳。最后一跳发现报文携带的标签值为0，则将标签弹出。0标签只有出现在栈底时才有效。
1	Router Alert Label	只有出现在非栈底时才有效。类似于IP报文的“Router Alert Option”字段，节点收到Router Alert Label时，需要将其送往本地软件模块进一步处理。实际报文转发由下一层标签决定。如果报文需要继续转发，则节点需要将Router Alert Label压回标签栈顶。
2	IPv6 Explicit NULL Label	表示该标签必须被弹出，且报文的转发必须基于IPv6。如果出节点分配给倒数第二跳节点的标签值为2，则倒数第二跳节点需要将值为2的标签正常压入报文标签值顶部，转发给最后一跳。最后一跳发现报文携带的标签值为2，则直接将标签弹出。2标签只有出现在栈底时才有效。
3	Implicit NULL Label	倒数第二跳LSR进行标签交换时，如果发现交换后的标签值为3，则将标签弹出，并将报文发给下最后一跳。最后一跳收到该报文直接进行IP转发或下一层标签转发。
4~13	保留	-
14	OAM Router Alert Label	MPLS OAM (Operation Administration & Maintenance) 通过发送OAM报文检测和通告LSP故障。OAM报文使用MPLS承载。OAM报文对于Transit LSR和倒数第二跳LSR (penultimate LSR) 是透明的。
15	保留	-

LDP 分配标签的空间有两种。

一种是基于接口的标签空间:每个接口通告的标签范围是唯一的，，LER1 为同一条 FEC 在不同接口通告的标签是不同的。
二是基于平台的标签空间:标签分配时并不是在每个接口下唯一，而是从整台 LSR 中来分配标签的， LER1 为同一条 FEC 通告一个标签。（常见）

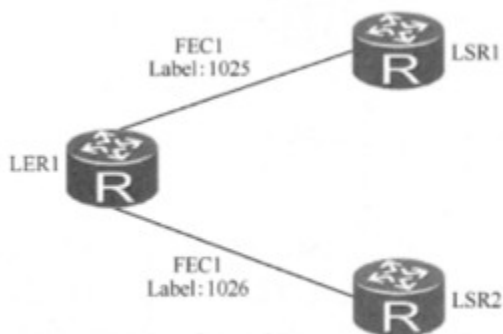


图 8-8 基于接口的标签空间

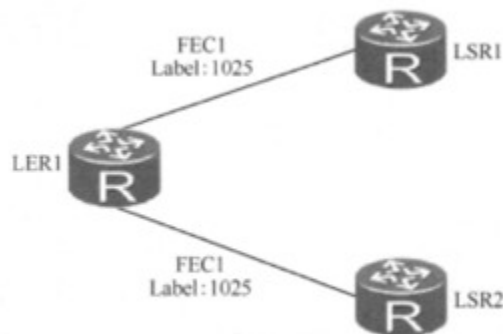


图 8-9 基于平台的标签空间

标签空间

标签空间是什么？设备有几种标签空间？

(1) 标签空间，决定本台设置的入标签如何产生。有基于平台的和基于接口的。

1 基于平台标签空间：

- a) 设备上的所有 FEC 共同使用 1024--- 2^{20} 的标签空间
- b) 标签分配时并不是在每个接口下唯一
- c) 帧模式下使用的标签空间为基于平台

2 基于接口标签空间：

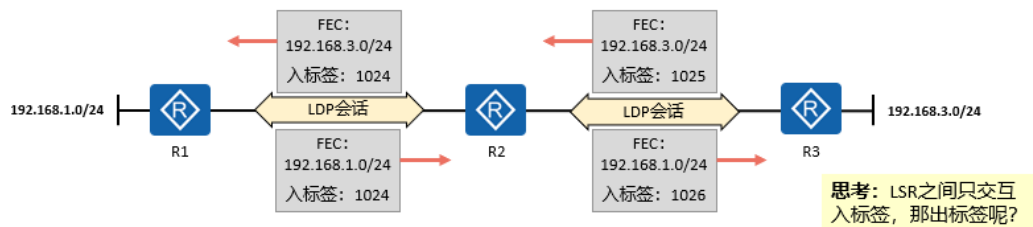
- a) 每个接口通告的标签范围是唯一的
- b) LSR 为同一条 FEC 在不同接口通告的标签是不同的 (小概率会相同)

=====

- MPLS 是一种根据标签报文中携带的短而定长的标签来转发数据的技术。
- MPLS 的一个基本概念就是两台 LSR 必须对在它们之间转发的数据的标签使用上“达成共识”。LSR 之间可以运行标签分发协议 (Label Distribution Protocol , LDP) 来告知其他 LSR 本设备上的标签绑定信息，从而实现标签报文的正确转发。
- 本课程将介绍 LDP 基本工作原理与特性，以及 LDP 的基本配置。

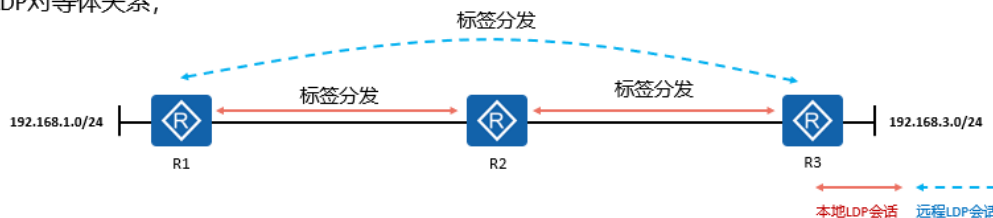
LDP协议概述

- LDP是MPLS的一种控制协议，相当于传统网络中的信令协议，负责FEC的分类、标签的分配以及LSP的建立和维护等操作。LDP规定了标签分发过程中的各种消息以及相关处理过程。
- LDP的工作过程主要分为两部分：
 - LSR之间建立LDP会话。
 - LSR之间基于LDP会话动态交换标签与FEC的映射信息，并根据标签信息建立LSP。



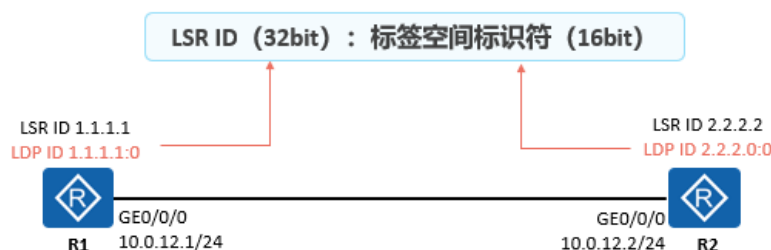
LDP会话、LDP邻接体、LDP对等体

- LSR之间交互标签绑定消息之前必须建立LDP会话。LDP会话可以分为：
 - 本地LDP会话 (Local LDP Session)：建立会话的两个LSR之间是直连的；
 - 远程LDP会话 (Remote LDP Session)：建立会话的两个LSR之间可以是直连的，也可以是非直连的。
- 两台LSR之间交互Hello消息之后，即建立起邻接体 (Adjacency) 关系；
- 在建立邻接体关系的基础上，两台LSR之间交互LDP会话消息，建立起LDP会话，两台设备之间形成LDP对等体关系；



LSR ID与LDP ID

- 每一台运行了LDP的LSR除了必须配置LSR ID，还必须拥有LDP ID。
 - LDP ID的长度为48bit，由32bit的LSR ID与16bit的标签空间标识符（Label Space ID）构成。
 - LDP ID以“LSR ID: 标签空间标识”的形式呈现。例如2.2.2.2:0。
- 标签空间标识一般存在两种形态：
 - 值为0：表示基于设备（或基于平台）的标签空间；
 - 值非0：表示基于接口的标签空间。



- 在本课程中，均使用基于设备（或基于平台）的标签空间。

LDP消息

运行LDP协议的LSR之间通过交换LDP消息来实现邻居发现、会话建立与维护以及标签管理等功能。

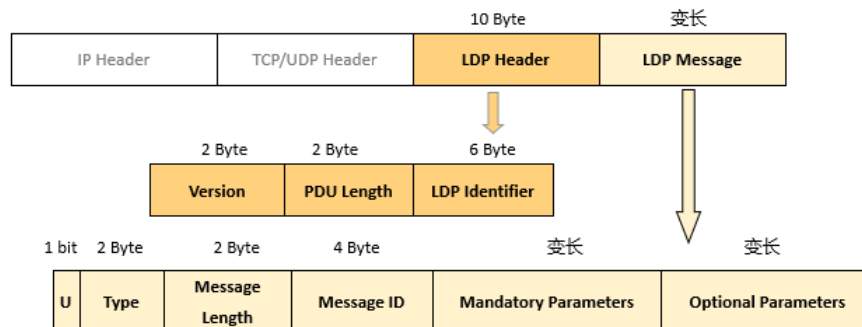
消息类型	消息名称	传输层协议	作用
Discovery Message	Hello	UDP	LDP发现机制中宣告本LSR并发现邻居
Session Message	Initialization	TCP	在LDP Session建立过程中协商参数
	KeepAlive		监控LDP Session的TCP连接的完整性
Advertisement Message	Address		宣告接口地址
	Address Withdraw		撤销接口地址
	Label Mapping		宣告FEC/Label映射信息
	Label Request		请求FEC的标签映射
	Label Abort Request		终止未完成的Label Request Message
	Label Withdraw		撤销FEC/Label映射
	Label Release		释放标签
Notification Message	Notification		通知LDP Peer错误信息

- 按照消息的功能，LDP 消息一共可以分为四大类型：发现消息（Discovery Message），会话消息（Session Message），通告消息（Advertisement Message）和通知消息（Notification Message）。

- 发现消息：用来宣告和维护网络中一个 LSR 的存在；用于通告和维护网络中 LSR 的存在，如 Hello 报文。
 - 会话消息：用于建立、维护和终止 LDP 对等体之间的会话，如 Initialization 报文、KeepAlive 报文。
 - 通告消息：用来生成、改变和删除 FEC 的标签映射；
 - 通知消息：用来宣告告警和错误信息。
 - LDP 消息承载在 UDP 或 TCP 之上，端口号均为 646。
- 其中发现消息用来发现邻居，承载在 UDP 报文上。其他消息的传递要求可靠而有序，所以 LDP 使用 TCP 建立会话，会话、通告和通知消息都基于 TCP 传递。

LDP报文封装

- LDP协议报文包括了LDP头部和LDP消息两部分。
 - LDP头部中携带了LDP版本、报文长度等信息；
 - LDP消息中携带了消息类型、消息长度等信息。

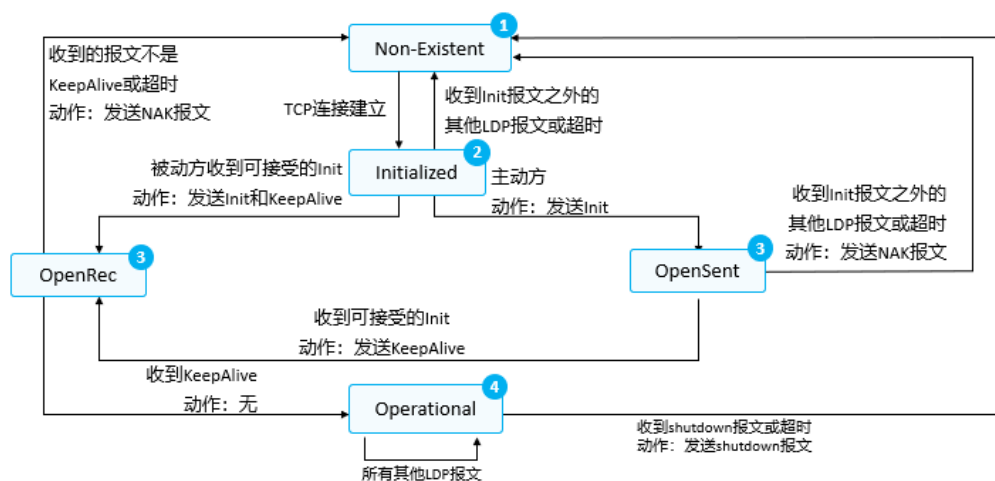


- LDP 头部长度为 10Byte，包括 Version，PDU Length 和 LDP Identifier 三部分。
- Version 占用 2Byte，表示 LDP 版本号，当前版本号为 1。
- PDU Length 占用 2Byte，以字节为单位表示除了 Version 和 PDU Length 以外的其他部分的总长度。
- LDP Identifier，即 LDP ID，长度 6Byte，其中前 4Byte 用来唯一标识一个 LSR，后 2Byte 用来表示 LSR 的标签空间。
- LDP 消息包含五个部分。

- U 占用 1 个 bit，为 Unknown Message bit。当 LSR 收到一个无法识别的消息时，该消息的 U=0 时，LSR 会返回给该消息的生成者一个通告，当 U=1 时，忽略该无法识别的消息，不发送通告给生成者。
- Message Length 占用 2 个 bytes，以字节为单位表示 Message ID、Mandatory Parameters 和 Optional Parameters 的总长度。
- Message ID 占用 32 个 bit，用来标识一个消息。
- Mandatory Parameters 和 Optional Parameters 分别为可变长的该消息的必须的和可选的参数。
- Message Type 表示具体的消息类型，目前 LDP 定义常用的消息有 Notification，Hello，Initialization，KeepAlive，Address，Address Withdraw，Label Mapping，Label Request，Label Abort Request，Label Withdraw，Label Release。

LDP会话状态机

LDP使用5种状态描述LDP会话状态机。



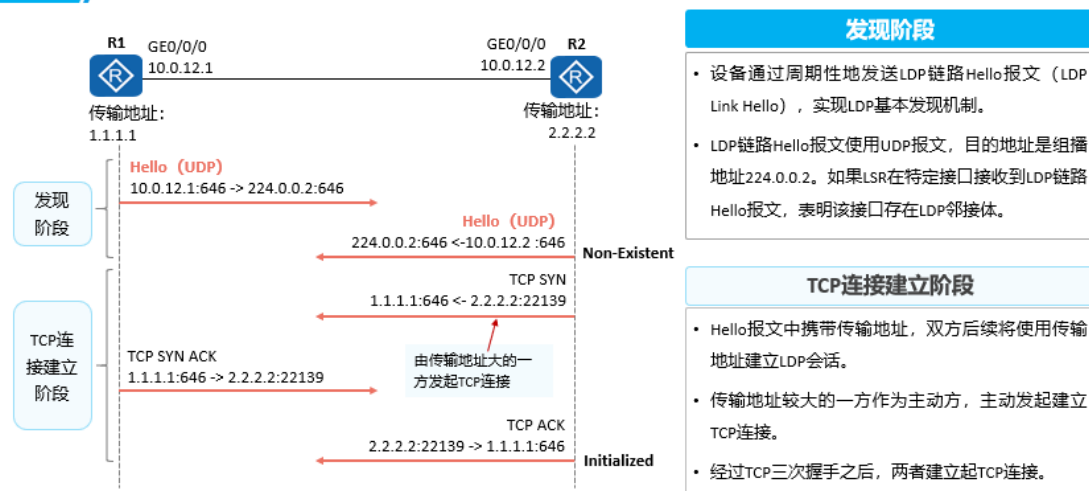
- LDP Session 协商过程可以通过状态机来描述。如图所示，有 5 种状态。分别是 Non-Existent，Initialized，OpenRec，OpenSent，Operational。
- Non-Existent 状态：该状态为 LDP Session 最初的状态，

在此状态双方发送 HELLO 消息，选举主动方，在收到 TCP 连接建立成功事件的触发后变为 Initialized 状态。

- Initialized 状态：该状态下分为主动方和被动方两种情况，主动方将主动发送 Initialization 消息，转向 OpenSent 状态，等待回应的 Initialization 消息；被动方在此状态等待主动方发给自己的 Initialization 消息，如果收到的 Initialization 消息的参数可以接受，则发送 Initialization 和 KeepAlive 转向 OpenRec 状态。主动方和被动方在此状态下收到任何非 Initialization 消息或等待超时，都会转向 Non-Existent 状态。
- OpenSent 状态：此状态为主动方发送 Initialization 消息后的状态，在此状态等待被动方回答 Initialization 消息和 KeepAlive 消息，如果收到的 Initialization 消息中的参数可以接受则转向 OpenRec 状态，如果参数不能接受或 Initialization 消息超时则断开 TCP 连接转向 Non-Existent 状态。
- OpenRec 状态：在此状态不管主动方还是被动方都是发出 KeepAlive 后的状态，在等待对方回应 KeepAlive，只要收到 KeepAlive 消息就转向 Operational 状态；如果收到其它消息或 KeepAlive 超时则转向 Non-Existent 状态。
- Operational 状态：该状态是 LDP Session 成功建立的标志。在此状态下可以发送和接收所有其它的 LDP 消息。在此状态如果 KeepAlive 超时或收到致命错误的 Notification 消息（Shutdown 消息）或者自己主动发送 Shutdown 消息主动结束会话，都会转向 Non-Existent 状态。
- LDP 状态切换信息可以通过指令 **debug mpls ldp session** 看到。

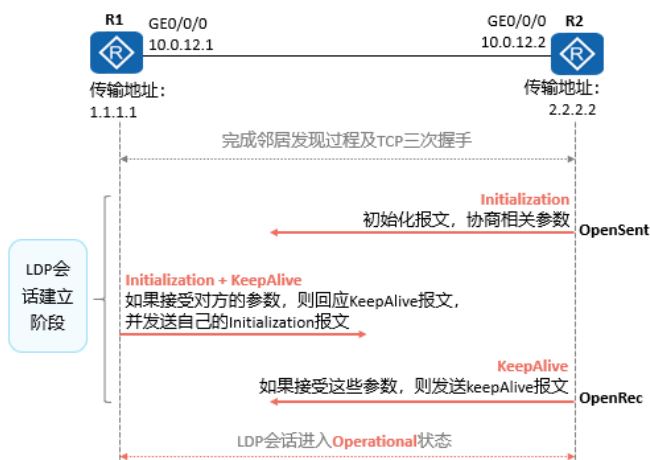


LDP会话建立 - 发现阶段与TCP连接建立



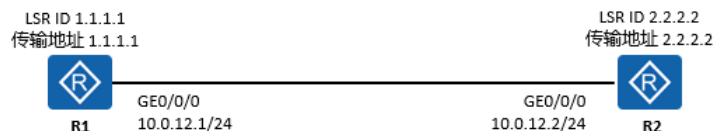
- 除了基本发现机制外，可以通过拓展发现机制发现非直连的远端邻接体，该内容不属于课程的重点，详细内容可以查阅 RFC5036 相关内容。
- LDP 的传输地址用于与邻居建立 TCP 连接。
- 两台 LSR 之间在建立 LDP 会话之前，需要先建立 TCP 连接，以便进行 LDP 协议报文的交换。
- 设备的传输地址被包含在 LDP Hello 报文中，LSR 通过 Hello 报文知晓邻居的传输地址。
- 在使用 Hello 报文发现邻居并且知道了对方的传输地址后，邻居之间就会开始尝试 TCP 三次握手（基于传输地址），并且交互 LDP 的初始化报文、标签映射报文等，这些报文都使用双方的传输地址作为源、目的 IP 地址。
- LSR 必须拥有到达邻居的传输地址的路由。
- 缺省情况下，公网的 LDP 传输地址等于设备的 LSR ID，私网的传输地址等于接口的主 IP 地址。
- 在接口视图下，使用 `mpls ldp transport-address` 命令，可以修改传输地址。

LDP会话建立 - 会话建立与保持



- TCP连接建立成功后, 主动方R2 (传输地址大的一方) 发送LDP初始化报文, 协商建立LDP会话的相关参数。
- LDP会话的相关参数包括LDP协议版本、标签分发方式、KeepAlive保持定时器的值、最大PDU长度和标签空间等。
- 被动方R1收到初始化报文后, 若接受R2的相关参数, 则回应KeepAlive报文作为确认, 为了提高发送效率同时发送自己的初始化报文。
- R2收到R1的初始化报文后, 若接受相关参数, 则回复KeepAlive报文给R1。
- 双方都收到对端的KeepAlive报文后, 会话建立成功。后续通过周期性发送的KeepAlive报文

LDP邻居状态



<R1>display mpls ldp peer

LDP Peer Information in Public network
A '*' before a peer means the peer is being deleted.

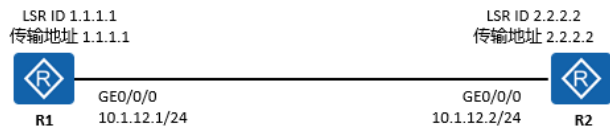
PeerID	TransportAddress	DiscoverySource
2.2.2.2:0	2.2.2.2	GigabitEthernet0/0/0

TOTAL: 1 Peer(s) Found.

- PeerID: LDP邻居的LDP ID;
 - 2.2.2.2代表的是邻居节点的LSR ID;
 - 0代表的是标签空间是基于平台的;
- TransportAddress: LDP邻居的传输地址;
 - 2.2.2.2代表邻居用来建立TCP连接的IP地址。



LDP会话状态



```
<R1>display mpls ldp session
```

LDP Session(s) in Public Network

Codes: LAM(Label Advertisement Mode), SsnAge Unit(DDDD:HH:MM)

A '*' before a session means the session is being deleted.

PeerID	Status	LAM	SsnRole	SsnAge	KASent/Rcv
2.2.2.2:0	Operational DU		Passive	0000:00:33	133/133

TOTAL: 1 session(s) Found.

- Status: LDP会话的状态;

- Operational表示LDP会话建立成功;

- LAM: 标签发布模式:

- 标签发布模式有DU和DoD两种模式(后文介绍);

- 此例中采用的是DU(下游自主)模式;

- SsnRole: LSR在LDP会话中的角色;

- Active表示建立LDP会话的主动方; Passive表示建立LDP会话的被动方;

- LDP 会话的状态：
- NonExistent：表示 LDP 会话的最初状态。在此状态双方互相发送 Hello 消息，在收到 TCP 连接建立成功事件的触发后变为 Initialized 状态。
- Initialized：表示 LDP 会话处于初始化状态。
- Open Sent：表示 LDP 会话进入初始化状态后，主动方给被动方发送了 Initialized 消息，并等待对方的回应。
- Open Recv：表示 LDP 会话进入初始化状态后，当双方都收到了对方发送的 KeepAlive 消息后，LDP 会话进入 Operational 状态。
- Operational：表示 LDP 会话建立成功。



标签的发布和管理

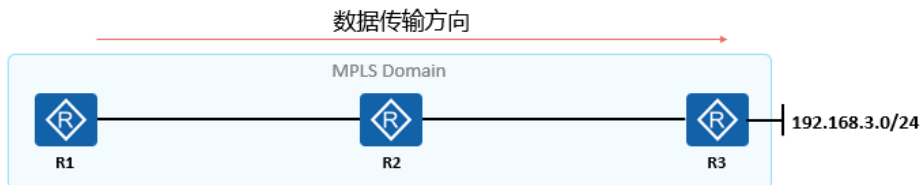
- 在MPLS网络中，下游LSR决定标签和FEC的绑定关系，并将这种绑定关系发布给上游LSR。
- LDP通过发送标签请求和标签映射消息，在LDP对等体之间通告FEC和标签的绑定关系来建立LSP
- 标签的发布和管理**由标签发布方式、标签分配控制方式和标签保持方式来决定。

内容	名称	默认	含义
标签发布方式 (Label Advertisement Mode)	下游自主方式 (Downstream Unsolicited, DU)	是	对于一个特定的FEC，LSR无需从上游获得标签请求消息即进行标签分配与分发。
	下游按需方式 (Downstream on Demand, DoD)	否	对于一个特定的FEC，LSR获得标签请求消息之后才进行标签分配与分发。
标签分配控制方式 (Label Distribution Control Mode)	独立方式 (Independent)	是	本地LSR可以自主地分配一个标签绑定到某个FEC，并通告给上游LSR，而无需等待下游的标签。
	有序方式 (Ordered)	否	对于LSR上某个FEC的标签映射，只有当该LSR已经具有此FEC下一跳的标签映射消息、或者该LSR就是此FEC的出节点时，该LSR才可以向上游发送此FEC的标签映射。
标签保持方式 (Label Retention Mode)	自由方式 (Liberal)	是	对于从邻居LSR收到的标签映射，无论邻居LSR是不是自己的下一跳都保留。
	保守方式 (Conservative)	否	对于从邻居LSR收到的标签映射，只有当邻居LSR是自己的下一跳时才保留。



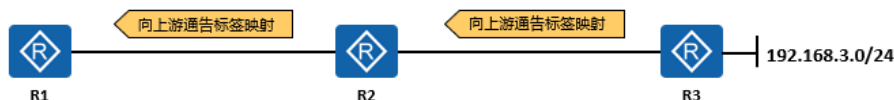
上游与下游

- MPLS根据数据的转发方向确定上、下游关系。标签报文从上游LSR发出，被下游LSR接收并处理。
- 如图所示，对于到达192.168.3.0/24的LSP而言，R3是R2的下游LSR，R1是R2的上游LSR。



标签发布方式 - DU模式

- DU模式
 - 对于一个特定的FEC，LSR无需从上游获得标签请求消息即进行标签分配与分发。
 - LSR会主动将自己为FEC捆绑的标签通告给上游邻居，无需邻居先发起请求再通告。

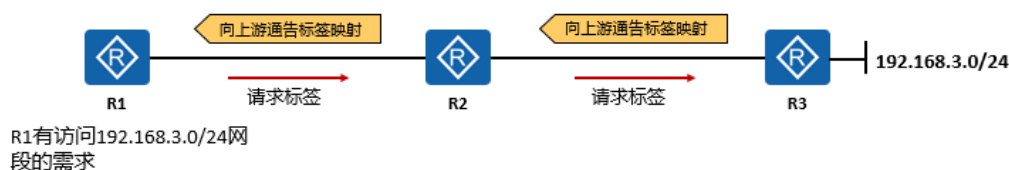


- 标签分配：**LSR 从本地标签空间中取出一个标签与 FEC 绑定。

- 标签分发：LSR 将标签与 FEC 的绑定关系通知给上游 LSR。
- 标签发布方式为 DU 时，系统默认支持 LDP 为所有对等体分标签，即每个节点都可以向所有的对等体发布标签映射关系，不再区分上下游关系。因为在只给上游对等体分标签情况下，发送标签映射消息的时候，要根据路由信息对会话的上下游关系进行确认。

标签发布方式 - DoD模式

- DoD模式
 - 对于一个特定的FEC，LSR获得标签请求消息之后才进行标签分配与分发。
 - 一般情况下，对特定FEC的访问需求会触发标签请求消息。

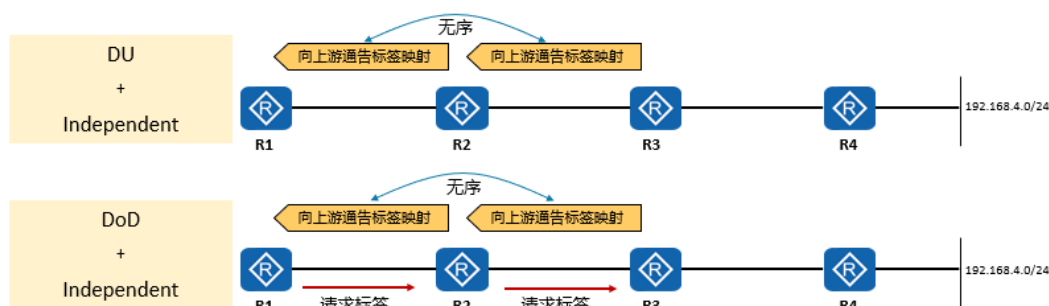


- 只有上游邻居向自己请求标签映射时，LSR 才会通告标签映射给该邻居

标签分配控制方式 - 独立模式

独立 (Independent) 模式

- 本地LSR可以自主地分配一个标签绑定到某个FEC，并通告给上游LSR，而无需等待下游的标签。

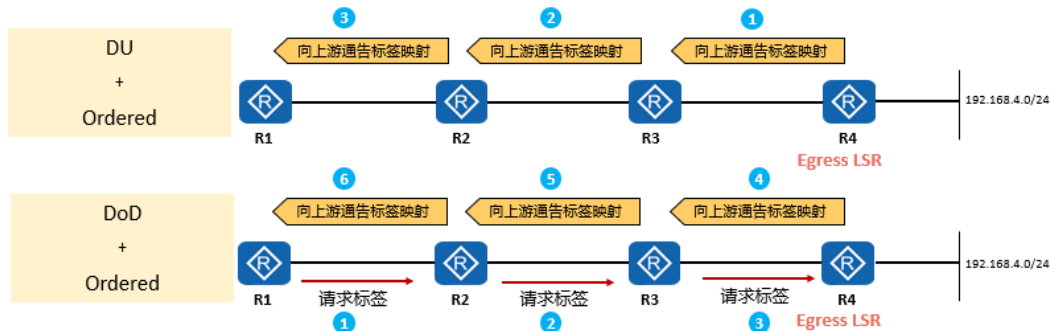


- 标签分配控制方式需要与标签发布方式结合使用：
- 在使用 DU 作为标签分发方式的情况下，如图所示，R2 和 R3 对 192.168.4.0/24 这条 FEC，可以在上游 LSR 无请求，且自身没有收到下游 LSR 的标签绑定信息的情况下，主动向上游 LSR 通告标签绑定信息。
- 采用 DoD 作为标签发布方式时，如图所示，R2 和 R3 对 192.168.4.0/24 这条 FEC，只要收到上游 LSR 的标签请求消息，可以在自身没有收到下游 LSR 的标签绑定信息的情况下，向上游 LSR 通告标签绑定信息。

标签分配控制方式 - 有序模式

- 有序 (Ordered) 模式

- 对于 LSR 上某个 FEC 的标签映射，只有当该 LSR 已经具有此 FEC 下一跳的标签映射消息、或者该 LSR 就是此 FEC 的出节点时，该 LSR 才可以向上游发送此 FEC 的标签映射。



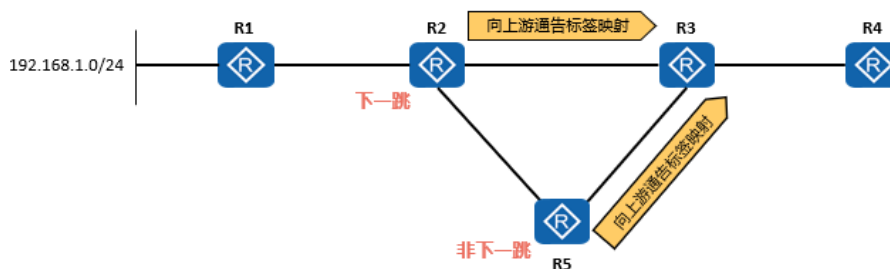
- 当标签控制方式为 Ordered，只有当 LSR 收到特定 FEC 下一跳发送的特定 FEC 标签映射消息或者 LSR 是 LSP 的出口节点时，LSR 才可以向上游发送标签映射消息。
- 当标签分发方式为 DU 时，如图所示，对于 192.168.4.0/24 这条 FEC，不论上游 LSR 是否有请求，必须收到下游 LSR 对此 FEC 的标签绑定信息才向上游 LSR 发布标签绑定信息，所以必须由 Egress LSR，也就是 R4 作为 LSP 建立的“起点”。
- 当标签发布方式采用 DoD 时，如图所示，对于 192.168.4.0/24 这条 FEC，只有收到上游 LSR 请求的请求，且自身已经收到下游 LSR 的标签绑定信息的情况下，才向上游 LSR 通

告标签绑定信息。因此，必须由 Ingress LSR (R1) 发起请求，逐跳请求到 Egress LSR (R4)，最终由 R4 开始，向上游建立 LSP。

标签保留 - 自由模式

自由 (Liberal) 模式

- LSR收到的标签映射可能来自下一跳，也可能来自非下一跳。
- 对于从邻居LSR收到的标签映射，无论邻居LSR是不是自己的下一跳都保留。

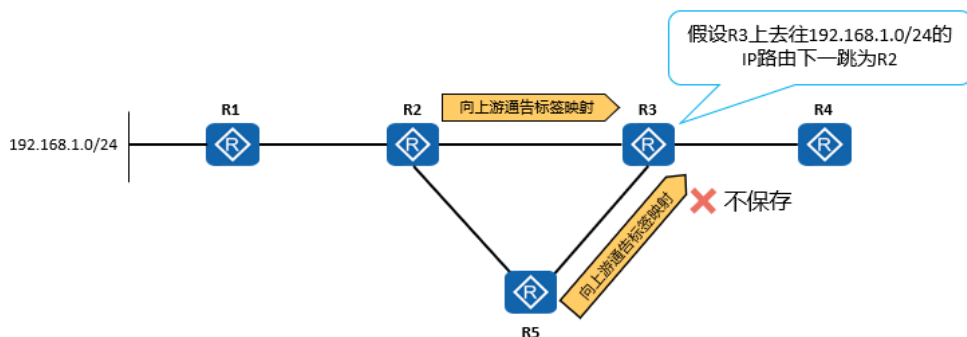


- 当基于 IP 网络部署 MPLS 时，LSR 根据 IP 路由表判断接收到的标签映射是否来自下一跳。
- Liberal 方式的最大优点在于路由发生变化时能够快速建立新的 LSP 进行数据转发，因为 Liberal 方式保留了所有的标签。缺点是需要分发和维护不必要的标签映射。
- DU 标签分发方式下，如果采用 Liberal 保持方式，则 R3 保留所有 LDP 邻居 R2 和 R5 发来的关于 192.168.1.0/24 这个 FEC 的标签，无论该 R2 和 R5 是否是 IP 路由表中到达 192.168.1.0/24 的下一跳。
- DoD 标签分发方式下，如果采用 Liberal 保持方式，LSR 会向所有 LDP 邻居请求标签。但通常来说，DoD 分发方式都会和 Conservative 保持方式搭配使用。

标签保留 - 保守模式

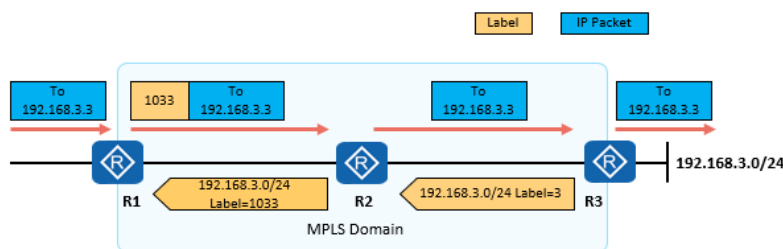
- 保守 (Conservative) 模式

- 对于从邻居LSR收到的标签映射，只有当邻居LSR是自己的下一跳时才保留。



- Conservative 方式的优点在于只需保留和维护用于转发数据的标签，以达到节约标签的目的。
- 当使用 DU 标签分发方式时，LSR 可能从多个 LDP 邻居收到到同一网段的标签映射消息，如图中 R3 会分别从 R2 和 R5 收到网段 192.168.1.0/24 的标签映射消息。如果采用 Conservative 保持方式，则 R3 只保留下一跳 R2 发来的标签，丢弃非下一跳 R5 发来的标签。
- 当使用 DoD 标签分发方式时，LSR 根据路由信息只向它的下一跳请求标签。
- 当网络拓扑变化引起下一跳邻居改变时：
- 使用自由标签保持方式，LSR 可以直接利用原来非下一跳邻居发来的标签，迅速重建 LSP，但需要更多的内存和标签空间。
- 使用保守标签保持方式，LSR 只保留来自下一跳邻居的标签，节省了内存和标签空间，但 LSP 的重建会比较慢。
- 保守标签保持方式通常与 DoD 方式一起，用于标签空间有限的 LSR。

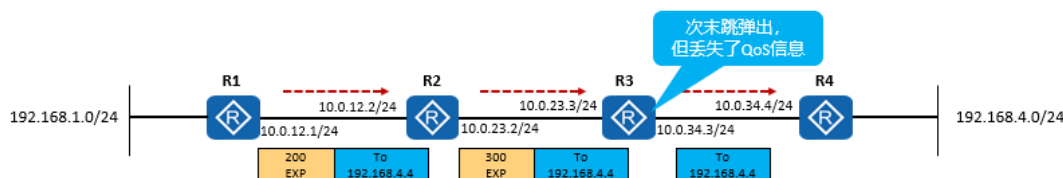
PHP特性



- PHP (Penultimate Hop Popping, 次末跳弹出), 如果激活了PHP特性, 那么egress节点在为本地路由分配标签的时候, 会分配一个特殊标签—3, 该标签被称为隐式空标签 (Implicit NULL Label)。当LSR转发一个标签报文时, 如果发现对应的出标签值为3, 则LSR会将栈顶标签弹出, 并将里面所封装的数据转发给下游LSR。
- 在标签发布时, R3 为作为 192.168.3.0/24 这条 FEC 的 Egress LSR。分配标签时, R3 为该 FEC 分配了标签 3, 并将该标签绑定信息通告给 R2。
- 在数据转发时, R2 作为到达 192.168.3.0 的次末跳 (倒数第二跳), 发现出标签值为 3, 于是将标签头部弹出, 将 IP 报文转发给 R3, 而 R3 则仅需执行一次查询操作 (查询 FIB 表) 即可获得相应的转发信息, 转发效率得到了提升。

隐式空标签与显式空标签 (1)

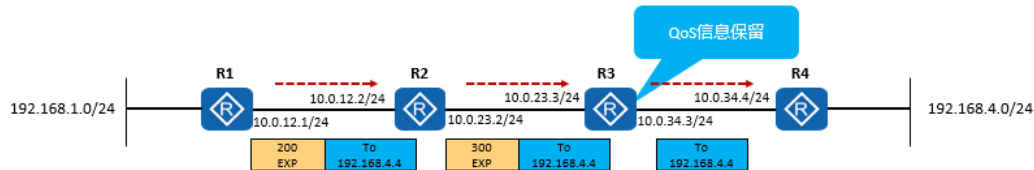
- 缺省情况下, Egress节点向倒数第二跳分配隐式空标签 (implicit-null), 即特殊标签3。
- 但在部署Qos的场景下, 标签被弹出后, 其中的优先级也会一并丢失。





隐式空标签与显式空标签 (2)

- 显式空标签机制，Egress节点向倒数第二跳分配特殊标签0。
- R3在转发标签报文时，若出标签封装为0，则不会将标签头部弹出，标签头部中的QoS信息得以保存。R4在收到带0标签的报文的时候，直接弹出标签，不用去查找ILM表项。
- 缺省情况下，Egress分配的是隐式空标签，通过label advertise explicit-null使能Egress节点向倒数第二跳分配显式空标签。

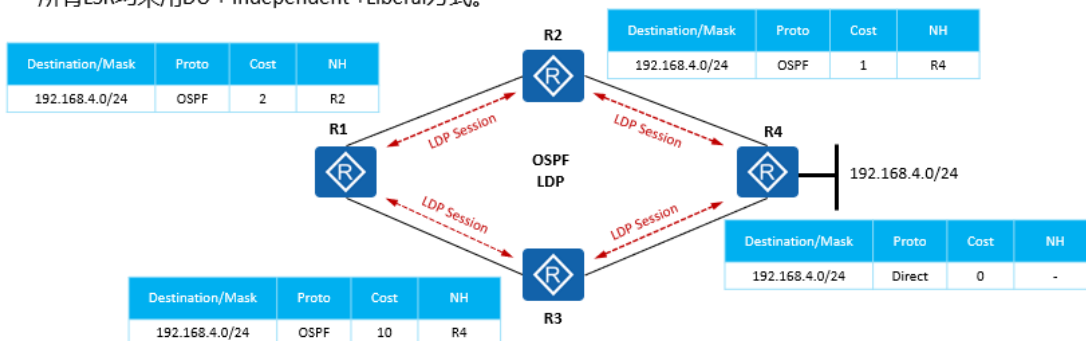


- 在 MPLS 视图下，执行命令 **label advertise { explicit-null | implicit-null | non-null }**，配置向倒数第二跳分配的标签。
- 根据参数的不同，可以配置 Egress 向倒数第二跳分配不同的标签。
- 缺省情况下，使用的是 **implicit-null**，Egress 向倒数第二跳节点分配隐式空标签，值为 3。
- 如果配置的是 **explicit-null**，Egress 节点向倒数第二跳分配显式空标签，值为 0。当需要支持 MPLS QoS 属性时，可以选用 **explicit-null**。
- 如果配置的是 **non-null**，Egress 向倒数第二跳正常分配标签，即分配的标签值不小于 16。



组网介绍

- 网络中已经部署OSPF路由协议且各设备之间能够正常学习到对方的路由信息。
- 已在各设备及相应接口上激活MPLS及LDP，且在相邻的设备之间已正常建立本地LDP会话。
- 所有LSR均采用DU + Independent + Liberal方式。

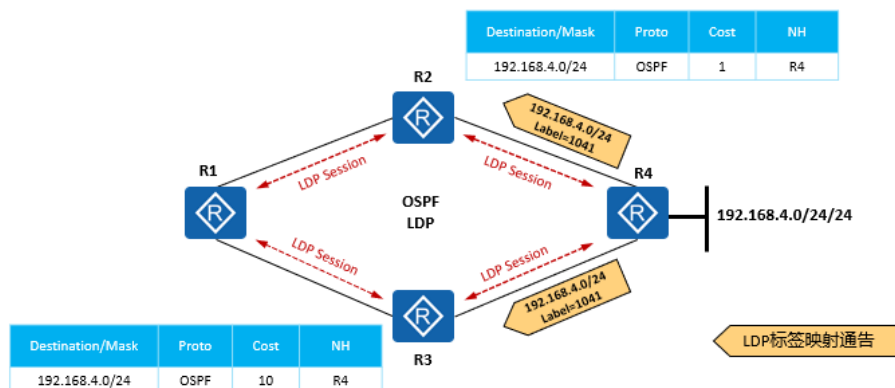


- 华为设备目前缺省模式为下游自主方式（DU）+ 有序标签分配控制方式（Ordered）+ 自由标签保持方式（Liberal）。
- 对于从R1进入，到达192.168.4.0/24的数据，R1为Ingress LSR，R4为Egress LSR。



标签分发 - Egress LSR

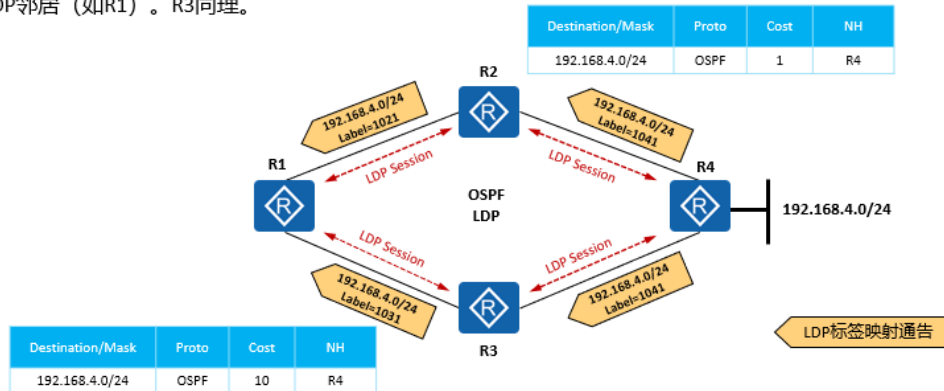
- R4直连网段192.168.4.0/24，R4将主动为到达该网段的路由分配标签，如1041，并主动通过LDP协议报文将标签映射通告给LDP对等体R2和R3。



- 注：缺省情况下，根据32位的主机IP路由触发LDP建立LSP。可以通过手工配置触发非32位路由的LSP建立。

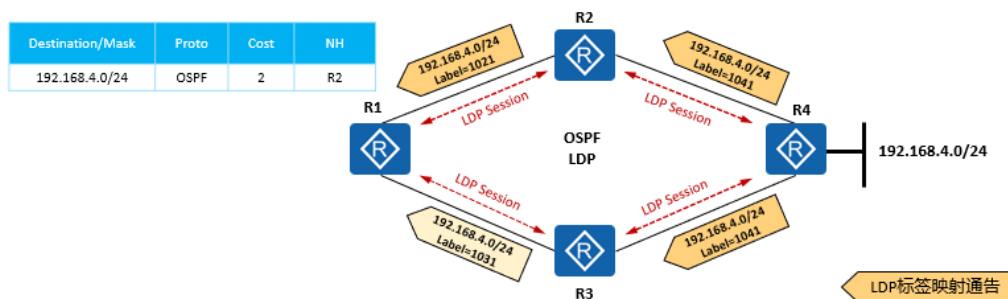
标签分发 - Transit LSR

- 以R2为例，在其路由表中，192.168.4.0/24路由的下一跳为R4，当它从R4收到关于192.168.4.0/24的标签映射通告时，由于该通告来自下游LDP邻居，因此这将触发它自己为该路由分配标签1021，并将标签映射通告给LDP邻居（如R1）。R3同理。



标签分发 - Ingress LSR

- R1收到LDP邻居R2及R3通告过来的关于192.168.4.0/24路由的标签映射后，将这两个标签都**存储**起来，但是由于在自己的路由表中，到达192.168.4.0/24的下一跳是R2，因此当前它只会**使用**R2所通告的标签1021。



- 注：当 R2 发生故障时，OSPF 路由将会重新收敛，此时 R1 的路由表中 192.168.4.0/24 路由的下一跳将会切换至 R3，此时 R1 将启用 R3 所通告的、关于 192.168.4.0/24 的标签。

标签转发 - Ingress LSR

R1作为Ingress LSR，需要对接收的IP报文执行Push操作压入标签，并进行标签转发。

R1的路由表

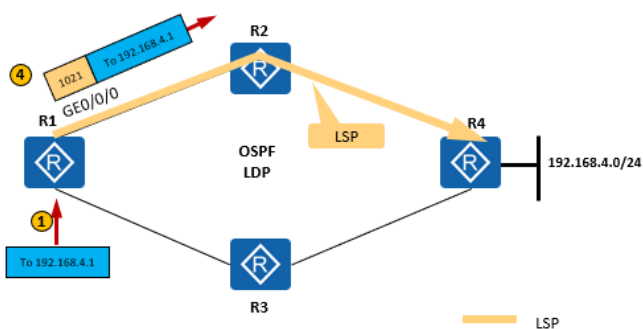
Destination/Mask	Proto	Cost	NH
192.168.4.0/24	OSPF	2	R2

R1的FIB表 ②

Destination/Mask	Proto	Cost	NH	Tunnel ID
192.168.4.0/24	OSPF	2	R2	0x12

R1的NHLFE ③

Tunnel ID	Out intf	OPER	NH	Out Label
0x12	GE0/0/0	push	R2	1021



- 当 R1 收到发往 192.168.4.1 的 IP 报文时，首先在其 FIB 表中查询该目的 IP 地址，它发现所匹配的表项的 Tunnel ID 为非 0，因此继续在 NHLFE 中查询该 Tunnel ID，然后意识到需要将该 IP 报文压入标签并进行标签转发，出接口为 GE0/0/0、下一跳为 R2、出站标签为 1021，于是为报文插入标签头部并转发出去。

标签转发 - Transit LSR

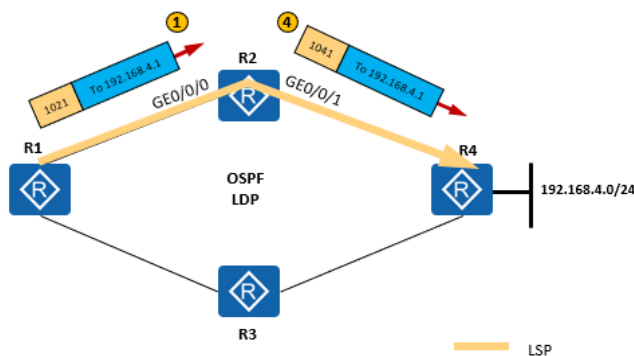
- R2作为Transit LSR，需要对接收的IP报文执行Swap操作交换标签，并进行标签转发。

R2的ILM表 ②

In Label	Tunnel ID
1021	0x12

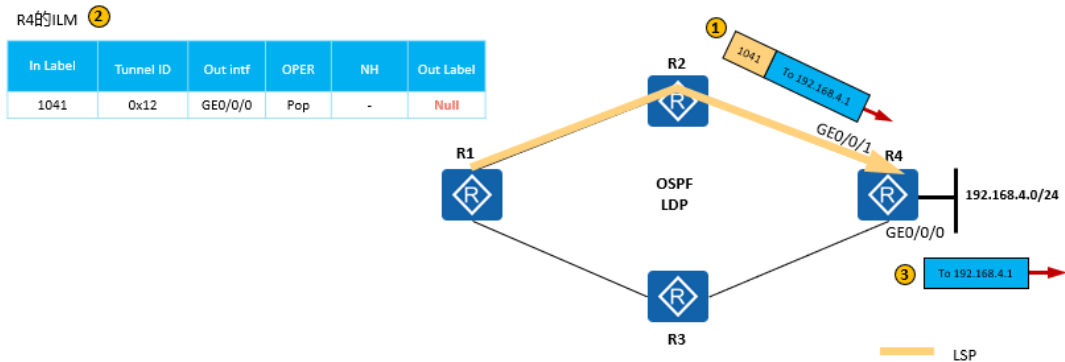
R2的NHLFE ③

Tunnel ID	Out intf	OPER	NH	Out Label
0x12	GE0/0/1	Swap	R4	1041



- 当 R2 收到携带 1021 标签的标签报文时，查询 ILM，根据 ILM 对应到 NHLFE 中的表项。于是，R2 对该标签报文通过 swap 操作将标签更换为 1041，并从相应的接口转发出去。

R4作为Egress LSR, 需要对接收的IP报文执行Pop操作交换标签, 并进行IP转发。



- 当 R4 收到携带 1041 标签的标签报文时，查询 ILM，根据 ILM 查询到操作为 Pop。于是，R4 对该标签报文通过 Pop 操作将最外层标签剥离，此时该报文已经变成了标准 IP 报文，R4 将对该 IP 报文执行标准的 IP 转发流程。
- R4 在转发该报文时分别查询了 LFIB 和 FIB 表，作为最后 Egress LSR，其存在转发效率提升的可能性，怎么做？

在 MPLS 中，运行 LDP 协议的 LSR 的操作小结

- LSR 首先通过运行 IGP 协议（例如 OSPF、IS-IS 等）来构建路由表、FIB 表；
- LDP 根据相应的模式，为路由表中的路由前缀（FEC）分配标签；
- LDP 根据相应的模式，将自己为路由前缀分配的标签，通过 LDP 标签映射报文通告给 LDP 邻居；
- LSR 将自己为路由前缀分配的标签，以及 LDP 邻居为该路由前缀通告的标签存储起来，并与出接口、下一跳地址等信息形成关联（标签转发表项）；
- 当 LSR 转发到达目的网络的标签报文时，所使用的出站

标签总是下游 LDP 邻居所通告的标签，此处所指的下游邻居，是设备的路由表中到达该目的网络的下一跳设备。

LDP基本配置命令 (1)

1. 使能LDP

```
[Huawei] mpls ldp
```

mpls ldp命令用来使能本节点的LDP能力，并进入LDP视图。

```
[Huawei-GigabitEthernet0/0/0] mpls ldp
```

在接口视图下，使能当前接口的LDP功能。需先使能全局LDP能力后才能执行接口下的LDP使能命令。

2. 配置LDP远端会话

```
[Huawei] mpls ldp remote-peer remote-peer-name
```

mpls ldp remote-peer命令用来创建远端对等体并进入远端对等体视图。

```
[Huawei-mpls-ldp-remote-PeerName] remote-ip ip-address
```

remote-ip命令用来配置LDP远端对等体的IP地址。

●

LDP基本配置命令 (2)

3. 配置触发LSP的策略

```
[Huawei-mpls] lsp-trigger { all | host | ip-prefix ip-prefix-name | none }
```

lsp-trigger命令用来指定哪些静态路由及IGP路由会触发LDP LSP的建立，缺省情况下，根据32位地址的IP路由触发LDP建立LSP。

- **all**: 所有静态路由和IGP路由项触发建立LSP。配置该命令后，所有静态路由及IGP路由会触发LDP建立LSP，会导致LSP数量庞大，占用过多的标签资源，以及整网LSP收敛速度变慢。故一般情况下**不推荐**配置此命令。
- **host**: 32位地址的IP路由触发建立LSP。
- **ip-prefix ip-prefix-name**: 根据IP地址前缀列表触发建立LSP。
- **None**: 不触发建立LSP。

4. 配置LDP标签发布模式

```
[Huawei-GigabitEthernet0/0/0] mpls ldp advertisement { dod | du }
```

缺省情况下，标签发布模式为下游自主标签分发（Downstream Unsolicited）。

- 如果标签分配方式为DU，则标签保持模式为Liberal。
- 如果标签分配方式为DOD，则标签保持模式为Conservative。

● BGP 路由也可以触发 LDP LSP 的建立，但此部分内容不在本课程的讨论范围内。



LDP基本配置命令 (3)

5. 配置LDP标签分配控制方式

```
[Huawei-mpls-ldp] label distribution control-mode { independent | ordered }
```

缺省情况下，LDP标签分配控制方式为有序标签分配控制（Ordered）。

6. 配置PHP特性

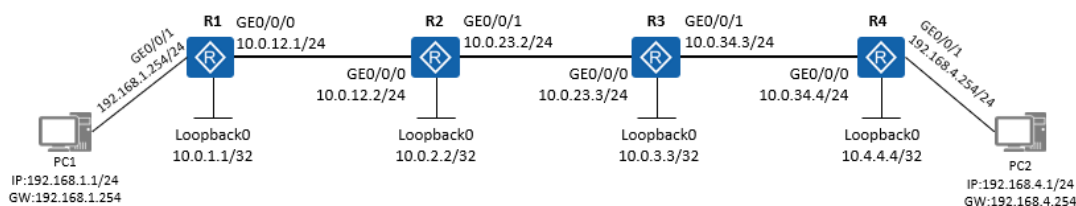
```
[Huawei-mpls] label advertise { explicit-null | implicit-null | non-null }
```

缺省情况下，出节点向倒数第二跳分配隐式空标签（implicit-null）。

- **explicit-null**：出节点向倒数第二跳分配显式空标签。
- **implicit-null**：出节点向倒数第二跳分配隐式空标签。
- **non-null**：出节点向倒数第二跳正常分配标签。



配置案例

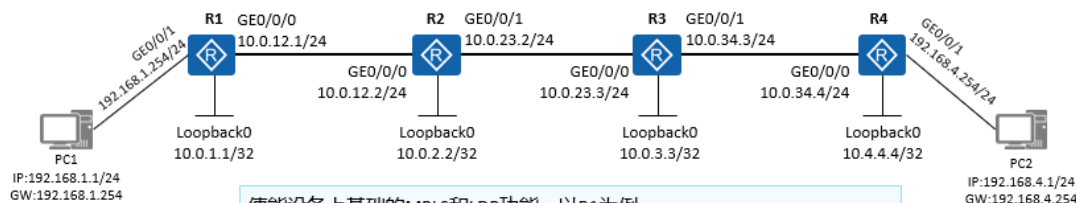


背景：R1、R2、R3和R4之间运行了IGP协议，已经实现了IP层面的互联互通。

需求：通过配置MPLS以及LDP，使得192.168.1.0/24网段和192.168.4.0/24网段可以通过标签交换的方式实现互访。



配置步骤 (1)

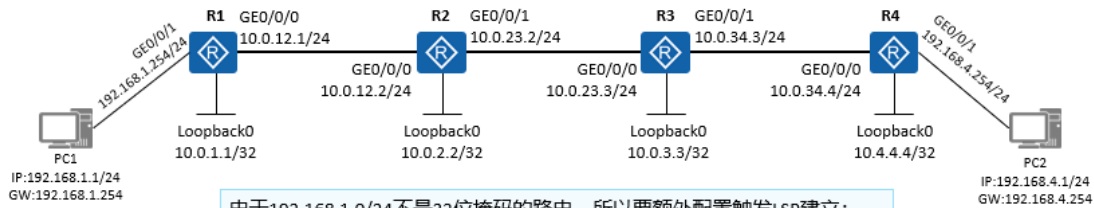


使能设备上基础的MPLS和LDP功能，以R1为例：

```
[R1]mpls
[R1-mpls]quit
[R1]mpls ldp
[R1-mpls-ldp]quit
[R1]interface GigabitEthernet 0/0/0
[R1-GigabitEthernet0/0/0]mpls
[R1-GigabitEthernet0/0/0]mpls ldp
[R1-GigabitEthernet0/0/0]quit
```



配置步骤 (2)



由于192.168.1.0/24不是32位掩码的路由，所以要额外配置触发LSP建立：

```
[R1]ip ip-prefix ldp permit 192.168.1.0 24
[R1-mpls]lsp-trigger ip-prefix ldp
```

同理，R4也需要进行相应的配置

```
[R4]ip ip-prefix ldp permit 192.168.4.0 24
[R4-mpls]lsp-trigger ip-prefix ldp
```



检查配置 - 查看LSP

#查看R1使用LDP创建的LSP相关信息

```
[R1]display mpls ldp lsp
```

LDP LSP Information

DestAddress/Mask	In/OutLabel	UpstreamPeer	NextHop	OutInterface
10.0.2.2/32	1024/3	10.0.2.2	10.0.12.2	GE0/0/0
10.0.3.3/32	1025/1025	10.0.2.2	10.0.12.2	GE0/0/0
192.168.1.0/24	3/NULL	10.0.2.2	192.168.1.254	GE0/0/1
*192.168.1.0/24	Liberal/1027		DS/10.0.2.2	
192.168.4.0/24	1027/1028	10.0.2.2	10.0.12.2	GE0/0/0

思考题：

- （ 单选 ） 以下哪条命令用于查看已经为特定 FEC 分配的标签？（ ）
 - A、display mpls ldp
 - B、display mpls ldp interface
 - C、display mpls lsp
 - D、display mpls ldp session
- （ 单选 ） 华为设备默认的标签发布方式、标签分配控制方式和标签保持方式的组合是（ ）

A、 DU + Independent + Conservative

B、 DU + Ordered+ Liberal

C、 DoD + Independent+ Liberal

D、 DoD + Ordered + Conservative

参考答案：

- C
- B
-