

# IS-IS 技术笔记

## 红茶三杯 CCIE 学习文档

文档版本： 2.0

更新时间： 2013-03-20

文档作者： 红茶三杯

文档地址： <http://ccietea.com>

文档备注： 请关注文档版本及更新时间

## 1 基础知识

### 1.1 协议概述

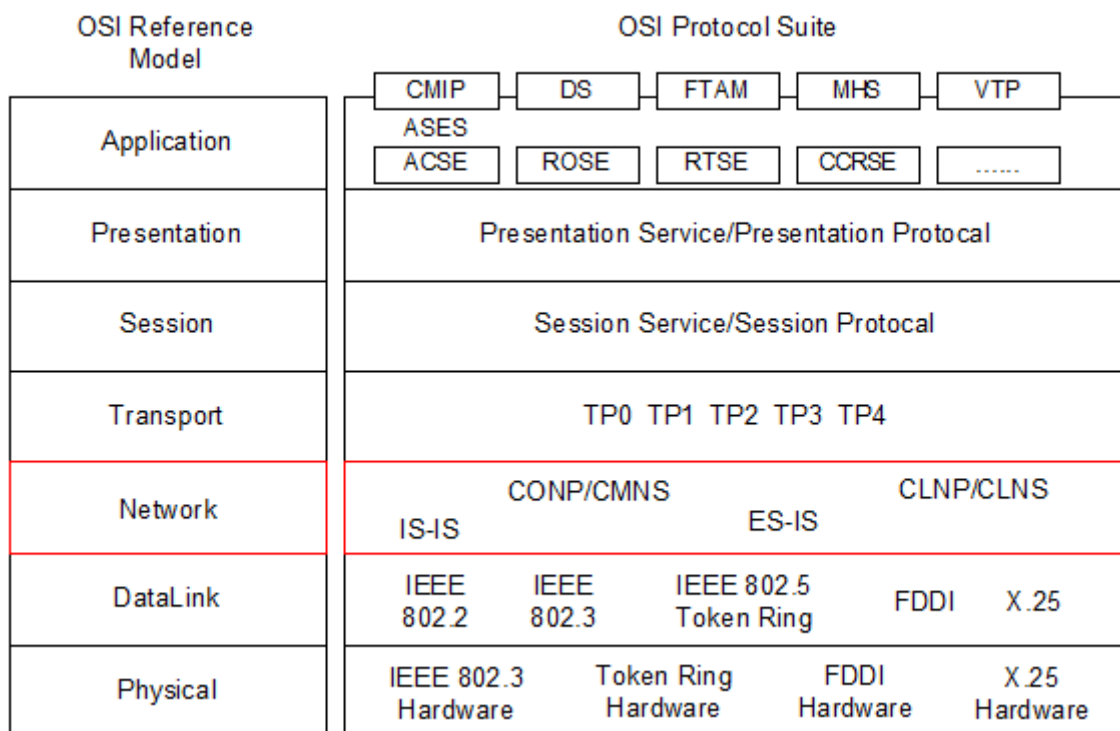
1. IS-IS 是一个链路状态的、内部网关协议。它被设计成适用在 OSI 无连接网络服务 ( CLNS : Connctionless Network Service ) 的环境中。为了提供对 IP 的路由支持, IETF 在 RFC1195 中对 IS-IS 进行了扩充和修改, 使它能够同时应用在 TCP/IP 和 OSI 环境中, 称为集成化 IS-IS ( Integrated IS-IS 或 Dual IS-IS )
2. 集成型 IS-IS 和 OSPF 都是 20 世纪 80 年代后期定义的, 大约是 1988 年。OSPF 实际上是从 IS-IS 的早期版本进化而来的, 不过它采用 IP 作为前提。
3. IS-IS 是路由选择协议, 传递 CLNP 的路由信息。在 ISO 协议族, 我们可以把 ES 终端系统理解为主机, 把 IS 中间系统理解为路由器。因此 IS-IS 的出生, 其实是为 CLNS 服务的, 往后发展成了集成性 IS-IS 就做了扩展, 也就可以支持 IP 路由了。
4. IS-IS 协议管理距离 115

### 1.2 OSI 协议栈

#### 1. OSI 协议簇及模型

OSI( Open System Interconnect )参考模型是一个国际化标准, 用于增强不同厂商设备之间的互操作性。它定义了一个 7 层的模型, 并且详细规定了各层的功能, 同时也确定了计算机网络的标准。制定 OSI 七层参考模型的是 ISO ( International Organization for Standardization , 国际标准化组织 )。对于数据通信和信息技

术的发展来说，OSI 参考模型起到了重要的作用。它提供了开放式的标准架构，使不同厂商生产的通信设备之间可以进行互联和互操作。ISO 七层模型的每一层都定义了单一的功能，可以将相关功能组合成功能层，从而简化和方便了协议的设计。



OSI 参考模型中的网络服务规范定义了网络设备之间使用无连接通信的功能，也就是 CLNS（Connectionless Network Service，无连接网络服务）。顾名思义，使用 CLNS，无需在发送数据之间建立端到端的路径。下图中展示的是 CLNS 中所包括的协议组件，这些协议组件都由 ISO 所定义。**OSI CLNS** 类似于 TCP/IP 中的**网络层服务**。

CLNP（Connectionless Network Protocol，无连接网络协议）、IS-IS、ES-IS（End System—Intermediate System，终端系统—中间系统）都是 ISO 定义的独立的 OSI 第三层（网络层）的协议，这些协议分别在不同的 ISO 标准中定义。

## 2. OSI CLNS 包含以下三个协议：

- CLNP 等价于 TCP/IP 模型中的 IP，他提供尽力而为的传输。
- ES-IS 终端系统到中间系统的协议，类似 TCP/IP 中的 ARP、ICMP 等协议
- IS-IS 中间系统到中间系统，路由选择协议。IS-IS 是路由选择协议，传递 CLNP 的路由信息。

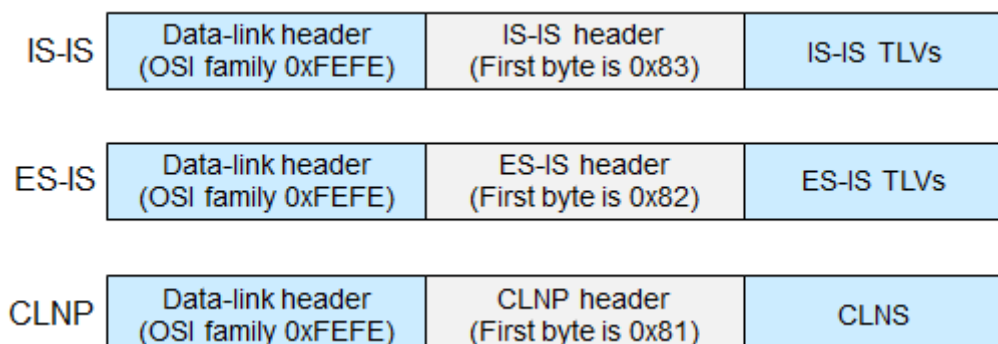
CLNP 这个名词可能很多人都比较陌生，它是一个 OSI 网络层协议。打个比方来说，它就相当于我们所熟悉的 IP 协议，而 IP 定义为用来为 TCP/IP 协议栈提供网络层服务。与 IP 一样，CLNP 也是一个无连接的协议，不提供可靠的数据连接，而且也独立于下层（数据链路层）协议。我们都知道，IP 是 TCP/IP 协议栈中唯一的网络层协议，高层的协议和数据全都封装在 IP 数据包中进行传输。这不同于 CLNS 网络环境，在

CLNS 中, CLNP、IS-IS、ES-IS 都是独立的网络层协议, 它们都直接被封装到数据链路层的帧中进行传输。

在 ISO 协议族, 我们可以把 ES 终端系统理解为主机, 把 IS 中间系统理解为路由器。因此 IS-IS 的出生, 其实是为 CLNS 服务的, 往后发展成了集成性 IS-IS 就做了扩展, 也就可以支持 IP 路由了。

### 3. 关于 CLNP

CLNP 类似 IP 协议, 只不过它是为 ISO 传输层提供服务的。IS-IS、ES-IS、CLNP 都是网络层协议, 都是直接封装在数据链路层帧内的。相比于 TCP/IP 中的 OSPF 报文是躲在 IP 头后的, 前者的协议报文的封装效率就要高一些了。



### 4. ES-IS

**OSI 已经为用于路由选择的两种网络协议制定了标准: ES-IS 和 IS-IS:**

在 OSI 术语中, 主机 (例如 PC) 被称为 ES (终端系统), 路由器被称为 IS (中间系统)。ES-IS 可以说是一种终端系统和路由器之间的“语言”或路由协议。它用来使同一网段或链路的终端系统和路由器之间可以彼此发现对方, 并可以让 ES 能够获悉其网络层地址。总结来说, **ES-IS 主要有以下几种功能:**

- 1) 使 ES 获悉其所在的区域, 即区域前缀
- 2) 在 ES 与 IS 之间建立邻接关系
- 3) 建立数据链路层地址到网络层地址 (CLNP 地址) 的映射

可以看出, ES-IS 在 CLNS 网络环境中的作用就好像 IP 网络中的 ICMP、ARP 与 DHCP 协议的协同工作。

在 ES-IS 工作过程中, 终端系统通过发送 ESH (ES Hello) 报文到特定的地址, 目的是向路由器通告自己的存在。路由器通过监听 ESH 报文, 以发现网络中存在的 ES, 以便后续将到达特定 ES 地址的数据包转发给 ES。在 ES-IS 中, 路由器通过发送 ISH (IS Hello) 报文到特定地址, 也向 ES 通告其自身的存在。ES 也监听 ISH, 如果收到多个 IS 发送的 ISH, ES 将随即进行选择, 并将所有数据都发送给这个 IS。

需要注意的是, 通常我们现网环境中的终端系统, 例如 PC, 都不使用 ES-IS, 因为这些 PC 都运行的是 TCP/IP 协议栈, 类似 ES-IS 的工作都由 TCP/IP 协议栈中的 ARP、ICMP、DHCP 协议来完成。

每一个 ES 都属于一个特定的区域, 当 ES 通过侦听中间系统的 HELLO (ISH) 分组发现最近的 IS 的时

候, OSI 路由就开始了。当一个 ES 想发一个分组给另一个 ES 的时候, 它就把分组发给它的直连网络上某个 IS ( 第 0 层 routing ), 然后这个 router 就会查找目的地址并沿着最佳路径转发分组。

## 5. IS-IS

IS-IS 是 CLNS 中一个重要的组成部分, 它是一个用来在 CLNS 网络环境中使路由器与路由器( IS 与 IS ) 之间动态的交换路由信息的协议, IS-IS 在 ISO 10589 中进行了定义。IS 与 IS, 即路由器与路由器之间的通信使用 IIH ( IS-IS Hello ) 报文。IS-IS 的设计主要是为了满足 CLNS 网络中的如下需求:

- 1) 在路由域内执行路由选择协议功能
- 2) 为网络提供最佳路由
- 3) 当网络出现故障后, 能够快速收敛
- 4) 提供无环路的网络
- 5) 提供网络的稳定性
- 6) 提供网络的可扩展性
- 7) 合理利用网络资源

为了满足如上需求, IS-IS 被设计成一种链路状态路由协议, 并且使用 SPF 最短路径优先算法以实现快速的收敛和无环路网络。

## 6. 集成 IS-IS

之前所提到的 IS-IS, 它仅支持 CLNS 网络环境, 而不支持 IP 网络环境中的路由信息交换。后来, IETF 在 RFC 1195 中对 IS-IS 进行了修改和扩展, 称之为集成 IS-IS( Integrated IS-IS )或双重 IS-IS( Dual IS-IS )。集成 IS-IS 的制定是为了使其能够同时应用在 TCP/IP 网络和 OSI 网络中, 使其能够为 IP 网络提供动态的路由信息交换。

集成 IS-IS 是一个能够同时处理多个网络层协议( 例如 IP 和 CLNP ) 的路由选择协议。相反, OSPF 只支持 IP 一种网络层协议, 即 OSPF 仅支持 IP 路由。而集成 IS-IS 可以支持纯 CLNP 网络或纯 IP 网络, 或者同时支持 CLNP 和 IP 两种网络环境, 并为其提供路由功能。集成 IS-IS 协议经过多年的发展, 已经成为一个可扩展的、功能强大的、易用的 IGP 路由选择协议, 并且在运营商网络中得到了更多的应用和部署, 主要用来实现域内的 IP 路由选择。

- 集成 IS-IS ( integrated IS-IS ) 使得 IS-IS 协议可以传播除 CLNP 之外的其他协议的路由信息。
- IS-IS 能在混合模式下同时路由 CLNP 和 IP。
- IS-IS 可以纯粹地用做 IP 路由选择, 也可以纯粹地用做 ISO 路由选择, 或同时用于两者。
- 即使只为 IP 提供路由选择功能, 也需要 CLNS 地址。

## 1.3 OSI 协议栈术语

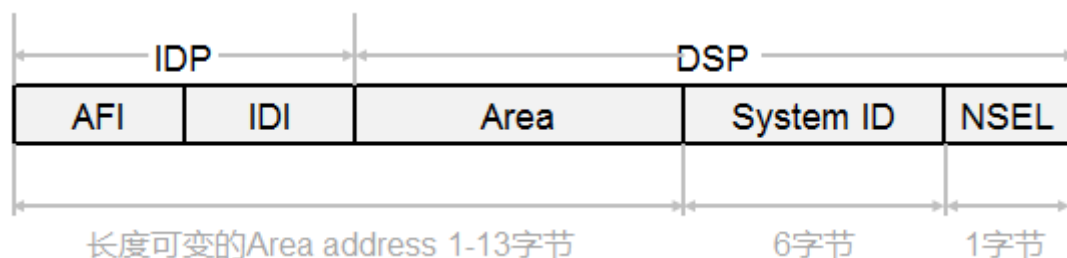
OSI 术语	OSI 术语解释	IP 中“类似”的概念
IS	Intermediate System 中间系统	Router 路由器
ES	End System 终端系统	Host 主机
PDU	Packet data unit 数据单元	Packets 包
NSAP	Network Service Access Point。是在网络层和传输层之间的边界上的概念性的点，它是 OSI 网络层为传输层提供服务的位置，每个传输层实体都会分配到一个唯一的 NSAP 地址。NSAP 的最后一个字节用来标识在同一个设备上的程序，类似于 TCP/IP 中的端口。	NSAP 可以理解为类似 IP+端口号的这么一个东西。
NET	Network Entity Title 是最后一个字节为 0 的 NSAP，它就是标识一个设备（NSAP 中的端口号为 0 嘛）。因此每个 router 都有唯一的 NET	-
SNPA	subnet point of attachments 是提供子网服务的点，它等价于对应的第三层地址（NET 或 NSAP）的第二层地址。通常是 LAN 上的 MAC 或者是 X.25，FR 或 ATM 中虚电路的 ID	MAC 地址等
SysID	System ID 系统 ID	OSPF 的 routerID
LSP	Link state Protocol Data Unit 链路状态数据单元	OSPF 的 LSA
LSPDB	LSP Database LSP 数据库	OSPF 的 LSADB
DIS	Designated Intermediate System 指定中间系统	OSPF 的 DR

## 2 ISO 编址

OSI 网络层编址是通过使用两类层次化地址：NSAP 和 NET 来实现的

### 2.1 NSAP

每一个传输层的实体都会分配到一个 NSAP 地址。**NSAP 地址是 CLNS 分组的网络层地址**。它用来标识设备。它由初始域部分（IDP）和域内自定义部分（DSP）组成，这两部分下面又做了详细的设定，在咱们理解这个 IDP 和 DSP 的时候，可以借 IP 里头的概念来套，IDP 有那么点像主网络号的意思，DSP 有那么点像子网号和主机 IP，当然，这个 IDP 的含义相对要复杂一点，见下图：



- AFI Authority and Format Identifier , 1 字节的授权和格式标识符。AFI 字段标识与 NSAP 相关的高层寻址域和 DSP 部分的语法。AFI 字段的取值范围为 0-99 的十进制数。高层地址域提供各种各样的子域，其值由 IDI 字段分配。每一个高层域定义自己的 IDI 字段格式。下面就是地址域 AFI 和 DSP 语法类型。

Address Domain	DSP Syntax		
	Decimal	Binary	Character
X.121	36	37	
ISO DCC	38	39	
F.69	40	41	
E.163	42	43	
E.164	44	45	
ISO 6523 ICD	46	47	
Local	48	49	50

- IDI Initial Domain Identifier , 可变长的初始域标识符，标识 AFI 下的子域  
AFI+IDI 用于标识 Domain
- Area 2 字节的 area 标识符，也称为 Domain 内自定义部分的高位部分（HODSP）用来把 Domain 细分为 area，大致类同 IP 中的子网
- System ID 6 字节的系统 ID，ES 或 IS 的标识符，类似于 OSPF 的 router ID，每台设备都有一个系统 ID，而在 IP 网络中每个接口都有一个 IP，这是区别之一。要注意：SystemID 必须在整个 Area 和主干（Level2）上保持唯一。
- N-SEL 1 字节的选择符，英文：NSAP-Selector 类似 TCP/IP 中的端口，识别设备上的进程（或服务），在 NET 中为 00

**IS-IS 使用简单的两层结构。把 IDP 和 HODSP 合在一起用作第 2 层路由的区域 ID，而剩下的系统 ID 就做第一层路由选择。**



对于 IP 应用程序而言,在 NSAP 地址中,1 字节定义 AFI,最少 2 字节定义实际区域信息,6 字节定义系统 ID 和 1 字节定义 NSEL。因此 NSAP 地址最少为 10 字节。在 CISCO IOS 中,NSAP 配置为点分 16 进制形式。AFI 值为 49 的只能本地使用,是 RFC1618 定义的预留私有地址空间。

## 2.2 NET

NSEL 值为 0 的 NSAP 地址用来标识设备,这就是这个设备的网络地址 NET。因此 NET 由区域 Id 和系统 ID 所决定。总的来说,NSAP 编址风格和 IP 编址风格之间的最大区别**就是 NSAP 仅使用一个地址标识一台 router**,而 IP 则是每个端口都分配一 Ip 地址。

### 分配 NET 的一些指导方针:

1. 一个中间系统(可以理解为路由器)至少有一个 NET(最多可有 254 个)且系统 ID 必须相同。Cisco 路由器默认支持最多三个 NET 地址,可使用如下命令修改:

```
Router(config-router)# max-area-address xx
```

2. 在一个路由选择区中的全部 IS 和 ES 必须有相同长度的系统 ID
3. 在一个区域中的所有 router 必须有相同的区域 ID
4. 所有的 2 层 router 必须有域内唯一的系统 ID
5. 所有的 1 层 router 必须有区域内唯一的系统 ID
6. 如果 ES 和 IS 有相同的区域 ID,那么同一个区域的所有 ES 都会和它同在一段共享媒介质上的 1 层 router 建立毗邻关系。
7. 如果一个 router 上分配了多个 net, 则这些 NET 的系统 ID 必须是相同的。



## 2.3 Cisco Router 的 ISO 编址

前面说了，NSAP 地址有 Area address 部分是变长的，那么当我们拿到一个地址，我怎么知道这个地址的哪一部分对应的是什么呢？Cisco IOS 软件从右边开始解释 NSAP 地址（这里使用 NSAP 的 OSI 格式）。首先从右边数起，第一个 B 是 NSEL，往左的 6B 是 SystemID，剩下的部分是 AreaID。

如 39.0f01.0002.0000.0c00.1111.00

NSEL 是 00（右边起 1B）

系统 ID 是 0000.0c00.1111

区域 ID 是 39.0f01.0002

对于 SystemID 的设定，习惯上把 router 的接口 MAC 或 IP 地址编码成系统 ID 或它的一部分。在集成 IS-IS 中普遍将回环 IP 用于这一目的，例如我有配置一个 loopback：192.168.1.24/32，然后我将这个 IP 补全，不够 3 位数的在前面加 0 形成 192.168.001.024 然后再形成一串儿：192168001024，再根据 4 位数一组形成 SystemID：1921.6800.1024。

## 2.4 SNPA

NSAP 或 Net 地址用来区分设备。在 ISO 术语中，数据链路层地址（LAN MAC，帧中继 DLCI 等）被称为子网连接点 SNPA（subnet point of attachments）。由于一个网络设备可能连接多个链路，因此需要多个 SNPA 地址，但只需要一个 OSI 网络地址。

ES-IS 协议的主要功能就是为网络设备提供 NSAP 地址到 SNPA 地址的映射。

SNPA 相当于 NSAP 或 NET 的二层地址。

**SNPA 地址可以通过下列几种方法获得：**

- Lan 接口的 MAC 地址
- X.25 或 ATM 的虚电路 ID
- FR 的 DLCI
- 对于 HDLC（高级数据链路控制）接口，SNPA 被设置为“HDLC”

**术语电路（circuit）相当于接口。**由于 NET 地址用于标识整个设备，所以 Circuit ID 用来表示不同的接口。路由器按照如下方式为接口指定 1 字节的 Circuit ID。

- 对于点到点接口，SNPA 是电路的唯一标识符。例如在 HDLC 点到点链路上，电路 ID 为 0x00
- 对于 LAN 接口，将 1 字节的电路 ID 附加到 6 字节的指定中间系统（DIS）的 SysID 的后面，形成一个 7 字



节的 LAN ID 如 1921.6811.1001.**03**。"the routers connected to the multiaccess network (Ethernet) all have the **same circuit ID**. The circuit ID is a one-octet number that the router uses to uniquely identify the IS-IS interface. If the interface is attached to a multiaccess network, the circuit ID is concatenated with the system ID of the DIS"

**术语链路**，是位于两个 IS（路由器）之间的路径，当两个相邻的 SNPA 可以通信是，链路就处于 UP 状态

## 2.5 NSAP 地址到主机名的映射

这 NSAP 地址呢，还是长了点，用起来呢，也不方便，这其中 SystemID 主要是用来标识几个属性，例如邻接信息、IS-IS 数据库中的 LSP 标示符等等，大多数情况下，我们比较适应那种有字面含义的符号，例如 R1、R2 神马的，而不是用一串数字。因此 CISCO IOS 为我们提供了一个非常好的方法：路由器 hostname 到 NSAP 地址的映射。本质上就是 hostname 到 SystemID 的映射。有静态映射和动态映射两个方法。

### 1. 静态主机名映射

RouterA 的配置如下：

```
router isis
  net 49.0001.1111.2222.3333.00
  exit
clns host RouterA 49.0001.1111.2222.3333.00
clns host RouterB 49.0001.4444.5555.6666.00
```

RouterB 的配置如下：

```
router isis
  net 49.0001.4444.5555.6666.00
  exit
clns host RouterA 49.0001.1111.2222.3333.00
clns host RouterB 49.0001.4444.5555.6666.00
```

在全局配置模式下使用 clns host 命令来建立主机名到 NSAP 地址的映射。这样一来在链路状态数据库中 LSP ID 中可以使用 hostname 来替代 SysID 部分。从而为我们的网络维护和排错带来极大的便利。

### 2. 动态主机名映射

静态映射，当网络规模大起来了，配置量就大得惊人，RFC2763 中定义了 IS-IS 的增强特性，其中就有动态主机名的映射定义。RFC2763 引入了新的 TLV，类型 137，在参与动态映射的路由器发送的 LSP 中包含着

这个 TLV，这样一来就提供了一个简单可靠的通告主机名信息的机制。如下图，这是一个 LSP 的报文：

```
ISO 10589 ISIS InTRA Domain Routeing Information Exchange Protocol
  Intra Domain Routing Protocol Discriminator: ISIS (0x83)
  PDU Header Length: 27
  Version (==1): 1
  System ID Length: 0
  PDU Type          : L2 LSP (R:000)
  Version2 (==1): 1
  Reserved (==0): 0
  Max.AREAs: (0==3): 0
  ISO 10589 ISIS Link State Protocol Data Unit
    PDU length: 109
    Remaining lifetime: 1200
    LSP-ID: 0000.0000.0004.00-00
    Sequence number: 0x00000037
    + Checksum: 0xd2b0 [correct]
    + Type block(0x03): Partition Repair:0, Attached bits:0, Overload
    + Area address(es) (4)
    + Protocols supported (1)
    - Hostname (2)
      Hostname: R4 TLV 137
    + IP Interface address(es) (4)
    + IS Reachability (23)
    + IP Internal reachability (36)
```

使用 show isis hostname 可以查看到主机名的映射

## 3 IS-IS 基础模块

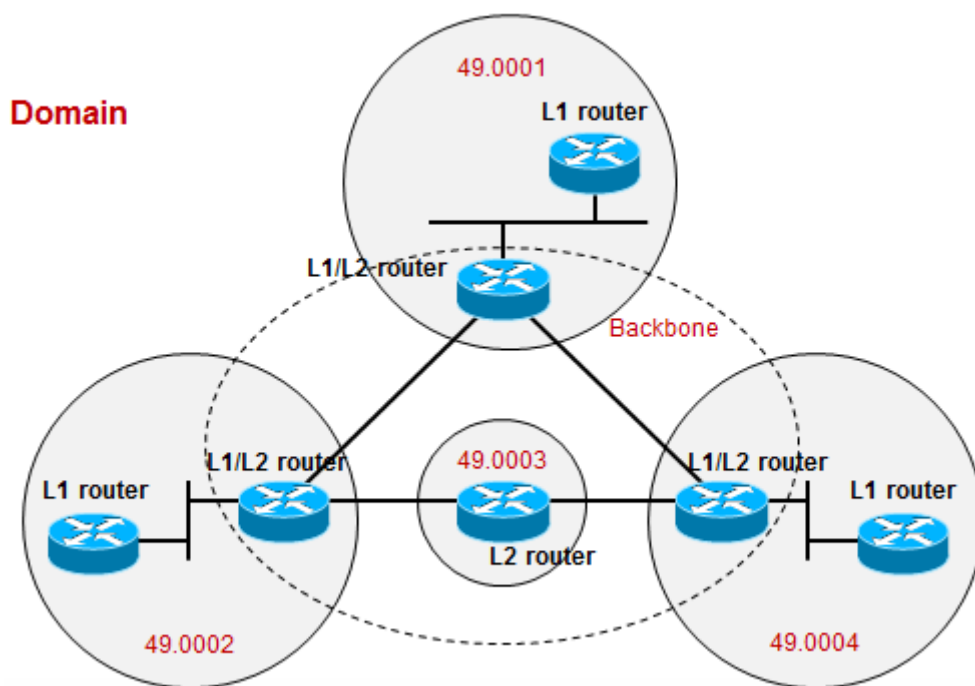
### 3.1 IS-IS 路由选择层次

#### 1. IS-IS area

- IS-IS 允许将整个 Domain 划分为多个区域，采用两级的分层结构
- 区域之间只能通过 L2 或 L1/L2 路由器进行互联
- 一台路由器最多可有 254 个 area ID。当然，一般一台路由器就属于一个区域，多区域的情况，可能用于区域的合并、分割或变更等（请见下文）
- 和 OSPF 不同的是，一台路由器必须整台属于一个区域，区域的边界不能在路由器上，也就是说，不能一个接口属于某个区域，另一个接口属于其他区域
- 区域内的所有路由器（除边界路由器外）只能和本区域内的路由器建立邻居关系

#### 2. Nodes level and area level

- **L1 router**，维护着第 1 层的 LSDB，默认情况下，L2 的 LSP 不会被泛洪过来，因此 L1 router 的 LSDB 比较轻量，只有本区域内的 LSP。
- **L2 router**，连接了不同的 Level1 area。储存着一个独立的 Level2 LSDB。
- **L1/L2 router**，同时保存着两套独立的 LSPD。同时运行着第 1 层和第 2 层的路由选择进程。在第 1 层的 LSDB 中维护着第 1 层的 LSP 信息 同时向其他 L1 router 通告它们是该区域的一个出口( 这些 L1 router 根据这个通告来形成一条指向给区域边界路由器的默认路由 );它们也支持第 2 层的功能来与主干中的其他 router 通信，并维护一个独立于第 1 层 LSDB 的第 2 层拓扑数据库。
- Level1 area 是 L1 routers 和 L1/L2 routers 的集合，Level2 area 是 L2 routers 和 L1/L2 routers 的集合。



### 3. L1 router

- 位于普通 area 内部
- L1 router 只与本区域内的 L1 router ( 或 L1/L2 router ) 形成邻接关系
- L1 router 只有本区域内 Level1 的链路状态数据库，包含本区域内所有的 L1 router 的路由信息
- 通过与自己最近的 L1/L2 router 的 ATT bit 生成指向此台设备的默认路由作为出口路由
- 在转发时，如果目的地址在本 area 内，就直接利用 L1 LSDB 生成的路由转发报文，如果目的地址不在本区域内，则利用本 area 最近的 L1/L2router 作为区域外网络的出口，由此可能造成次优路由

### 4. L2 router

- 位于骨干 area
- 可以和其他的 L2 ( 或 L1/L2 ) router 形成邻接关系
- L2 router 有 Level2 链路状态数据库，它包含所有区域间的路由信息

- 接收来自本区域内其他 L2 router 的报文，并按照目的地址将报文转发给其他 area 的 L2 router ( 或者转发到同一 area 的 L2 router )。接收来自其他区域的 L2 路由器的报文，并按照目的地址将报文转发

## 5. L1/L2 router

- L1/L2 router 通常位于 area 边界上
- 可以和本 area 内任何级别的路由器形成邻接关系；可以和其他 area 相邻的 L2 或 L1/L2 router 形成 L2 邻接关系
- 可能有两个级别的链路状态数据库：
  - L1 ( 区域内 ) 及 L2 ( 区域间 )
- 完成它所在的 area 和骨干之间的路由信息交换，既承担 L1 的职责也承担 L2 的职责
- 注意：一台 L1/L2 router 如果和其他 area 的路由器形成邻接关系，那么它将向本 area 的 L1 router 通告自己是个出口，具体方法是在生成本区域的 L1 LSP 时将报文中的 ATT bit 置 1

## 6. IS-IS 的层次性

### ISIS 由两个层次：

- Level2：Backbone，连续的 L2 router 的集合（含 L1/L2 router）；Backbone 是由所有的 L2（含 L1/L2 router）组成，Backbone 必须是连续的。  
注意：IS-IS 的 Backbone 不是某个特定的什么区域
- Level1：相对于单个区域的概念，由本区域中的 L1 router 构成，其路由信息发布到 Backbone 中。

**注意：**一个 IS-IS 路由域（routing domain）并不一定需要有两个层次，如果只部署一个区域的话，可能全都是 L1 或者全部是 L2，推荐用 L2，以得到比较好的扩展性。

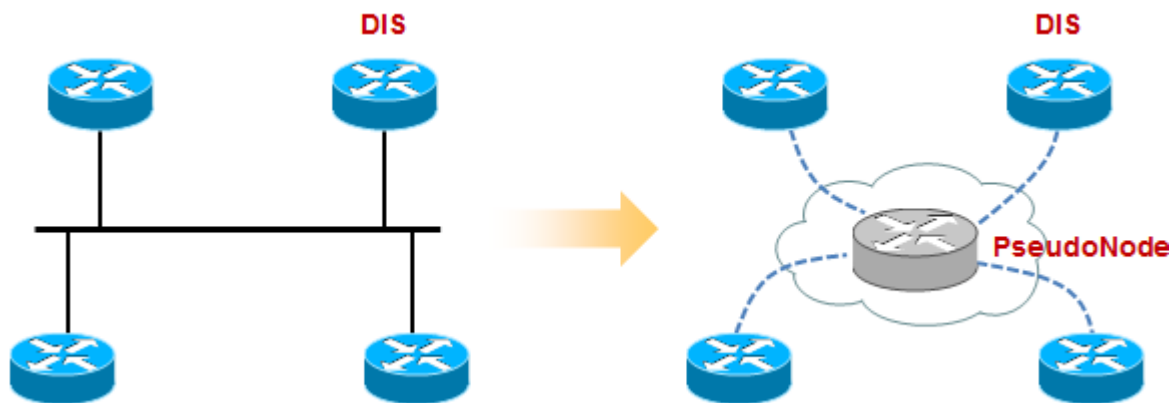
## 3.2 Designated IS and PseudoNode

### 1. Designated IS ( DIS )

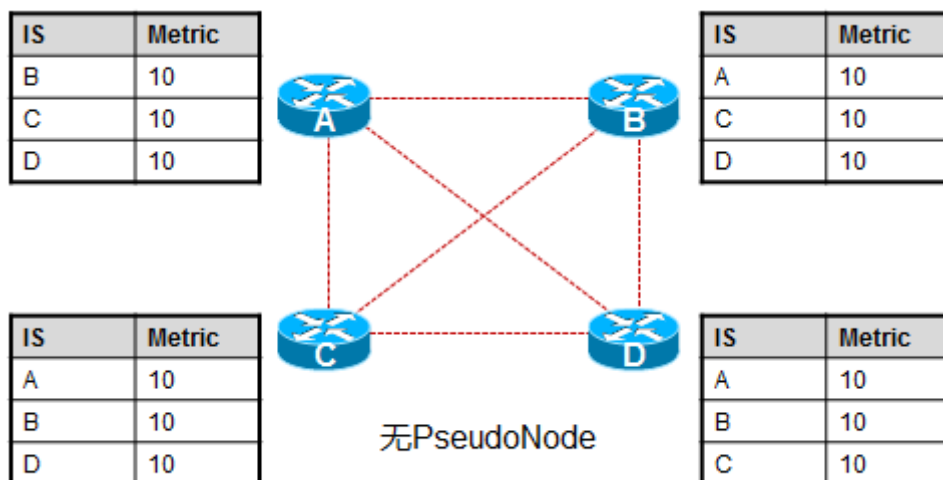
- 类似 OSPF 中 DR 的概念
- 在广播多路访问网络中，一台 router 会被选举为 DIS，点对点网络不需要 DIS
- Level1 有 level1 的 DIS，Level2 有 Level2 的 DIS，选举结果有可能不一样
- 与 OSPF 不同的是，DIS 是可抢占的，并且不存在备份 DIS。当一个 DIS 挂掉了，直接再选举。
- DIS 发送 HELLO 数据包的时间间隔是普通 router 的 1/3（默认是 3.3S），这样可以确保 DIS 出现故障的时候能够被更快速的被发现。

## 2. PseudoNode ( PSN )

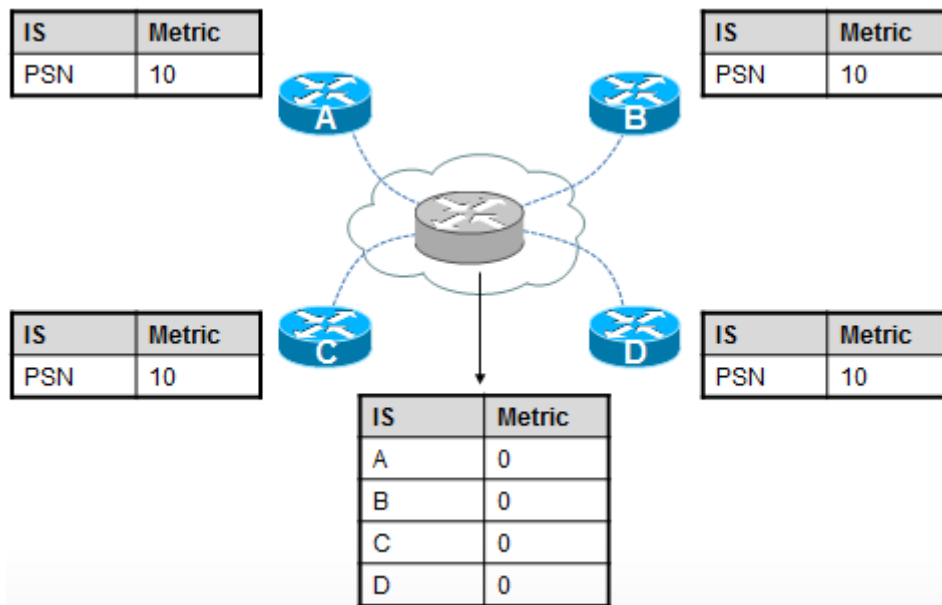
- 也叫伪节点，是广播多路访问网络中的一台虚拟路由器，由 DIS 创建
- DIS 在 “伪节点 LSP” ( 类似于 OSPF 中的 LSA ) 中通告 LAN 中的所有邻居
- LAN 中的所有路由器在它们的 LSP 中通告自己与伪节点的连接性 ( 自己和伪节点的邻接关系 )
- PSN 可以减少泛洪扩散和数据库同步的数量。



伪节点代表了一个广播多路访问网络 ( 包括连接到这个 LAN 上的所有 IS ), 实际上就是将整个广播多路访问网络 “当成 ”一台路由器 ( 如上图, 灰色的路由器 ), 在网络的其他地方看来, 这就是一个整体。那么这个整体是不是得产生一个 LSP ? 这个 LSP 就由 DIS 来产生。



如果没有 PSN 伪节点, 情况就像上图, LAN 上所有路由器的 LSPDB 如图所示, 图中红色虚线为邻接关系, 我们看到邻接关系多而且 LSPDB 庞大, 那么有了 PSN 伪节点呢? 看下图:



那么这样一来，LSPDB 的体积就大大减小了，注意，上面这张图只是一个抽象化、模拟化的概念，ABCD 四台路由器之间实际还是都有 ISIS 邻接关系的（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，这一点与 OSPF 是不同的），只不过有了 DIS 之后，大家产生的 LSP 体积减小了。也就是说，虽然 IS-IS 广播网上所有的路由器之间都形成邻接关系，但 LSDB 的同步仍然依靠 DIS 来保证。而 LSP 的稳定性和可靠性也得到了保证。SPF 算法的运算自然也就快了。

包括 DIS 在内的 LAN 中的所有 router 都跟 PSN 建立毗邻关系，依靠伪节点在 LAN 上形成的这种毗邻关系。一旦 DIS 在 LAN 中产生了伪节点，可靠性就得到了保证。它将周期性发送第 1 层和第 2 层的 Hello PDU 及 CSNP。Hello PDU 指明他是某一层的 DIS，CSNP 描述了所有 LSP 的汇总信息，包括 LSP ID、序列号、检验和剩余生存时间。它用组播来泛洪。CSNP 只是用来更正任何丢失的 PDU。

【补充】 非伪节点 LSP 代表一台 router，包括连接到这台 router 的所有 IS 和 LAN。

### 3. Designated IS ( DIS ) 的作用

- 在广播子网中创建并向子网中的所有路由通告“伪节点 LSP”（LSP 就是类似 LSA 的东东）
- 在广播子网中每隔 10S 周期性的发送 CSNP 来泛洪 LSP

DIS 发送 HELLO 数据包的时间间隔是普通 router 的 1/3（默认是 3.3S）

DIS 发送 CSNP 数据包的时间间隔默认是 10S

当我们选定一个 DIS 后，DIS 就代表了这个广播网络，包括连接到这个广播网络上的所有的 IS

### 4. 选举 DIS 的顺序

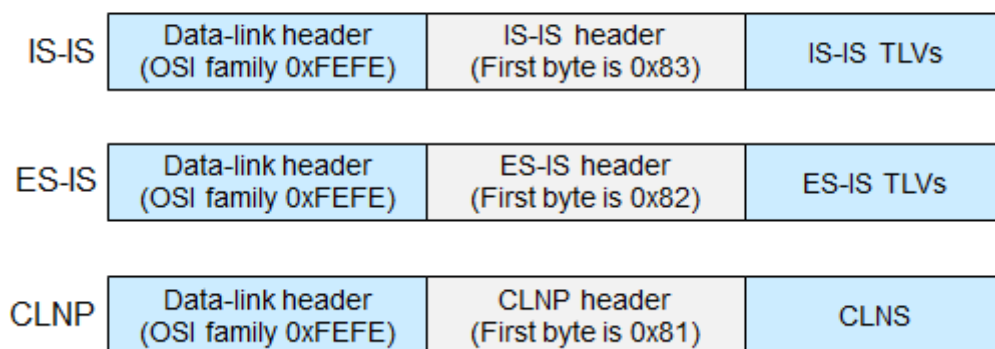
- 1) 接口优先级 (默认 64), 注意与 OSPF 不同的是, 优先级为 0 的 IS 也参与 DIS 的选举
  - 2) SNPA ( Subnetwork point of attachment ), 比大
    - 在 LAN 中, SNPA 指的是 MAC 地址
    - 在帧中继网络中, SNPA 是 DLCI
  - 3) 系统 ID, 比大
- DIS 是可抢占的, 而且没有备份的 DIS, 也就是说当 DIS 挂掉了, 立即重选。

## 3.3 OSI PDU

### 1. OSP PDU

- OSI 协议栈定义数据的单元为协议数据单元 PDU, 因此 OSI 视数据 Frame 为数据链路层的 PDU.
- 有三种类型的 PDU: IS-IS 的 ES-IS 的和 CLNP
- IS-IS 和 ES-IS 不将路由信息放在 IP 或 CLNP 分组中, 而是直接放在数据链路层 Frame 中 ( 这样封装效率更高 )。
- 真正的 CLNP 数据会被封装在 CLNP 头后, 然后进行二层的封装。

- Network PDU = datagram , packets
- Data-link PDU = frame



### 2. IS-IS 的 PDU 有 4 种:

- HELLO PDU( ESH 、 ISH 和 IIH)
  - 用于建立和维护毗邻关系
  - ESH 是 ES 发送到 IS 的
  - ISH 是由 IS 发送到 ES
  - IIH 则是 IS 之间传送的。



注意 ESH 和 ISH PDU 都是 ES-IS 的 PDU 而 不是 IS-IS 的 PDU

- **LSP**

用来发布链路状态信息，就是有点类似 OSPF 的 LSA 的东东

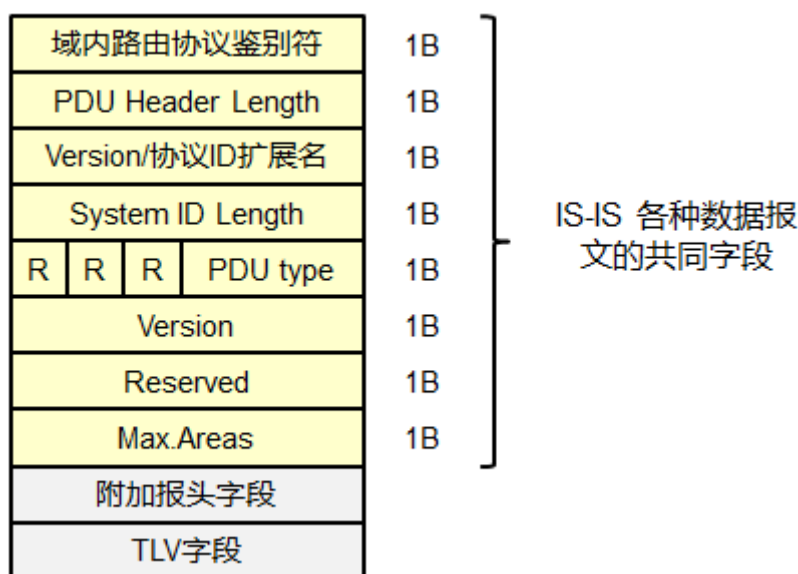
- **CSNP (完全序列号 PDU)**

用来发布一台 router 上的完整的链路状态数据库，CSNP 用来告知其他 router 他们自己的数据库中可能过时或者丢失的 LSP。

- **PSNP (部分序列号 PDU)**

用来**确认和请求**链路状态信息

### 3. IS-IS 报文格式



上图为 IS-IS 报文的格式，黄色背景色部分为 IS-IS 各种类型报文的共同字段。接下来每种数据包会有自己的附加报头字段，然后就是 TLV。

- 域内路由协议鉴别符：这是分配给 IS-IS 的网络层标示符，值为 0x83
- PDU header Length：数据报头字段的字节数（这个字节总数包括了附加报头字段的字节数在内）
- Version/协议 ID 扩展名：值为 1
- SystemID Length：标识系统 ID 的长度，值为 0 表示长度为 6B，值为 255 表示长度为 0，其他系统 ID 字段长度可能值为 1-8 字节。
- PDUType：PDU 类型。这是一个 5bit 的字段，用于标识 IS-IS 数据包的类型。值为 15 表示 L1 LAN IIH；值为 16 表示 L2 LAN IIH；值为 18 表示 L1 LSP；值为 20 表示 L2 LSP；值为 24 表示 L1 CSNP；值为 25 表示 L2 CSNP；值为 26 表示 L1 PSNP；值为 27 表示 L2 CSNP。
- Version：值为 1
- MAX.Areas：最多区域地址。表示我们可以为一个路由器配置多少个不同的区域前缀。值

为 0 表示最多支持 3 个区域地址数，默认情况下值为 0，取值范围为 1-254 的实际数字。

IS-IS 的数据包都有上面的共同 8B，接下去附加报头字段各有不同，然后就是 TLV 字段。在抓包的时候，TLV 中的三元：类型、长度、可变长的内容，中的类型及长度，抓包软件可能不呈现，但是可以从抓包结果的二进制编码中找到类型和长度的内容。

#### 4. IS-IS Hello 消息 ( IIH PDU )

- Hello 数据单元：定期发送，用来确定相邻的其他 IS 是否在运行 IS-IS，已建立邻接关系，交换 LSP，达到 LSDB 同步
- 在 IS-IS 里有三种 HELLO 包：一种是点对点接口的 ( Point-to-Point IS to IS hello PDU ) 一种是 LAN 上对 L1 router ( Level1 LAN IS to IS Hello PDU ) 的、一种是在 LAN 上对 L2 router 的 ( Level2 LAN IS to IS Hello PDU )
- 在 LAN 上 L1 和 L2 IIH PDU 发送到不同的组播 MAC 地址 L1 为 0180-C200-0014 ,L2 为 0180-C200-0015

更详细内容，见下文。

#### 5. IS-IS LSP 消息及 CSNP、PSNP 请见 IS-IS 链路状态数据库章节

### 3.4 IS-IS 毗邻关系及网络类型

#### 1. 对不同的网络类型，IS-IS 的邻接关系建立方式有所不同，IS-IS 支持如下两类网络：

- 点对点网络
- 广播网络

#### 2. IS-IS 邻接关系建立需要遵循的基本原则

- 只有同一层次的相邻路由器才有可能建立邻接关系
- 对于 Level1 router 来说要求 area 号一致
- 要进行同一网段检查 ( 邻接必须在同一个网段 )
- MTU 隐含检查

IS-IS 的 IIH 会打上许多 padding，hello 报文填充为出接口 MTU。这种机制在低带宽链路上会比较伤，可以在接口上 no isis hello padding 取消 padding，这条命令，如果要打，建议在链路两端都加上。实验亲测的结果是，如果一边 no 掉 hello padding，邻接仍然能建立。

### 3. IS-IS 的 IIH 概述

- Hello 数据单元：定期发送，用来确定相邻的其他 IS 是否在运行 IS-IS，已建立邻接关系，交换 LSP，达到 LSDB 同步
- 在 IS-IS 里有三种 HELLO 包：
  - 一种是点对点接口的 ( Point-to-Point IS to IS hello PDU )
  - 一种是 LAN 上对 L1 router ( Level1 LAN IS to IS Hello PDU ) 的、
  - 一种是在 LAN 上对 L2 router 的 ( Level2 LAN IS to IS Hello PDU )
- 在 LAN 上 L1 和 L2 IIH PDU 发送到不同的组播 MAC 地址： L1 为 0180-C200-0014，L2 为 0180-C200-0015
- 和所有的 IS-IS 数据包一样，HELLO 包也由报头和 TLV 组成。

### 4. HELLO 间隔、HELLO 乘数和 HELLO 保持间隔

- HELLO interval 是 hello 数据包的发送周期，有 25% 的抖动来减少网络中 IIH 同步传输的可能性。在 CISCO IOS 中，普通路由器 hello interval 默认 10S，DIS 默认 3.3S。
- HELLO multiplier ( HELLO 乘数 )，是丢失多少 HELLO 包后，邻接路由器宣告邻居失效，默认为 3
- HELLO holdtime 是 IS-IS 路由器收到 2 个连续 hello 数据包的最大允许时间间隔。Hello holdtime=hello interval × hello multiplier。因此 hello holdtime 默认为 30S。如果一台路由器在 hello holdtime 超时了还没收到任何 hello，那么宣告邻居失效。

### 5. 点对点链路上的 IS-IS 邻接

在点对点链路上默认 10S 发送一次 IIH。

域内路由协议鉴别符			
PDU Header Length			
Version/协议ID扩展名			
System ID Length			
R	R	R	PDU type
Version			
Reserved			
Max.Areas			
预留 ( 6bits )			Circuit type
SystemID ( sender )			
Holding timer			
PDU Length			
Local Circuit ID			
TLV			

路由器会检查收到的 IIH，确认里面的各项参数。

- Circuit type : 01 表示 L1 路由器，10 表示 L2 路由器，11 表示 L1/2 路由器
- SystemID : 产生该 hello 的 IS 的 systemID (有的书上写的是 SourceID，同义)
- Holding timer : 保持时间，默认 30S
- PDU length : 整个 PDU 的长度，包括头
- Local Circuit ID 本地电路 ID 是一个链路标示符。由发送 Hello PDU 的路由器分配给这条电路的标识，并且在路由器的接口上是唯一的。在点到点链路的另一端，Hello 报文中的本地电路 ID 可能或也可能不为同样的值。
- TLV 部分包含了系统特性的字段，如下图：

```

+ Restart Option (3)
+ Point-to-point Adjacency State (1)
+ Protocols Supported (1)
+ Area address(es) (4)
+ IP Interface address(es) (4)
  Padding (255)
  Padding (255)
  Padding (255)
  Padding (255)
  Padding (255)
  Padding (255)
  Padding (169)
  
```

#### 详细考虑一下邻接关系建立过程中的细节：

- 当路由器在一个点对点链路上收到 ISH 报文，路由器会检查本地的邻接数据库看看报文的源 systemID 是否存在，如果存在，则忽略该 ISH，如果不存在，则接收路由器创建新的邻接关系，设为“初始化”状态，系统类型为“未知”。随后发送 IIH 回应。如果这台路由器又从这个新的邻居收到一个 IIH 报文，路由器将邻接关系设为 UP，系统类型为 IS，但是这里有个问题，就是本地的 IIH 发出去后，我并不知道对方是否收到，那么就有可能报文在中途丢失而出现一边 UP 一边 DOWN 的情况。因此我们引入了三次握手机制，请看下文。
- 在默认的实现模式中，IIH 报文会被填充至出接口 MTU 大小(使用增加 padding 的方式，如上图)，路由器会对比本地接口 MTU 和收到的 IIH 大小，确保在形成邻接前能够处理来自邻接点的最大可能的数据包。
- 路由器始发 IIH 时，根据我们所做的配置来设置报头中的 circuit type：L1、L2 或 L1-2。
- 路由器采用 8bits 的 local circuit ID 给每个点对点链路分配一个本地唯一的链路标示符。
- IS-IS 要求整个域内的 SystemID 长度一致。如果收到的 IIH 中，“System ID Length”字段与本地的 sysID 长度不一致，则忽略该 IIH。
- 单台路由器配置的区域地址最大数目必须与邻接邻居一致，默认情况下 CISCO IOS 路由器最多支持 3 个区域地址。最新的 IOS 版本可以设置为 255。如果两台路由器支持的最大区域地址数目不一致，也就是 MAX.Areas 字段不匹配，则忽略 IIH 报文。

### 三次握手机制：

我们引入一个新的 TLV ( 类型为 240 ) ,称作 “point-to-point adjacency state” , 如果不支持该新特性的设备收到包含该 TLV 的 IIH , 将忽略这个 TLV , 这样可以保证向后兼容。

▣ Point-to-point Adjacency State (1)  
Adjacency State: Up

这个 state 有三种取值：down、init、up

## 6. 多路访问链路上的 IS-IS 邻接

LAN 上的 IS-IS 邻接与 P2P 链路还是有些不同的：

- 不是通过 ISH 触发，一旦接口激活了，路由器就在接口上泛洪 IIH 报文
- 产生伪节点，伪节点作用由选举的 DIS 扮演
- Lan 中的 router 与同一 LAN 中的所有具有相同区域 ID 和层次的其他 router 都建立毗邻关系。
- 在 LAN 上 L1 和 L2 IIH PDU 发送到不同的组播 MAC 地址 L1 为 0180-C200-0014 ,L2 为 0180-C200-0015

以下就是 LAN 中的 IIH 的报文格式：

域内路由协议鉴别符			
PDU Header Length			
Version/协议ID扩展名			
System ID Length			
R	R	R	PDU type
Version			
Reserved			
Max.Areas			
预留（6bits）			Circuit type
SystemID（sender）			
Holding timer			
PDU Length			
R	Priority		
SystemID（DIS）			
TLV			

- **Priority：** 优先级。接口的 DIS 优先级，用来在广播 LAN 中选举 DIS。优先级数值越高，路由器成为 DIS 的可能性越大。

- **SystemID(DIS) :** 局域网 ID。由 DIS 的 SysID 与 1 字节的伪节点 ID 组成, LAN ID 用来区分同一台 DIS 上的不同 LAN。

```

❏ ISIS HELLO
  Circuit type           : Level 1 and 2, reserved(0x00 == 0)
  System-ID {Sender of PDU} : 0000.0000.0002
  Holding timer: 30
  PDU length: 1497
  Priority               : 64, reserved(0x00 == 0)
  System-ID {Designated IS} : 0000.0000.0002.02
  + Protocols Supported (1)
  + Area address(es) (4)
  + IP Interface address(es) (4)
  + Restart signaling (3)
    Padding (255)
    Padding (255)
    Padding (255)
    Padding (255)
    Padding (255)
    Padding (163)
  
```

用到的 TLV 有 :

```

❏ Protocols Supported (1)
  NLPID(s): IP (0xcc)
❏ Area address(es) (4)
  Area address (3): 49.0001
❏ IP Interface address(es) (4)
  IPv4 interface address: 10.1.123.1 (10.1.123.1)
❏ Restart Signaling (3)
  ❏ Restart Signaling Flags: 0x00
    .... .0.. = Suppress Adjacency: False
    .... ..0. = Restart Acknowledgment: False
    .... ...0 = Restart Request: False
❏ IS Neighbor(s) (12)
  IS Neighbor: cc:02:2d:0c:00:00
  IS Neighbor: cc:01:2d:0c:00:00
  Padding (255)
  Padding (255)
  Padding (255)
  Padding (255)
  Padding (255)
  Padding (149)
  
```

其中, IS neighbor 放的是邻接的 MAC

Lan 上的路由器与其他所有路由器建立毗邻关系, 而 OSPF 中, LAN 上的路由器仅与 DR 建立毗邻关系。

**形成 LAN 邻接关系 :**

- 当 LAN 接口激活 IS-IS, 路由器立即发送 IIH, 其中 IIH 包含一个 LANID, 这个 LANID 在报文中显示为

DIS, 初始的时候是本地的 SystemID+唯一的本地 circuit ID 构成。路由器同时开始侦听 ESH、ISH、IIH 等报文从而视图发现邻接体。接下来开始 DIS 选举进程。

- 路由器收到 IIH 的时候, 会进行一系列的校验动作, 例如 SystemID 长度是否匹配啊, max area address 是否匹配啊, 验证密码是否正确啊, 区域 ID 什么的。
- 一旦路由器收到 IIH, 会检查 hello 包发送者的邻接情况, 如果邻接关系已经建立, 则重置该邻接的 holdtimer, 如果邻接关系没有建立, 接收路由器创建邻接关系并标识邻接类型 (L1、L2), 设置为初始化状态, 直到后续收到 hello 数据包确认邻接关系才算建立。
- 路由器发送的 hello 包里包含整个 LAN 上发送 hello 的所有邻居的 MAC 地址, 路由器使用一种简单的机制确认二步通信, 也就是在收到的 hello 包中, 如果看到自己的 MAC 那就算二步通信正常。如果没有看到自己的 MAC, 那么证明二步通信不正常, 邻接关系则维持在初始化状态。

## 3.5 IS-IS 链路类型

### 1. IS-IS 支持的链路类型

- **点对点链路**

如 PPP、HDLC

不选举 DIS, 链路建立起来后用 CSNP 来触发数据库同步。

- **广播型多路访问链路**

选举 DIS。在该网络中该 DIS 会在它所参与的每一路由选择层上 (无论是 L1 还是 L2) 以及它连接着的每一个 LAN 上都生成并泛洪扩散新的伪节点 LSP。LAN 上每台路由器都与其他所有路由器和 DIS 建立毗邻关系, 不选举备份 DIS 路由器。选举出的 DIS 并不能保证一直为 DIS。

- **非广播型多路访问链路 (实际上 IS-IS 并不理解 NBMA 介质)**

IS-IS 在 NBMA 网络中运行的时候, 有一些问题。

如果是帧中继 P2P 子接口, 那么不会有问题, 这是推荐的部署方法

如果是帧中继主接口或者是 P2MP 子接口, 则要求是全互联的 PVC, 具体请见本文档相关章节



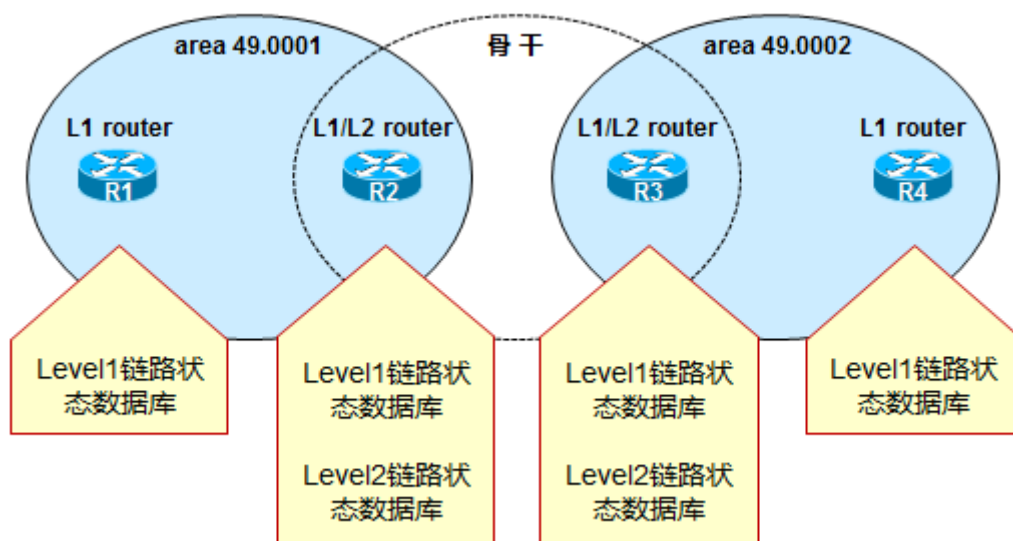
## 4 IS-IS 链路状态数据库

### 4.1 概述

区域内的路由器交换 LSP 的过程称为 flooding，从区域内的邻接路由器收到的 LSP 存储在本地路由器的 Level1 链路状态数据库里。IS-IS 域中的每一个区域都有一个唯一的 Level1 链路状态数据库。通过 SPF 算法再计算出来最短路径。

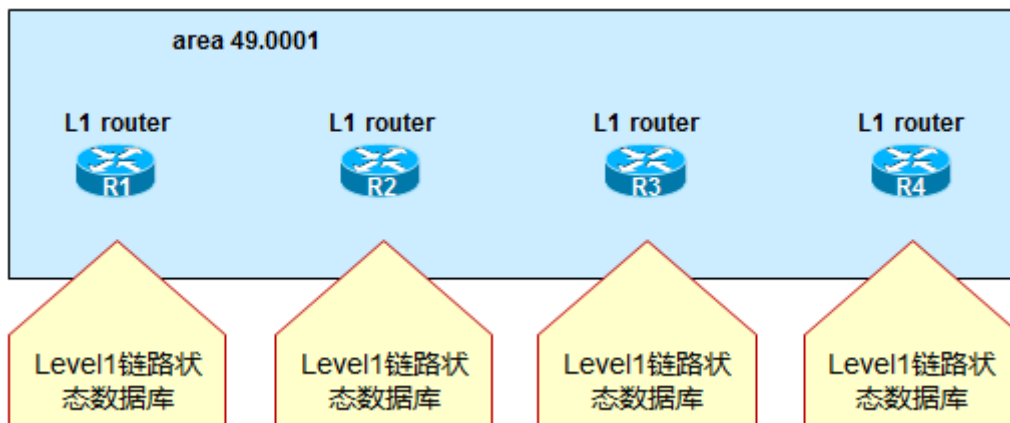
**IS-IS 支持两个层次的路由选择：Level1 及 Level2。** Level1 支持区域内的路由、Level2 支持区域间的路由。Level2 LSP 中封装了从骨干网络相连的路由器中获取的描述 Level2 网络拓扑结构信息。完整的 Level2 网络拓扑结构可以通过在 Level2 链路状态数据库中使用 SPF 算法获得。

在一个典型的网络设计中，多个独立的 Level1 区域通过骨干网相连，骨干网由具有 Level2 路由选择功能的路由器构成。Level1-2 路由器拥有分别支持 Level1 及 Level2 路由选择的两个独立的链路状态数据库。

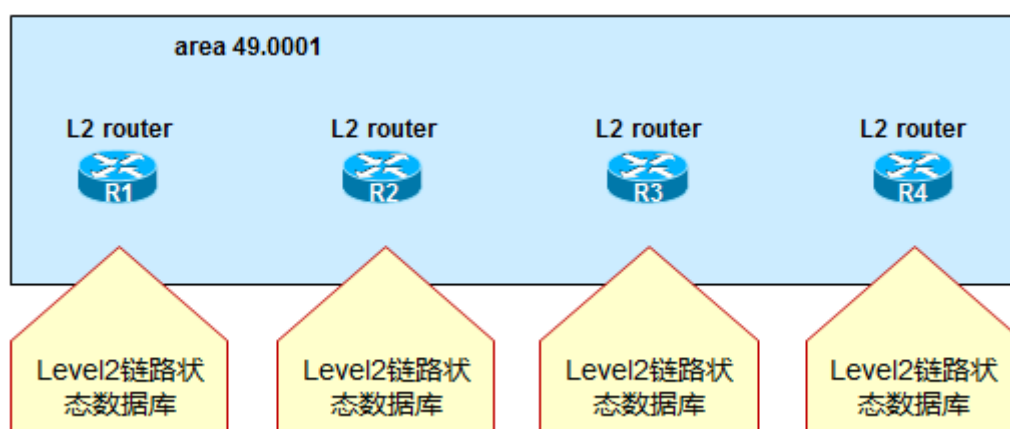


从上面这个图，我们可以看出，R1、R4 作为 L1 router，只维护 Level1 的链路状态数据库，R2、R3 是 L1/L2 router，因此维护两个互相独立的 L1 链路状态数据库及 L2 链路状态数据库。那么，如果骨干中，若有一台路由器并不与 L1 router 相连，可以配置为 L2 router。

当然，如果是单区域运行，你可以采用全 L1 router，也可以采用全 L2 router 的解决方案。



上图是单区域的 Level1 路由选择域



上图是单区域的 Level2 路由选择域

## 4.2 LSP ( LinkState Protocol Data Unit ) 格式

### 1. LSP 概述

- LSP 链路状态数据单元
- 链路状态报文用来在区域中传播链路和节点信息（描述本地网络的拓扑结构）
- LSP 的泛洪和交互最终构成 IS-IS 链路状态数据库
- LSP 分为两种：Level1 LSP PDU 和 Level2 LSP PDU
- Level2 LSP 包含在 IS-IS 里所有可能到达的前缀信息。Level1 LSP 只用于本地区域

### 2. LSP 主要包含如下信息

- 区域信息
- 邻接路由器
- IP 子网
- 度量值
- 认证信息

### 3. LSP 报文结构

域内路由协议鉴别符			
PDU Header Length			
Version/协议ID扩展名			
System ID Length			
R	R	R	PDU type
Version			
Reserved			
Max.Areas			
PDU Length			
Remaining Lifetime			
LSP-ID			
Sequance num			
checksum			
Partition repair	Attach ed	overload	IS类型
TLV			

2B  
2B  
ID长度+2  
4B  
2B  
1B

**Remaing lifetime** LSP 到期前的生存时间  
**LSP ID** LSP 的身份标识  
**Sequence num** LSP 的序列号  
**Type block** 这是 1B，里头包含不少重要的位：

Partition repair	Attached	overload	IS类型
1bit	4bits	1bit	2bits

```

Type block(0x03): Partition Repair:0, Attached bits:0, Overload bit:0, IS type:3
0... .... = Partition Repair: Unsupported
.000 0... = Attachment: 0
0... = Error metric: Unset
.0.. = Expense metric: Unset
..0. = Delay metric: Unset
...0 = Default metric: Unset
.... .0.. = Overload bit: Unset
.... ..11 = Type of Intermediate System: Level 2 (3)
    
```

**Partition repair:** 第 8 位如果 LSP 始发路由器支持区域划分修复则设置该字段

**Attached bits :** 第 4-7 位, 用于设置该 LSP 与带有如下可用度量的另一个区域的关  
联关系: 第 4 位-默认; 第 5 位-延迟; 第 6 位-开销; 第 7 位-误差

**Overload :** 第 3 位 如果设置则表示 LSP 始发路由器的链路状态数据库已超载,  
内存空间和 CPU 资源已受到限制

**IS type :** 第 1、2 位, 用于指示路由器的类型是 L1router 还是 L2router

不同的 TLV 字段值可以包含在 LSP 中用来通告各种不同的路由选择信息。

## 4.3 LSP header 详解

### 1. Remaning timer

**LSP remaining time 字段中两个重要的阈值是 : LSP maxage 及 LSP refresh interval**

**LSP maxage**, 定义了 LSP 的生存时间, 默认为 20 分钟, 也就是 1200S, 可使用 `lsp-max-lifetime` 命令修改。当一条 LSP 从一台路由器始发, 其 remaning timer 就被设置为 LSP maxage, 然后在区域内泛洪。同时在路由器本地, 也会为这个 LSP 的 remaining time 进行倒计时, 当倒计时达到 LSP refresh interval 时, 也就是减少的时间等于 LSP refresh interval 时, LSP 的始发路由器会重新产生 LSP 并泛洪出去。否则该 LSP 将继续生存下去, 知道剩余生存时间字段值减少到 0, 然后被清除出网络。

**LSP refresh interval**, 是一个 LSP 周期性的泛洪时间, 默认 15 分钟, 也就是 900S, 可通过 `lsp-refresh-interval` 来修改。

如果某台路由器在 LSP refresh interval 来临后没有刷新其产生的 LSP, 并且 LSP remaining timer 字段已经减少到 0, 那么所有拥有该 LSP 拷贝的路由器, 将在 60S 宽限期后将该 LSP 拷贝从自己的链路状态数据库中清除。这个 60S 的时间是 ZeroAgeLifetime, 零寿命生存时间, 不可配置。

### 2. LSP ID (链路状态数据包标示符)

## 用来唯一地标识一个 LSP 及鉴别源路由器

如 R1.01-00 其中 R1 为 SystemID (这是基于主机名的 SystemID), 01 为 PseudonodeID 伪节点 ID, 如果是伪节点发的, 则为非 00, 00 为分片号。

我们的每个 IS 产生的 LSP 一般就一个, LSP 的第一段编号为 0。但是如果我这 LSP 特别大, 超过了接口 MTU, 那么就有可能先进行分片然后再传, 再用这个分片号来表示 (0 为第一段, 1 位第二段, 依此类推), 单个 LSP 最大长度是 1492 字节, 最多可以分 255 片。IS-IS 通过在产生 LSP 的时候就进行分片, 而不是先产生大 LSP, 然后在网络层 (例如 IP) 分片, 从而达到降低 CPU 资源浪费的负面效应。

ISO 10589 中要求把 hello 数据填充到输出接口的 MTU 大小, 这意味着邻接关系的两端通常使用同一个 MTU 来构成邻接关系。

### 3. LSP 序列号

4B, 是一个无符号整数。从 0 开始, 每次增加 1。当一个路由器最先连接到网络中, 它产生的第一个 LSP 的序列号就为 1。当网络发生变化时而产生的 LSP 序列号也加 1。LSP 序列号在数据同步过程中起到关键作用, 主要体现在用于区分 LSP 的新旧。

### 4. 链路状态数据包校验和

Checksum 的计算从 LSP 中剩余生存时间字段之后的字段开始知道数据包末尾。LSP 拷贝在网络中路由器之间传播时其 checksum 字段不被修改。

一般来说, 任何一个检测到被破坏的 LSP 的路由器都会通过发送一个剩余生存时间重置为 0 的被破坏的 LSP 的拷贝来启动一个清除进程。这种方式有效的保证了网络中的其他路由器清除这个被破坏的 LSP。被破坏的 LSP 不参与路由计算, 当然也不会在网络中被泛洪。

如果一条 LSP 在传输过程中不断被破坏, 它会被其他路由器不断清除, 同时源路由器不断重发。这就产生了 LSP 破坏风暴。CISCO IOS 允许路由器忽略被破坏的 LSP 并只在本地将错误写进日志文件。通过 lsp-ignore-errors 命令开启。

### 5. Type Block

Partition repair	Attached	overload	IS类型
1bit	4bits	1bit	2bits

#### ● Partition repair

ISO10589 中定义了如何通过建立一条穿过 Level2 骨干的 Level1 路由来修复一个被隔离的 Level1 区域。这主要通过在每个区域中选举出一台具有 level2 功能的路由器作为区域指定的 level2 IS 并在区域间建立一条被称作虚邻接或虚链路的特殊邻接关系来实现。

虚链路实现了穿过骨干的 level1 修复路径。区域指定路由器通过设置其 level1 LSP 的 partition repair

位来通告虚链路，从而把虚链路存在的信号通知给 level1 router 以便实现区域间的数据转发。当然，这个特性极其少用，在 CISCO IOS 上的支持，还得查查相关的文档。

- **Attached**

区域关联。L1/L2 路由器在其生成的 L1 LSP 中设置该字段以通知同一区域中的 L1 路由器自己与其他区域相连。通常来说就是 L2 骨干区域相连。当 L1 区域中的路由器收到 L1/2 路由器发送的 ATT 位被置位的 L1 LSP 后，它将创建一条指向 L1/2 路由器的默认路由，以便数据可以被路由到其他区域。虽然 ATT 位同时在 L1 LSP 和 L2 LSP 中进行了定义，但是它只会在 L1 LSP 中被置位，并且只有 L1/2 路由器会设置这个字段。

ISO 10589 中定义的 IS-IS 区域是 stub 区域，Attached 定义了四种 metric 类型，CISCO IOS 只支持其中的 default metric。

这里就有一个隐含的问题，因为 Level1 router，并不知道 level2 的（本区域外的）路由信息，取而代之的是寻找最近的 L1/L2 router 前往目的地，那么就可能会出现次优路径的问题，为了解决这个问题，IS-IS 引入了一个叫做路由泄露的机制。

## 6. Overload

表示一台路由器资源的可用状态。如果该 bit 置位了，则表明该路由器发生了超载。超载是指路由器的性能因为 CPU 与内存资源不足而受到抑制。被设置了 overload 字段的那些 LSP 不会在网络中扩散并且在计算通过 overload 路由器的路由时也不会采用这些 LSP。换句话说，overload 路由器被传输路绕开了，那些 overload 路由器作为最后一跳的路由才参与计算。

使用 set-overload-bit 来手动设置 overload 位。

## 7. IS-type

表明了该 LSP 是来自 L1 router 还是 L2 router，这个字段 2bits。

00-未使用；01-level1；10-未使用；11-level2

## 4.4 LSP 的各类 TLV

根据不同的 IS-IS PDU 类型和特定的网络环境，紧跟在各种类型 IS-IS PDU 之后的是 TLV( Type/Length/Value ) 字段，PDU 报头与 TLV 字段构成了一个完整的 IS-IS PDU。在 ISO10589 和 RFC1195 这二种当前 IS-IS 标准中，使用代码 ( code ) 这个词，而不是类型 ( type )，但由于 TLV 用于其他协议标准中，故 TLV 比 CLV 在网络文献中使用的多，在这里我们也使用 TVL 代替 CLV。在 IS-IS PDU 所使用的各种 TLV 中，既有 ISO 10589 中定义的，也有 RFC 1195 中定义的。ISO 中定义的 TLV 用于 CLNP 网络环境，但是其中的大多数也用于 IP 网络环境。RFC 中定义的 TLV 只用于 IP 环境。对于一个 IS-IS PDU，后面既可以携带支持 CLNP 协议的 TLV，又可以携带支持

IP 协议的 TLV。如果一个路由器不能识别一个 TLV，那么将忽略它。

## 1. Level1 路由的 LSP TLV 字段

TLV	类 型 值	定义来源	备注
Area address	1	ISO10589	列出了路由器上配置的区域地址组。它只出现在非伪节点的 LSP 中，并且如果 LSP 被划分成不同的片段，那么该 TLV 字段出现在第一段。
中间系统邻接路由器 IS reachability	2	ISO10589	<p>IS Reachability (23) Reserved value 0x00, must == 0</p> <div> <div>IS Neighbor: 0000.0000.0005.00 Default Metric: 10, Internal Delay Metric: Not supported Expense Metric: Not supported Error Metric: Not supported</div> <div>IS Neighbor: 0000.0000.0003.00 Default Metric: 10, Internal Delay Metric: Not supported Expense Metric: Not supported Error Metric: Not supported</div> </div> <p>IS邻居 0000.0000.0005 IS邻居 0000.0000.0003</p> <p>列举邻接的 L1-router，注意啊，我们这小节说的是 L1 router LSP 的 TLV 字段啊。这里头除了包含邻接 IS 的系统 ID，还有一堆 metric</p>
终端系统邻接路由器	3	ISO10589	只在 L1 router 的 LSP 中出现。列举了邻接的 L1-router 和终端主机。例如通过 ES-IS 协议搜寻到的 ISO CLNP 工作站。
认证信息	10	ISO10589	用于认证
IP 内部可达性信息	128	RFC1195	列举了直连的 IP 地址前缀列表，它只在非伪节点 LSP 中使用。每个 IP 前缀分配了一堆的 metric。
支持协议 Protocols support	129	RFC1195	列举了集成型 IS-IS 协议支持的第三层协议，目前只有 CLNP ( NLPID 0x81 ) 和 IP ( 0xCC )
IP 接口地址	132	RFC1195	包含源路由器上配置的一个或多个 IP 地址，在 CISCO IOS 中为最高的 loopback 口 IP

## 2. Level2 路由的 LSP TLV 字段

TLV	类 型 值	定义来源	备注
Area address	1	ISO10589	列出了路由器上配置的区域地址组。



中间系统邻接路由器 IS reachability	2	ISO10589	<p>IS Reachability (23) Reserved value 0x00, must == 0</p> <div> <div>IS Neighbor: 0000.0000.0005.00 Default Metric: 10, Internal Delay Metric: Not supported Expense Metric: Not supported Error Metric: Not supported</div> <div>IS邻居 0000.0000.0005</div> </div> <div> <div>IS Neighbor: 0000.0000.0003.00 Default Metric: 10, Internal Delay Metric: Not supported Expense Metric: Not supported Error Metric: Not supported</div> <div>IS邻居 0000.0000.0003</div> </div>
分离指定层 2 中间系统	4	ISO10589	这种 TLV 通过在 L2 router 之间建立一条穿过骨干网的虚链路来支持区域修复
前缀邻接路由器	5	ISO10589	搜集可达的 NSAP 前缀信息, 它只用于区域间 ISO CLNP 路由选择 (level2 路由)
认证信息	10	ISO10589	用于认证
IP 内部可达性信息	128	RFC1195	列举了直连的 IP 地址前缀列表
支持协议 Protocols support	129	RFC1195	列举了集成型 IS-IS 协议支持的第三层协议, 目前只有 CLNP ( NLPID 0x81 ) 和 IP ( 0xCC )
IP 外部可达性信息	130	RFC1195	收集通过其他路由协议得到的 IP 路由信息
域间路由选择协议信息	131	RFC1195	暂不支持
IP 接口地址	132	RFC1195	包含源路由器上配置的一个或多个 IP 地址, 在 CISCO IOS 中为最高的 loopback 口 IP

## 4.5 IS-IS metric

1. 你看我们在 LSP 中的许多地方都出现了 IS-IS metric , 例如 :

```

❑ IP Interface address(es) (4)
    IPv4 interface address: 4.4.4.4 (4.4.4.4)
❑ IS Reachability (23)
    Reserved value 0x00, must == 0
❑ IS Neighbor: 0000.0000.0005.00
    Default Metric: 10, Internal
    Delay Metric: Not supported
    Expense Metric: Not supported
    Error Metric: Not supported
❑ IS Neighbor: 0000.0000.0003.00
❑ IP Internal reachability (36)
❑ IPv4 prefix: 4.4.4.0/24
❑ IPv4 prefix: 10.1.34.0/24
❑ IPv4 prefix: 10.1.45.0/24
    Default Metric: 10, Internal, Distribution: up
    Delay Metric: Not supported
    Expense Metric: Not supported
    Error Metric: Not supported
    
```

## 2. 一共有四种 metric 类型：

- Default metric 所有 IS-IS 路由器都支持，经常被解释成跟带宽成反比的度量方式，越小越优先
- Delay metric 可选，链路的传输延迟，不支持
- Expense metric 可选，链路的传输开销，不支持
- Error metric 可选，链路的出错概率，不支持

## 3. 每一种 metric 类型，都是 8bits，形式如下：

1bit	1bit	6bits
0	I/E	Default metric
1	I/E	Delay metric
1	I/E	Expense metric
1	I/E	Error metric

如果最高位(最左边的位)为 0，则表示 metric 类型为 default metric，如果为 1，则表示不支持该 metric 类型。第 7bit，是内部/外部标志位，设置为 0 表示是内部度量，设置为 1 表示是外部度量。

在 CISCO IOS 路由器中，IOS 无法自动分配基于带宽的链路(接口)度量，无论链路的带宽多少，所有的接口默认 IS-IS 度量都是 10，当然你可以修改。一完整路径的最大度量值是 1023。当然 IETF 对 default metric 做了扩展。

## 4. IS-IS 度量值的扩展

前面说了 IS-IS 四种 metric 类型，每种 metric 其实真正能用的 bits 就是低 6 位，等于说，接口 metric 最大也就是 63，我们提出几个新的 TLV，来弥补这个缺陷。

### 扩展的中间系统可达性 TLV ( 类型 22 )

是类型 2TLV 的扩展形式

### 扩展的 IP 可达性 TLV ( 类型 135 )

是类型 128TLV 的扩展形式

### 流量工程 RouterID TLV ( 类型 134 )

主要用于 TE

- **扩展的中间系统可达性 TLV ( 类型 22 )**

主要用于代替类型 2TLV 以提供更大的度量值。另外，它还支持基于 IS-IS 的 MPLS TE

Type	Length	Value
22	Value 的长度	3B 的 default metric，够你臭屁的吧？ 1B 的 sub-TLV 长度 6B 的 systemID+1B 的 PSN ID 0-244B 的 sub-TLV 主要用于 MPLS TE

- **扩展的 IP 可达性 TLV ( 类型 135 )**

被用来代替 TLV128。

全局配置命令 metric-style wide 可以让 cisco 路由器运行适当的 IOS 版本来发送携带 宽度量值 的 LSP 同时也能接受并理解解释携带 TLV22 和 TLV135 的 LSP

全局配置命令 metric-style transition 可以让原来的窄度量平滑扩展到新的 宽度量 ,这条命令允许一台路由器同时接受和发送窄、及宽度量。

MPLS TE 配置需要使用宽度量值

全局配置命令 metric-narrow 可以恢复到最初的默认配置

## 4.6 SNP 序列号数据包

- Complete sequence number packets ( CSNP ) 完全序列号数据包
- Partial sequence number packets ( PSNP ) 部分序列号数据包

两种数据包具有相同的格式并且各自携带了 LSP 摘要信息的集合。区别在于 CSNP 携带了该路由器链路状态数据库中所有已知 LSP 的摘要信息，而 PSNP 携带的是是其中一个子集。为了区分 Level1、Level2 链路状态数据库，路由器产生 L1、L2 各自独立的 SNP 报文。

## 1. CSNP

- 对于第一层的 LSP 数据库,使用的是第 1 层的 CSNP 和 PSNP,对于第 2 层的使用的则是第 2 层的 CSNP 和 PSNP。
- 描述了 LSP 数据库里所有的 LSP 信息。
- 用于两种情况:DIS 每隔 10s 周期性组播;链路刚刚建立起来时的点对点链路上。
- 如果 LSP 数据库过大,则要发送多个 csnp
- 在点对点链路上,一旦毗邻关系建立起来了,两个 IS 都会发送 CSNP 分组。当发现有丢失的 LSP 时会用 PSNP 来请求。
- 在 LAN 中存在着 DIS,DIS 生成并更新伪节点信息,并在 LAN 中扩散 LSP。DIS 每隔 10s 发送 CSNP,其中列出其链路状态数据库中保存的 LSP。这些 CSNP 通过组播发送到 LAN 上所有 IS-IS router。

### ISO 10589 ISIS InTRA Domain Routeing Information Exchange Protocol

Intra Domain Routing Protocol Discriminator: ISIS (0x83)

PDU Header Length : 33

Version (==1) : 1

System ID Length : 0

PDU Type : L2 CSNP (R:000)

Version2 (==1) : 1

Reserved (==0) : 0

Max.AREAs: (0==3) : 0

#### ISO 10589 ISIS Complete Sequence Numbers Protocol Data Unit

PDU length: 67

Source-ID: 0000.0000.0003.00

Start LSP-ID: 0000.0000.0000.00-00

End LSP-ID: ffff.ffff.ffff.ff-ff

#### LSP entries (32)

⊕ LSP-ID: 0000.0000.0002.00-00, Sequence: 0x0000000a, Lifetime: 1198s, Checksum: 0x00000000

⊕ LSP-ID: 0000.0000.0003.00-00, Sequence: 0x00000008, Lifetime: 1196s, Checksum: 0x00000000

- Start LSP ID: 起始 LSP ID。表示 TLV 字段中描述的 LSP 范围的第一个 LSP ID。
- End LSP ID: 结束 LSP ID。表示 TLV 字段中描述的 LSP 范围的最后一个 LSP ID。

## 2. PSNP

主要是两个功能:

- 路由器使用 PSNP 在点对点链路上确认多个 LSP (中的下一个接收)
- 路由器使用 PSNP 来请求 LSP 的最新版本或者被丢失的 LSP。这点上同时适用于点对点 and 广播链路

在点到点链路上,邻居用 **PSNP 来确认新的 LSP**。当源 router 收到了邻居发送过来的 PSNP 确认时,它会停止向这个邻居发送新的 LSP。在 LAN 中,没有明确的 PSNP 确认。丢失了的 LSP 可以被检测到。如果任何 LSP 丢失或过期了,router 都会以 **PSNP 的形式发送请求**来要求重传这些 LSP。

```
ISO 10589 ISIS InTRA Domain Routeing Information Exchange Protocol
  Intra Domain Routing Protocol Discriminator: ISIS (0x83)
  PDU Header Length : 17
  Version (==1) : 1
  System ID Length : 0
  PDU Type : L2 PSNP (R:000)
  Version2 (==1) : 1
  Reserved (==0) : 0
  Max.AREAS: (0==3) : 0
  ISO 10589 ISIS Partial Sequence Numbers Protocol Data Unit
    PDU length: 35
    Source-ID: 0000.0000.0002.00
  LSP entries (16)
    LSP-ID: 0000.0000.0003.00-00, Sequence: 0x00000008, Lifetime: 1194s, Checks
```

## 4.7 LSP 洪泛扩散和同步

### 1. 关于泛洪

LSP 报文的“泛洪”指当一个路由器向相邻路由器报告自己的 LSP 后，相邻路由器再将同样的 LSP 报文传送到自己相邻的路由器，并这样逐级将 LSP 传送到整个层次内的一种方式。通过这种“泛洪”，整个层次内的每一个路由器就都可以拥有相同的 LSP 信息，并保持 LSDB 的同步。每一个 LSP 都拥有其自己的一个 4 字节的序列号。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。

每个 LSP 在 LSDB 中都有一个最大经历时间 (MaxAge)，当这个时间到达后如果没有接收到新的 LSP 来更新 LSDB，则这个 LSP 会从 LSDB 中清除。在旧的 LSP 被从 LSDB 中清除后，它还会再保留一段时间 (ZeroAgeLifetime，这期间只保留这个 LSP 的报文头)，当这个时间也达到时它将会被真正删除。

一般缺省的 MaxAge 是 1200 秒，ZeroAgeLifetime 是 60 秒。当一个 IS 所发出的 LSP 的序列号达到 0xFFFFFFFF 时，这个路由器会将 IS-IS 进程暂停 MaxAge + ZeroAgeLifetime 秒，以便在整个路由域内和这个路由器对应的 LSP 都被删除掉。随后 IS-IS 进程再重新启动，并从序列号 1 开始发送 LSP。(这种情况其实不会出现，因为即使每秒钟产生一个 LSP，也要到 136 年后序列号才能从 1 达到 0xFFFFFFFF。)

### 2. LSP 产生的原因

IS-IS 路由域内的所有路由器都会产生 LSP，以下事件会触发一个新的 LSP：

- 邻接关系 Up 或 Down；
- IS-IS 相关接口 Up 或 Down；

- 引入的 IP 路由发生变化；
- 区域间的 IP 路由发生变化；
- 接口被赋了新的 metric 值；
- 周期性更新。

### 3. 路由选择信息扩散和数据库同步

- **SRM ( Send Routing Message )**: 发送路由信息消息
- **SSN ( send Sequence Number )**: 发送序列号标志

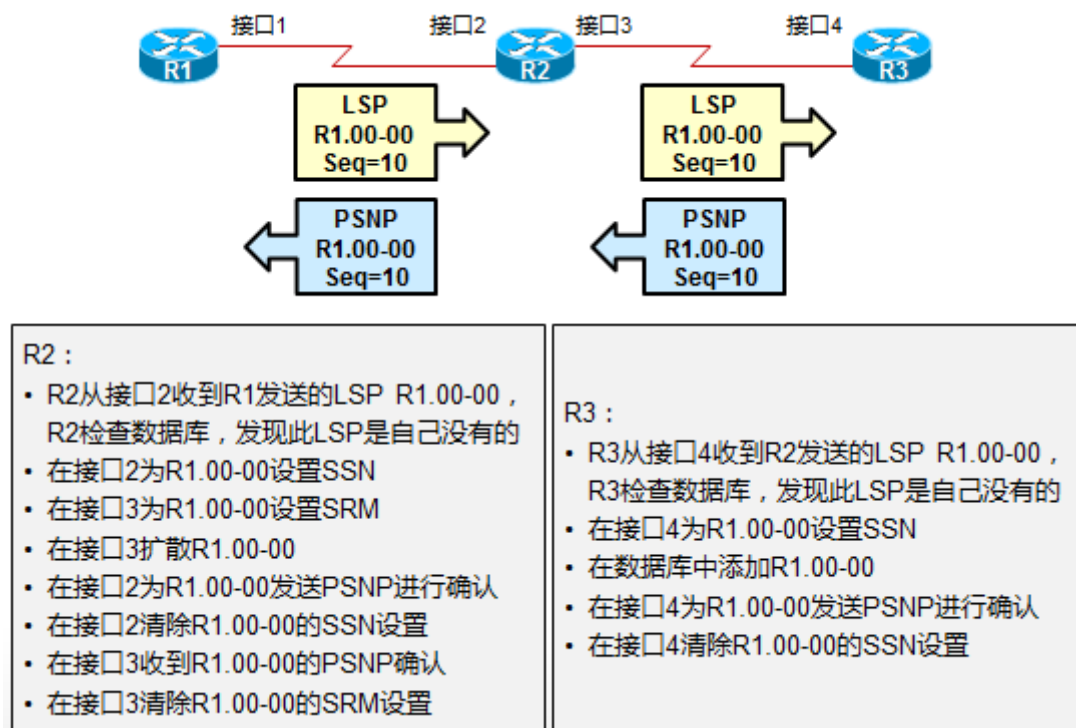
SRM：在更新过程中用来控制 LSP 传递到邻接路由器

SSN：主要用在如下两个方面：

- 在点对点链路上确保 LSP 可靠的扩散
- 广播链路上数据库同步时用于请求完整的 LSP 信息

**SRM 和 SSN 用来协调扩散和数据库同步**

### 4. 点对点链路中路由选择信息扩散

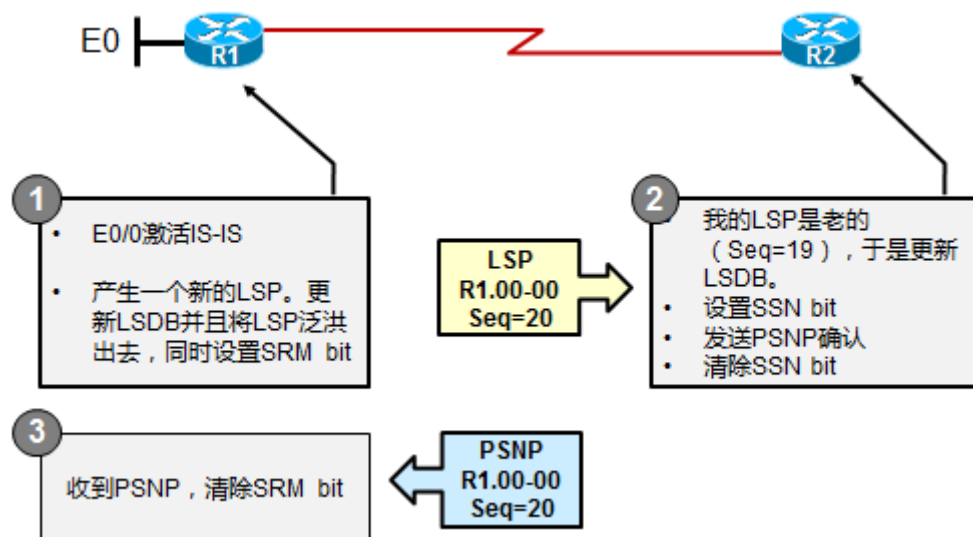


- ISs 将新的 LSP 泛洪给所有的邻居
  - LSP 发出后，SRM 被设置
  - 邻居要使用 PSNP 确认这个 LSP

如果 retransmission timeout 后仍然没有收到 PSNP 确认，则始发路由器将重传这个 LSP

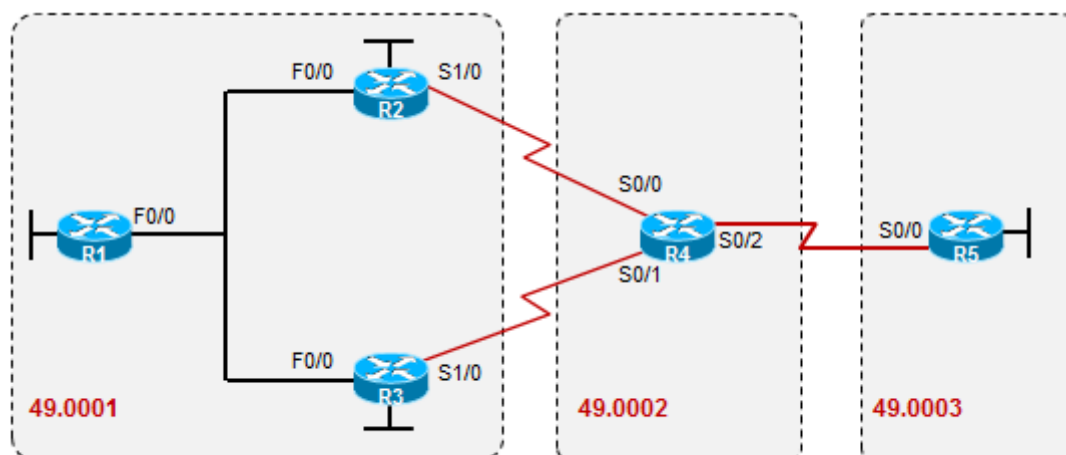
- 在收到关于这个 LSP 的 PSNP 确认后，SRM 被清除
- 在邻接关系建立过程中，邻接双方会互相发送 CSNP（邻接关系建立完成后不再发送 CSNP）
  - 如果收到 CSNP 且发现其中丢失了 LSPs，则路由器会对该丢失的 LSP 进行重传
  - 丢失的 LSPs 可以通过 PSNP 进行请求

再来看一个例子：



## 5 集成 IS-IS 的配置

### 5.1 基础实验





### R1 的配置如下：

#### interface Loopback0

ip address 1.1.1.1 255.255.255.0

ip router isis

!! 在接口上激活 IS-IS

#### interface FastEthernet0/0

ip address 10.1.123.1 255.255.255.0

ip router isis

#### router isis

net 49.0001.0000.0000.0001.00

!! 为设备配置 net 地址，可以配置多个，默认最多 3 个

is-type level-1

!! 配置设备的 is-type，默认为 level1-2

### R2 的配置如下：

#### interface Loopback0

ip address 2.2.2.2 255.255.255.0

ip router isis

#### interface FastEthernet0/0

ip address 10.1.123.2 255.255.255.0

ip router isis

#### interface Serial1/0

ip address 10.1.24.2 255.255.255.0

ip router isis

router isis

net 49.0001.0000.0000.0002.00

### R3 的配置如下：

#### interface Loopback0

ip address 3.3.3.3 255.255.255.0

ip router isis

#### interface FastEthernet0/0

ip address 10.1.123.3 255.255.255.0

ip router isis

#### interface Serial1/0

ip address 10.1.34.3 255.255.255.0

ip router isis

#### router isis

net 49.0001.0000.0000.0003.00

R4 的配置如下：

**interface Serial0/0**

ip address 10.1.24.4 255.255.255.0

ip router isis

**interface Serial0/1**

ip address 10.1.34.4 255.255.255.0

ip router isis

**interface Serial0/2**

ip address 10.1.45.4 255.255.255.0

ip router isis

**router isis**

net 49.0002.0000.0000.0004.00

is-type level-2-only

!! R4 为 L2 router，因此修改 is-type 类型为 level2-only

R5 的配置如下：

**interface Loopback0**

ip address 5.5.5.5 255.255.255.0

ip router isis

**interface Serial0/0**

ip address 10.1.45.5 255.255.255.0

ip router isis

**router isis**

net 49.0003.0000.0000.0005.00

is-type level-2-only

接下去就 show 一下：

**R1#show ip route**

```

          1.0.0.0/24 is subnetted, 1 subnets
C          1.1.1.0 is directly connected, Loopback0
          2.0.0.0/24 is subnetted, 1 subnets
i   L1      2.2.2.0 [115/20] via 10.1.123.2, FastEthernet0/0
          3.0.0.0/24 is subnetted, 1 subnets
i   L1      3.3.3.0 [115/20] via 10.1.123.3, FastEthernet0/0
          10.0.0.0/24 is subnetted, 3 subnets
i   L1      10.1.24.0 [115/20] via 10.1.123.2, FastEthernet0/0

```

```
i   L1      10.1.34.0 [115/20] via 10.1.123.3, FastEthernet0/0
C      10.1.123.0 is directly connected, FastEthernet0/0
i*   L1    0.0.0.0/0 [115/10] via 10.1.123.3, FastEthernet0/0
      [115/10] via 10.1.123.2, FastEthernet0/0
```

我们看到 R1 已经学习到了 2.2.2.0、3.3.3.3.0 以及 10.1.24.0、10.1.34.0。

R2 及 R3 的 loopback 接口来自本区域，因此为 L1 的路由这个没有问题，但是为啥 R2-R4、R3-R4 直连链路也进来了呢？那是因为这些是 R2 和 R3 的直连接口，可以通过在 R2 的 s1/0 及 R3 的 S1/0 口上，将接口配置为 level2，使用接口级别命令：isis circuit-type level2，那么这样一来 R2 及 R3 在区域 49.0001 里通告 LSP 的时候，就不会通告这些这直连链路了。因此 R1 的路由表变成：

#### R1#show ip route

```
      1.0.0.0/24 is subnetted, 1 subnets
C      1.1.1.0 is directly connected, Loopback0
      2.0.0.0/24 is subnetted, 1 subnets
i   L1      2.2.2.0 [115/20] via 10.1.123.2, FastEthernet0/0
      3.0.0.0/24 is subnetted, 1 subnets
i   L1      3.3.3.0 [115/20] via 10.1.123.3, FastEthernet0/0
      10.0.0.0/24 is subnetted, 3 subnets
C      10.1.123.0 is directly connected, FastEthernet0/0
i*   L1    0.0.0.0/0 [115/10] via 10.1.123.3, FastEthernet0/0
      [115/10] via 10.1.123.2, FastEthernet0/0
```

接下来再看看 R1 的 isis database：

#### R1#show isis database

IS-IS Level-1 Link State Database:

LSPID	LSP Seq Num	LSP Checksum	LSP Holdtime	ATT/P/OL
R1.00-00	* 0x0000000B	0xA538	687	0/0/0
R2.00-00	0x0000000F	0x7056	906	1/0/0
R3.00-00	0x0000000E	0x3B83	1096	1/0/0
R3.02-00	0x0000000C	0x4F65	880	0/0/0

在区域 49.0001 我们可以看到所有路由器都泛洪了自己的 LSP，LSP 使用 LSPID 来表示，例如 R1.00-00，这个 R1 就是设备的 hostname，这里我们有默认的 hostname 映射机制，请看上文。另外 R1 后面的 00 是伪节点标示符，为 00 则表示该 LSP 的始发路由器不是 DIS，再后头的 00 则是分片标记。

我们看到上面的输出，R1、R2、R3 都产生了 LSP，打星号的条目为 R1 自己发出的。另外，还有个特殊的 LSP：R3.02-00，这里伪节点标记为非 0，则表示 R3 是一台 DIS，而这个 LSP，是一个伪节点的 LSP。

接下去我们看看这些 LSP 的明细：

#### R1#show isis da R1.00-00 detail

IS-IS Level-1 LSP R1.00-00

LSPID	LSP Seq Num	LSP Checksum	LSP Holdtime	ATT/P/OL
R1.00-00	* 0x0000000C	0xA339	1174	0/0/0

Area Address: 49.0001

NLPID: 0xCC

Hostname: R1

IP Address: 1.1.1.1

Metric: 10 IP 10.1.123.0 255.255.255.0

Metric: 10 IP 1.1.1.0 255.255.255.0

Metric: 10 IS R3.02

可以看到其中包含的主要两个内容：一是 R1 的直连网段，这里是 10.1.123.0 及 1.1.1.0，另一是 R1 的直连 IS，由于这是一个 LAN 的环境，因此会选举 DIS，并且每一台非 DIS 路由器通告自己与 DIS 的邻接关系，而不通告与其他物理 IS 的邻接关系。R3\R3 的 LSP 类似，这里就不再赘述。看看这个伪节点的 LSP 吧：

#### R1#show isis da R3.02-00 detail

IS-IS Level-1 LSP R3.02-00

LSPID	LSP Seq Num	LSP Checksum	LSP Holdtime	ATT/P/OL
R3.02-00	0x0000000C	0x4F65	539	0/0/0

Metric: 0 IS R3.00

Metric: 0 IS R2.00

Metric: 0 IS R1.00

我们看到伪节点 LSP 主要的内容就是连接在这个伪节点上的 IS。

再查看一下 isis 的邻居：

#### R1#sh isis neighbors detail

System Id	Type	Interface	P Address	State	Holdtime	Circuit Id
R2	<b>L1</b>	a0/0	0.1.123.2	UP	3	R3.02

Area Address(es): 49.0001

SNPA: cc02.1ab0.0000 **!! 邻居的 SNPA，在 LAN 中为对方的接口 MAC**

State Changed: 00:20:27

LAN Priority: 64 **!! 用于选举 DIS 的接口优先级**

Format: Phase V

R3	<b>L1</b>	Fa0/0	10.1.123.3	UP	9	R3.02
----	-----------	-------	------------	----	---	-------

```
Area Address(es): 49.0001
SNPA: cc03.1ab0.0000
State Changed: 02:23:31
LAN Priority: 64
Format: Phase V
```

再看一下 R4 的 isis database :

#### R4#sh isis da

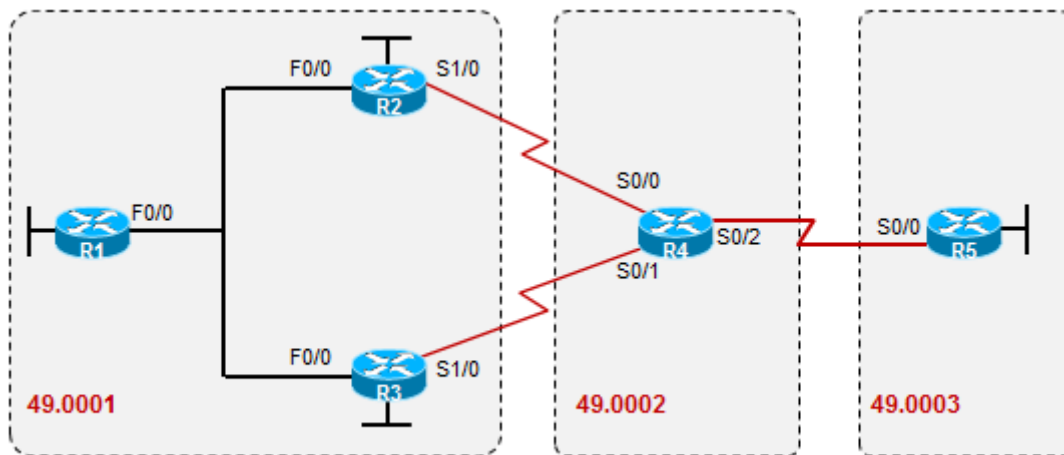
IS-IS Level-2 Link State Database:

LSPID	LSP Seq Num	LSP Checksum	LSP Holdtime	ATT/P/OL
R2.00-00	0x00000011	0x2B7C	558	0/0/0
R3.00-00	0x00000011	0x1B7C	562	0/0/0
R3.02-00	0x0000000C	0x10BA	908	0/0/0
R4.00-00	* 0x0000000F	0xD3FB	569	0/0/0
R5.00-00	0x0000000D	0x26DC	1144	0/0/0

#### R4 的路由表 :

```
1.0.0.0/24 is subnetted, 1 subnets
i L2   1.1.1.0 [115/30] via 10.1.34.3, Serial0/1
        [115/30] via 10.1.24.2, Serial0/0
2.0.0.0/24 is subnetted, 1 subnets
i L2   2.2.2.0 [115/20] via 10.1.24.2, Serial0/0
3.0.0.0/24 is subnetted, 1 subnets
i L2   3.3.3.0 [115/20] via 10.1.34.3, Serial0/1
5.0.0.0/24 is subnetted, 1 subnets
i L2   5.5.5.0 [115/20] via 10.1.45.5, Serial0/2
10.0.0.0/24 is subnetted, 4 subnets
C       10.1.24.0 is directly connected, Serial0/0
C       10.1.45.0 is directly connected, Serial0/2
C       10.1.34.0 is directly connected, Serial0/1
i L2   10.1.123.0 [115/20] via 10.1.34.3, Serial0/1
        [115/20] via 10.1.24.2, Serial0/0
```

## 5.2 Attached-bit



在完成基本配置后，由于 R2、R3 为 L1\L2 router，因此对于他两来说，都会向本地的 level1 area 通告自己为出去的一个口子，方法就是在自己产生的 LSP 中，ATT 置位，我们在 R1 上看 isis database 已经能看出来了。

### R1#show isis database

IS-IS Level-1 Link State Database:

LSPID	LSP Seq Num	LSP Checksum	LSP Holdtime	ATT/P/OL
R1.00-00	* 0x00000000B	0xA538	687	0/0/0
R2.00-00	0x00000000F	0x7056	906	<b>1/0/0</b>
R3.00-00	0x00000000E	0x3B83	1096	<b>1/0/0</b>
R3.02-00	0x00000000C	0x4F65	880	0/0/0

这将直接导致 R1 会从 R2、R3 中选择出最近的一个出口，并且生成一条默认路由指向这个出口，由于这个实验中，到达 R2、R3 等代价，因此 R1 的路由表中，有两条默认路由负载均衡。这种默认的、自动的机制对于 level1 area 来说是福利，因为他可以减小路由表，同时又提供了一个出去的口子，但是这也带来一个问题，例如，如果 R2 或者 R3 连接 backbone 的链路 DOWN 掉了，那么 R1 是否会傻乎乎的还是将数据包丢给 R2 或 R3 呢？庆幸的是，在新的 CISCO IOS 中，对 IS-IS 做了优化：在正常情况下，R2、R3 都会对在 49.0001 区域中泛洪的 LSP 进行 ATT 置位，拿 R2 举例，如果 R2 的 S1/0 口 DOWN 掉了，那么它就丢失了 Backbone 的连接，于是 R2 立即触发一个新的 LSP 并且 ATT 没有置位，如此一来，R1 将原先指向 R2 的默认路由撤销。

当然，上面的方法不保险。所以，在特殊的情况下，我们可能希望 L1\L2 router 有条件性的去设置这个 ATT 位、更加可控的去设置，那么可以关联 route-map，并且在 route-map 中去 match 某条特定的路由，只要这条路由在我路由表里，route-map 才满足，我才会设置 ATT。

例如上图，我们可以在 R2 上做测试，配置如下：

```
ip prefix-list test seq 5 permit 5.5.5.0/24
route-map test permit 10
```

```
match ip address prefix-list test
router isis
net 49.0001.0000.0000.0002.00
set-attached-bit route-map test
```

这样一来，只要 5.5.5.0 在 R2 的路由表中，那么 R2 就会在其向区域 49.0001 中泛洪的 LSP 中进行 ATT 置位，我们可以尝试将 R5 的 loopback shutdown，这样 5.5.5.0 就挂了，当 R2 检测到这个变化，发现路由表中没了 5.5.5.0/24 的路由，那么 R2 将触发一个 LSP 更新，这个 LSP 中 ATT 就没有置位了，这样一来 R1 就只会产生一条指向 R3 的默认路由。当 5.5.5.0 恢复了，那么 R2 又会产生一个 LSP，又将 ATT 置位（这条 5.5.5.0 的路由未必必须是 ISIS 路由，你完全可以耍耍，配一条指向 null0 的路由，丫就 ATT 置位了）。这里有一点要注意的是：用于匹配路由的工具建议采用前缀列表，我在 12.4 的 IOS 上测试过用标准的 ACL 去匹配路由，结果貌似不成功；

另一点要注意的是，只有 L1/L2 router 在保有 L2 连接的情况下，才会自动将 ATT 置位，如果上图中，R2 的 S1/0 口挂掉了，那么 R2 将没有任何 Level2 的邻接关系，也就是说，R2 与 backbone 或者 level2 area 断开了，R2 将立刻触发一个新的 LSP，并且将 ATT 清除，这就导致 R1 不在使用 R2 作为默认路由的下一跳。

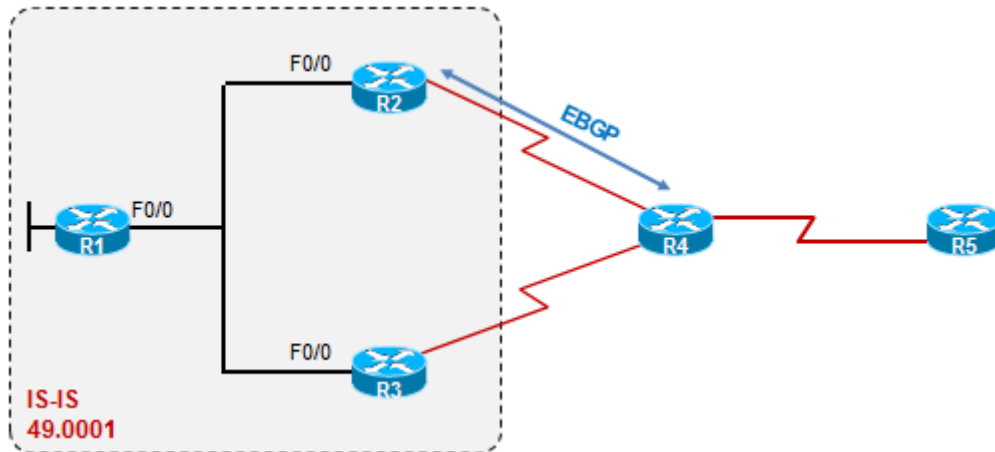
不希望 L1-router 产生这条默认路由，或者更严谨点说，我们不希望 L1\L2 router 压 att 位，方法就是在 L1/L2router 上用 set attached-bit route-map 关联一个 route-map test deny 的 route-map，就可以了。

## 5.3 Overload-bit

一台路由器如果资源不足，它可能就无法维护完整的链路状态数据库，数据流量到达此台设备可能就无法得到正常的转发，因此它有告知或者告警其他路由器的能力，具体的做法是在其自己产生的 LSP 中将 Overload-bit 置位。被设置了 overload 字段的那些 LSP 不会在网络中扩散并且在计算通过 overload 路由器的路由时也不会采用这些 LSP。换句话说，overload 路由器被传输路绕开（Bypass）了，那些以 overload 路由器作为最后一跳的路由（本地直连的网段）才参与计算。再说白点，就是接收这个 LSP 的 IS-IS 路由器，仍然可以产生到达 overload 路由器直连的路由，但是不会通过这台路由器去往其他的目的地。

单纯的由于设备资源不足而去设置 overload 的动机现如今已经不多了，毕竟现在设备的性能越来越好，而且 overload 这个特性毕竟是在设备性能不够强劲的时代的产品。然而 overload 的扩展应用却是非常值得关注的。我们来看一下下面这个图：





假设这是一个运营商内部的网络，R4 是国干，R2、R3 往左是省干，省干内部跑了 IS-IS 用于承载 Core 内的路由前缀信息，并且作为省干出口 R2、R3 都下发了特定的国干的路由前缀或缺省路由，我们这里拿缺省路由举例。R2、R3，都与 R4 维护 EIGRP 邻居关系，那么这时候假设 R2 要做设备升级，然后重启，重启之后由 IS-IS 的收敛速度比 BGP 要快，率先收敛完毕后，R2 向 IS-IS 区域内泛洪 LSP，R1 将使用 R2 作为某些路由前缀的下一跳，或默认路由的下一跳，然而这时候 R2 的 BGP 并未收敛完毕，这样一来，到达 R2 的数据就有可能被丢弃，黑洞就出现了。

我们可以借助 overload-bit 来避免这个问题，命令的解释如下（配置在 router isis 进程下）：

#### Set-overload-bit

本路由器产生的 LSP overload-bit 置位。这将导致这台路由器将不会被其他路由器作为去往任何目的地的下一跳，除了其本地直连的网段。这一般用在路由器是叶节点的时候，也就是说是网络的最末梢，下面直接连的就是用户。

#### Set-overload-bit on-startup <5-86400 sencondes>

后面跟一个秒数，用来说明设备重启需要设置 overload-bit 的时间，超过该时间后，overload-bit 被清除

#### set-overload-bit on-startup wait-for-bgp

只有等到 BGP 收敛完成后，overload-bit 才会被清除，这个特性不是特别的推荐。感觉不太靠谱。

#### set-overload-bit suppress external

不通告从其他路由协议重发布进来的路由

#### set-overload-bit suppress interlevel

不通告从其他 level 泄露进来的路由

### 1. 测试 1：set-overload-bit

在 R2 上配置如下：

```
Router isis
set-overload-bit
```

这将导致 R2 产生的 LSP 的 OL 位置位，在 R1 上看到的結果是这样的：

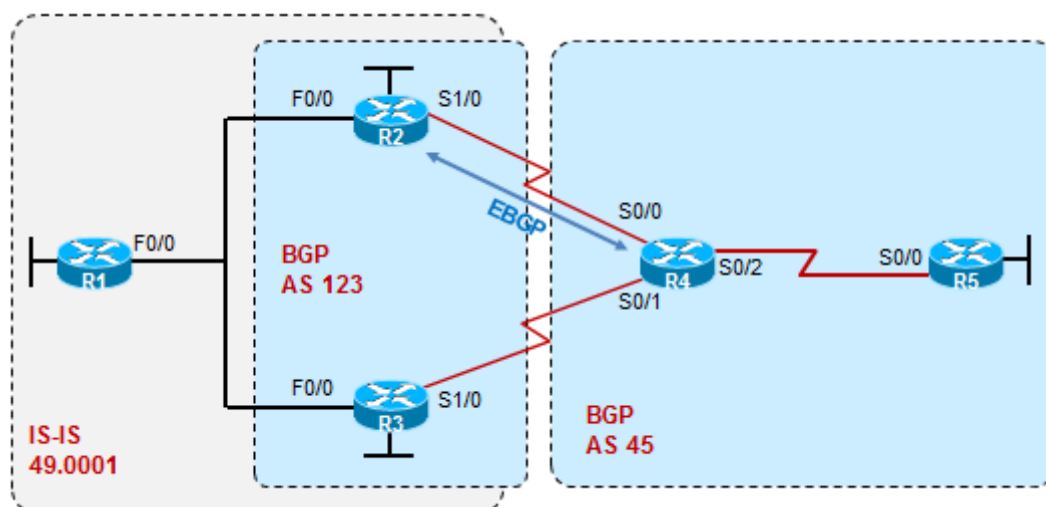
IS-IS Level-1 Link State Database:

LSPID	LSP Seq Num	LSP Checksum	LSP Holdtime	ATT/P/OL
R1.00-00	* 0x00000013	0x9540	724	0/0/0
R2.00-00	0x00000017	0x5C66	1178	<b>0/0/1</b>
R3.00-00	0x00000014	0x2F89	668	1/0/0
R3.02-00	0x00000010	0x4769	722	0/0/0

我们发现，R2 产生的 LSP 之前的 attached-bit 没了，但是设置了 Overload-bit。现在我们去查看路由表，R1 再计算最短路径的时候，就不会考虑通过 R2 出区域。于是 R1 上仅剩指向 R3 的默认路由。当然，R2 直连的路由还是能在 R1 上看到的。

## 2. 测试 2：set-overload-bit on-startup wait-for-bgp

我们要变更一下拓扑：



在 R2 上配置如下：

```
router isis
net 49.0001.0000.0000.0002.00
set-overload-bit on-startup wait-for-bgp
router bgp 123
no synchronization
neighbor 10.1.24.4 remote-as 45
no auto-summary
```

保存配置后重启 R2，在 R2 重启完成后，isis 率先收敛，然而这个时候 bgp 暂未收敛完成，因此 R2 产生的 LSP 中 overload 置位，从 R1 上能看出结果如下，在此期间，R1 不会将 R2 作为去往其他网络的下一跳（R2 直连的网络除外）：

#### R1 上 show isis database

IS-IS Level-1 Link State Database:

LSPID	LSP Seq Num	LSP Checksum	LSP Holdtime	ATT/P/OL
R1.00-00	* 0x0000002B	0x2995	1112	0/0/0
R2.00-00	0x00000036	0xE7D9	1178	<b>0/0/1</b>
R3.00-00	0x00000029	0x18B7	968	0/0/0
R3.01-00	0x00000026	0x2D6E	1019	0/0/0

随后我们等待 R2 上 BGP 的邻居关系建立，可以 debug 一下：

#### R2 上 debug isis update

**\*Mar 1 00:00:32.639: %BGP-5-ADJCHANGE: neighbor 10.1.24.4 Up**

R2#sh run |

**\*Mar 1 00:00:32.755: ISIS-Upd: Building L1 LSP**

**\*Mar 1 00:00:32.759: ISIS-Upd: Important fields changed**

\*Mar 1 00:00:32.759: ISIS-Upd: full SPF required

\*Mar 1 00:00:32.759: ISIS-Upd: Building L2 LSP

\*Mar 1 00:00:32.763: ISIS-Upd: Important fields changed

\*Mar 1 00:00:32.763: ISIS-Upd: full SPF required

\*Mar 1 00:00:32.783: ISIS-Upd: Sending L1 LSP 0000.0000.0002.00-00, seq 37, ht 1199 on FastEthernet0/0 via high priority queue

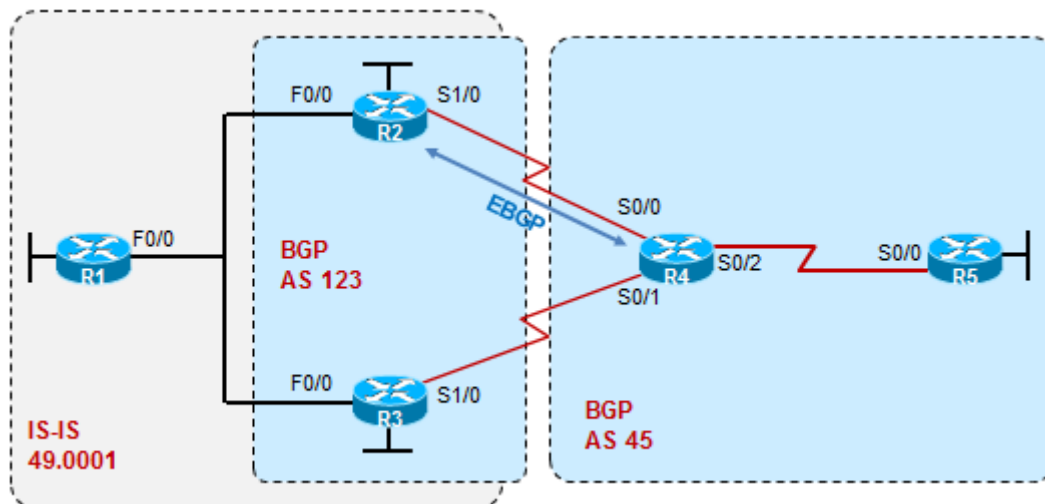
我们看到，BGP 邻居关系 up 后，R2 马上触发一个 LSP，这个 LSP 清除之前的 overload 置位。

#### 最终 R1 上 show isis database

IS-IS Level-1 Link State Database:

LSPID	LSP Seq Num	LSP Checksum	LSP Holdtime	ATT/P/OL
R1.00-00	* 0x0000002B	0x2995	1103	0/0/0
R2.00-00	0x00000037	0xE1E2	1194	0/0/0
R3.00-00	0x00000029	0x18B7	959	0/0/0
R3.01-00	0x00000026	0x2D6E	1010	0/0/0

### 3. 测试 3：set-overload-bit on-startup 120

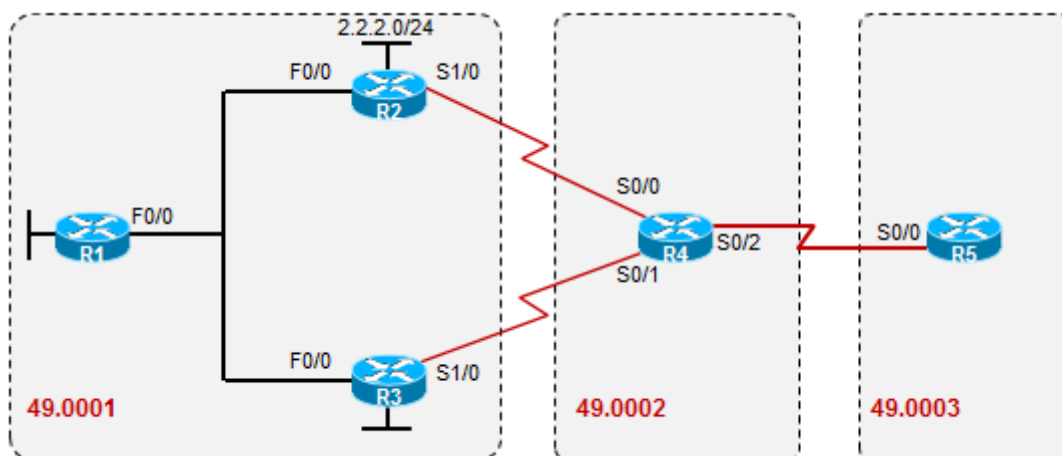


还是这个图，不过 R2 的配置变成：

```
router isis
net 49.0001.0000.0000.0002.00
set-overload-bit on-startup 120           !! 120s
router bgp 123
no synchronization
neighbor 10.1.24.4 remote-as 45
no auto-summary
```

保存配置后重启，在重启后，R2 产生的 LSP 先是设置了 overload-bit 的，这点从 R1 上 isis database 就能看到，我们在 R2 上同样开启 debug isis update，同时使用 show version | in uptime 查看路由器的启动时间，当启动时间到达 120s 后，从 debug 信息中能看到 R2 触发了一个 LSP，并且清除了 overload-bit。

#### 4. 测试 4：set-overload-bit suppress external



- Point1：

### R2 的配置如下：

```
route-map test permit 10
router isis
 net 49.0001.0000.0000.0002.00
 redistribute connected level-1           !!将本地直连的 loopback 2.2.2.0/24 重发布到 level1
 redistribute isis ip level-2 into level-1 route-map test
```

### R1 的路由表如下：

```
1.0.0.0/24 is subnetted, 1 subnets
C      1.1.1.0 is directly connected, Loopback0
      2.0.0.0/24 is subnetted, 1 subnets
i L1   2.2.2.0 [115/10] via 10.1.123.2, FastEthernet0/0
      5.0.0.0/24 is subnetted, 1 subnets
i ia   5.5.5.0 [115/158] via 10.1.123.2, FastEthernet0/0
      10.0.0.0/24 is subnetted, 4 subnets
i ia   10.1.24.0 [115/148] via 10.1.123.2, FastEthernet0/0
i ia   10.1.45.0 [115/158] via 10.1.123.2, FastEthernet0/0
i ia   10.1.34.0 [115/158] via 10.1.123.2, FastEthernet0/0
C      10.1.123.0 is directly connected, FastEthernet0/0
i*L1 0.0.0.0/0 [115/10] via 10.1.123.3, FastEthernet0/0
      [115/10] via 10.1.123.2, FastEthernet0/0
```

现在是初始环境，R1 能学习到 inter-area 路由，以及 R2 本地重发布到 level1 的路由 2.2.2.0。当然由于 R2 设置了 attached-bit，因此 R1 也会产生一条指向 R2 的默认路由。

## ● Point2

### R2 的配置修改如下：

```
route-map test permit 10
router isis
 net 49.0001.0000.0000.0002.00
 set-overload-bit           !! 设置 overload-bit
 redistribute connected level-1           !!将本地直连的 loopback 2.2.2.0/24 重发布到 level1
 redistribute isis ip level-2 into level-1 route-map test
```

### 再看看 R1 的路由表：

```
1.0.0.0/24 is subnetted, 1 subnets
C      1.1.1.0 is directly connected, Loopback0
```

```

2.0.0.0/24 is subnetted, 1 subnets
i L1    2.2.2.0 [115/10] via 10.1.123.2, FastEthernet0/0
5.0.0.0/24 is subnetted, 1 subnets
i ia    5.5.5.0 [115/158] via 10.1.123.2, FastEthernet0/0
10.0.0.0/24 is subnetted, 4 subnets
i ia    10.1.24.0 [115/148] via 10.1.123.2, FastEthernet0/0
i ia    10.1.45.0 [115/158] via 10.1.123.2, FastEthernet0/0
i ia    10.1.34.0 [115/158] via 10.1.123.2, FastEthernet0/0
C       10.1.123.0 is directly connected, FastEthernet0/0
i*L1 0.0.0.0/0 [115/10] via 10.1.123.3, FastEthernet0/0

```

我们看到 R1 这些 inter-area 路由都还在，但是 R1 不再指默认路由到 R2。

注意，这个时候 R2 产生的 LSP 已经设置了 overload-bit 了，在 R1 上 show isis database 能看到

### ● Point3

在上述配置的基础上，修改 R2 的配置：

```

route-map test permit 10
router isis
net 49.0001.0000.0000.0002.00
set-overload-bit suppress external      !! 设置 overload-bit，抑制外部路由
redistribute connected level-1          !!将本地直连的 loopback 2.2.2.0/24 重发布到 level1
redistribute isis ip level-2 into level-1 route-map test

```

再看看 R1 的路由表：

```

1.0.0.0/24 is subnetted, 1 subnets
C       1.1.1.0 is directly connected, Loopback0
5.0.0.0/24 is subnetted, 1 subnets
i ia    5.5.5.0 [115/158] via 10.1.123.2, FastEthernet0/0
10.0.0.0/24 is subnetted, 4 subnets
i ia    10.1.24.0 [115/148] via 10.1.123.2, FastEthernet0/0
i ia    10.1.45.0 [115/158] via 10.1.123.2, FastEthernet0/0
i ia    10.1.34.0 [115/158] via 10.1.123.2, FastEthernet0/0
C       10.1.123.0 is directly connected, FastEthernet0/0
i*L1 0.0.0.0/0 [115/10] via 10.1.123.3, FastEthernet0/0

```

R2 重发布的直连路由 2.2.2.0 没了，inter-area 路由都还在并且，R1 不再使用 R2 作为默认路由的下一跳，注意，这时候 R2 发的 LSP 也是 overload-bit 置位了。

注意，此时此刻，R4 上也不会学习到 2.2.2.0/24 了，这条外部路由在 R2 上被彻底抑制住了。

## 5. 测试 5 : set-overload-bit suppress interlevel

修改 R2 的配置：

```
route-map test permit 10
router isis
 net 49.0001.0000.0000.0002.00
 set-overload-bit suppress interlevel      !! 设置 overload-bit，抑制区域间路由
 redistribute connected level-1           !! 将本地直连的 loopback 2.2.2.0/24 重发布到 level1
 redistribute isis ip level-2 into level-1 route-map test
```

看一下 R1 的路由表：

```
1.0.0.0/24 is subnetted, 1 subnets
C      1.1.1.0 is directly connected, Loopback0
2.0.0.0/24 is subnetted, 1 subnets
i L1   2.2.2.0 [115/10] via 10.1.123.2, FastEthernet0/0
10.0.0.0/24 is subnetted, 2 subnets
i ia   10.1.24.0 [115/148] via 10.1.123.2, FastEthernet0/0
C      10.1.123.0 is directly connected, FastEthernet0/0
i*L1 0.0.0.0/0 [115/10] via 10.1.123.3, FastEthernet0/0
```

R2 重发布的 2.2.2.0 外部路由还在；inter-area 路由都没了，除了一条 R2 本地直连的 24.0；

另外，R1 仍然不会采用 R2 作为默认路由的下一跳。

还是那句话，只要使用了 set-overload-bit，R2 发出的 LSP 的 overload-bit 就会置位

## 6. 测试 6 : set-overload-bit suppress interlevel external

修改 R2 的配置：

```
route-map test permit 10
router isis
 net 49.0001.0000.0000.0002.00
 set-overload-bit suppress interlevel external
 redistribute connected level-1           !! 将本地直连的 loopback 2.2.2.0/24 重发布到 level1
 redistribute isis ip level-2 into level-1 route-map test
```

看一下 R1 的路由表：

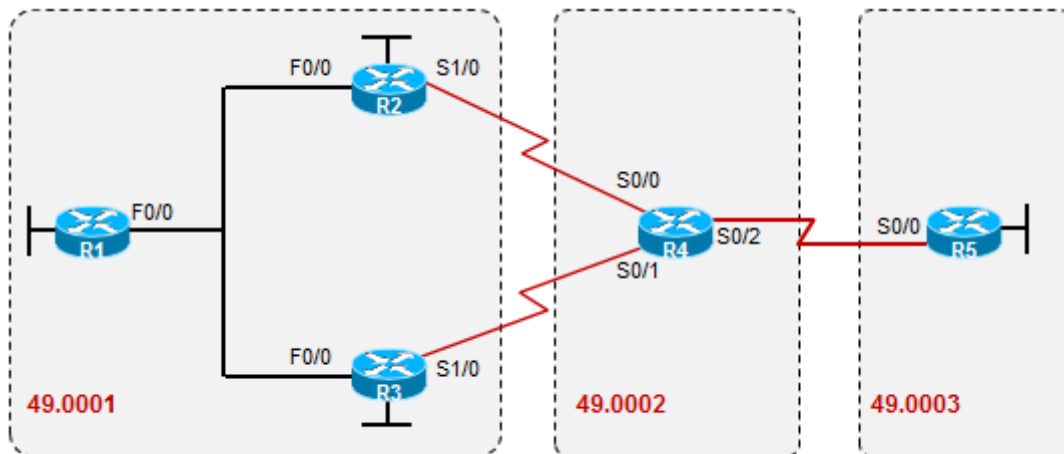
```
1.0.0.0/24 is subnetted, 1 subnets
```



```
C      1.1.1.0 is directly connected, Loopback0
      10.0.0.0/24 is subnetted, 2 subnets
i ia   10.1.24.0 [115/148] via 10.1.123.2, FastEthernet0/0
C      10.1.123.0 is directly connected, FastEthernet0/0
i*L1 0.0.0.0/0 [115/10] via 10.1.123.3, FastEthernet0/0
```

我们发现，R2 本地重发布的外部路由没了；R2 从 level2 注入 level1 的 inter-area 路由也没了；同时 R2 也不能作为 R1 默认路由的下一跳。只有一条 R2 本地直连的 24.0 路由。

## 5.4 路由泄露



我们知道默认的情况下，IS-IS 的 level1 area 很像一个 stub（但是人家允许重发布外部路由进来，所以只是说像），area 内的 L1 router 通过一条默认路由来到达区域外，这就有可能产生次优路径，因为他找的是本区域最近的 L1\L2 router。那么，为了解决这个问题，我们可以采用路由泄露的方法，例如在 R2 上配置：

```
access-list 1 permit 5.5.5.0
route-map test permit 10
  match ip add 1
router isis
redistribute isis ip level-2 into level-1 route-map test
```

如此一来，L2的路由中，被route-map匹配的路由，就会被注入到level1 area 中，那么R1就能学习到5.5.5.0。类型为ia，也就是isis的interarea路由，即isis区域间路由。

**R1#show ip route 5.5.5.0**

```
Routing entry for 5.5.5.0/24
  Known via "isis", distance 115, metric 168, type inter area
```

Redistributing via isis

Last update from 10.1.123.2 on FastEthernet0/0, 00:01:32 ago

Routing Descriptor Blocks:

\* 10.1.123.2, from 2.2.2.2, via FastEthernet0/0

Route metric is 168, traffic share count is 1

同时为了避免环路, R2 在向 level1 area 注入这些路由的时候, 会打上 downbit, 这样一来, R3 收到这条 LSP, 针对这些路由, 就不会再泛洪到 backbone 里。

```

ISO 10589 ISIS Link State Protocol Data Unit
  PDU length: 100
  Remaining lifetime: 1199
  LSP-ID: 0000.0000.0002.00-00
  Sequence number: 0x0000001c
  + Checksum: 0xca33 [correct]
  + Type block(0x0b): Partition Repair:0, Attached bits:1, ov
  + Area address(es) (4)
  + Protocols supported (1)
  + Hostname (2)
  + IP Interface address(es) (4)
  + IP Internal reachability (24)
  + IS Reachability (12)
  - IP Internal reachability (12)
    - IPv4 prefix: 5.5.5.0/24
      Default Metric: 30, Internal, Distribution: down
      Delay Metric: Not supported
      Expense Metric: Not supported
      Error Metric: Not supported
  
```

## 5.5 查看及验证

log-adjacency-changes

强烈建议开启此命令, 可以查看到 isis 邻接关系的一些 log, 默认是关闭的

- show clns neighbors
- show clns int
- show clns protocol
- show isis database
- show ip route isis
- show isis spf-log
- show isis lsp-log

## 5.6 多区域集成 IS-IS 的配置

1. 缺省地, Cisco IOS 在 IS-IS router 上同时启动 L1 和 L2 的操作, 如果要指定 router 仅仅作为区域 router 或主干 router, 则在路由器配置模式下用:

**is-type** 命令 如 is-type level-1

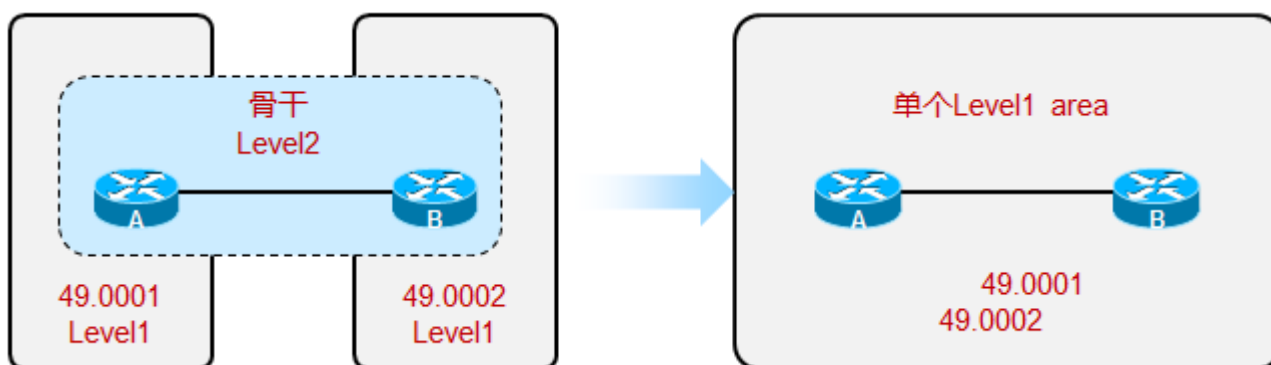
2. 尽管 L1/L2 router 同时具有 L1 和 L2 的功能, 但有时在特定的 接口上只需建立第 1 层毗邻关系, 而在其他接口上配置第 2 层毗邻关系。

在接口配置模式下: isis circuit-type 可以让接口工作在 level-1 level-1-2 或 level-2-only

3. 通过接口配置命令: isis metric value **level-1 | level-2** 来配置 metric 在同一接口上, 也可以分别为第 1 层和第 2 层设定不同的度量值。

## 5.7 为单个 IS-IS 进程配置多个 NET

一般来讲, 一个 IS-IS 进程 仅需要一个 NET, 当然你也可以为一个 IS-IS 进程分配多个 areaID 不同的 NET 地址, 注意这些 NET 地址的 SystemID 必须相同。这样一来这台路由器可以连接到多个区域, 在我们需要把不同区域合并为一个区域时, 这个方法很不错。一般来说, 一台 Level1 路由器仅仅在其连接到的区域内参与扩散 L1 LSP, 如果这台路由器连接了多个区域, 那么可以实现 L1 LSP 在多个区域的泛洪, 从而完成区域的有效合并。



看上面这个例子: A 在区域 49.0001, B 在 49.0002, 由于他们在不同的区域, 因此彼此成了 Level2 邻接关系。同时维护本区域内的邻接关系。如果 A 配置上:

router isis

```
net 49.0001.0000.0000.0001.00
```

```
net 49.0002.0000.0000.0001.00
```

那么 A-B 之间就形成了 Level1 邻接, 因为此刻它两同属一个区域。A 继续维护与区域 49.0001 的 Level1 邻接, B 也继续维护与区域 49.0002 的 Level1 邻接, 由于在 AB 之间形成了新的 Level1 邻接, 设备开始叫唤最初他们互相隔离的 Level1 数据库, 并扩散到各自的 Level1 邻居, 这样就实现了区域的合并。

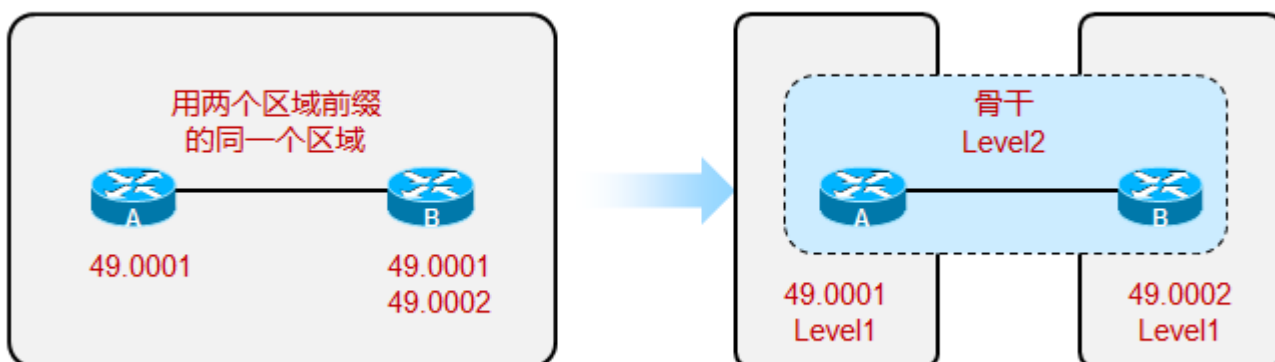
所以利用这个特性 ( 一个 IS-IS 进程配置多个 NET 地址 ), 可以实现 :

- 区域合并
- 区域分离
- 重编址

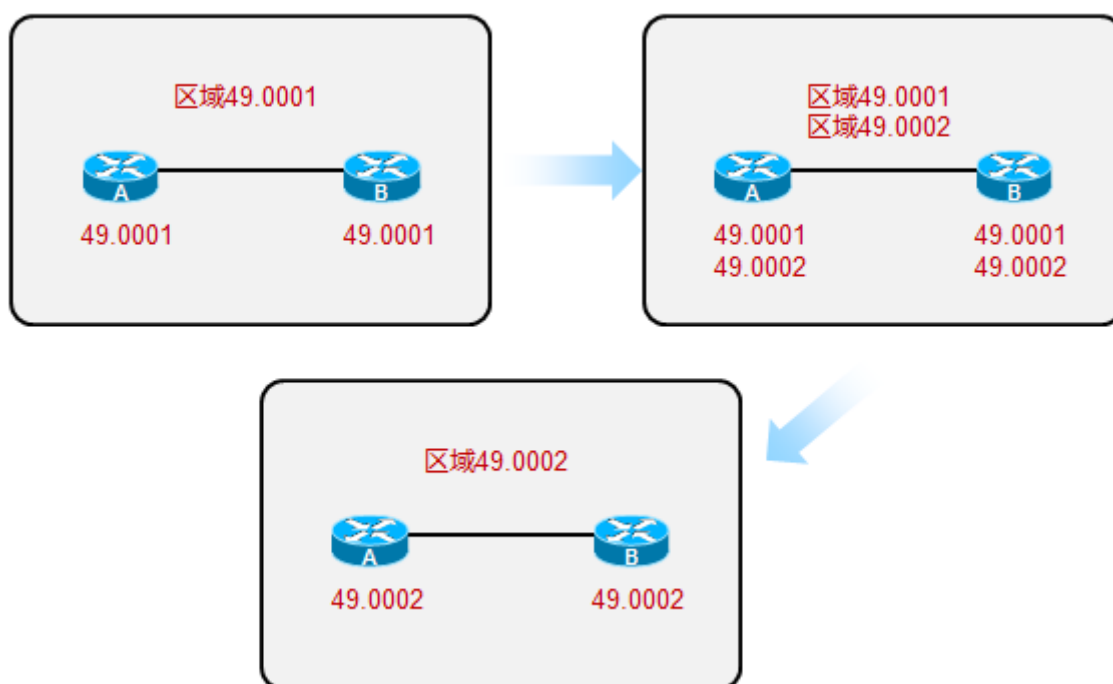
### 1. 区域合并

上面已经说过了

### 2. 区域分离



### 3. NSAP 地址重编址



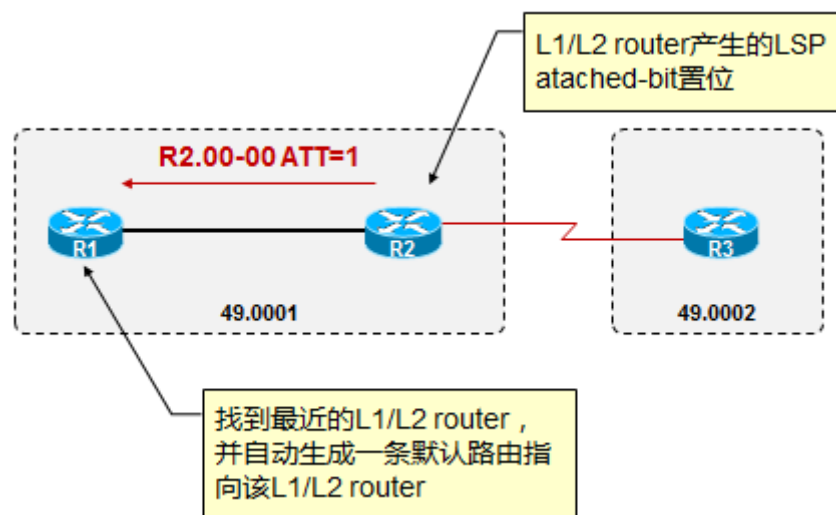
重编址过程和区域合并、分离类似，只是在重编址期间，需要清除一些或者全部路由器的区域前缀，用新的

区域前缀替换。如上图，希望把原先的区域前缀 49.0001 替换成 49.0002，为了实现这个目标，在 A 和 B 上配置一个新区域前缀的 NSAP 地址，注意这两个地址的 SystemID 必须一致。然后依次删除 A 和 B 的区域 ID 49.0001 的 NSAP 地址，这样就实现了路由器新的 NSAP 地址的无缝、无冲突和无破坏的重新配置。

## 5.8 默认路由的注入

在最初的 IS-IS 设计中，Level1 area 是 stub，L1 router 会寻找最近的 L1/L2 router，并生成一条默认路由指向该 L1/L2 router（前提是该 L1/L2 router 进行了 ATT 置位）；然而连接到 Backbone 的 Level2 area 路由器要了解 IS-IS domain 内的所有路由而不会自动产生默认路由。我们可以使用 default-information originate 命令来手工下发一条默认路由进 backbone。该命令产生的路由被装载进 level2 LSP 并泛洪给 backbone 的其他 IS-IS 路由器。

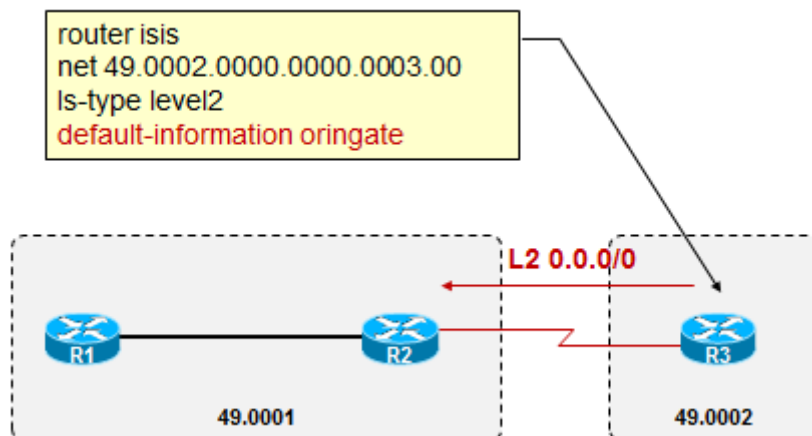
### 1. Level1 area 默认路由



关于 ATT-bit 前面已经讲很多了，这里就不啰嗦了。我在 CISCO IOS Version 12.4(10) -- C3640-JK9O3S-M 上做过测试，在 R1 也就是 L1 router 上，配 default-information originate，无法生成默认路由。

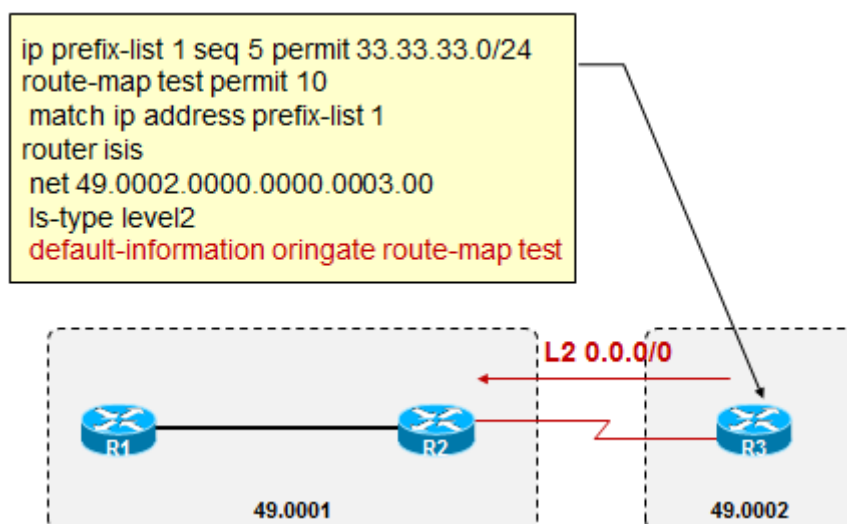
在 L1/L2 router 也就是 R2 上配置，会产生默认路由，并泛洪给 R3，但是不会给 R1。

### 2. Level2 area 默认路由



在 R3 上做如图所示的配置，R3 本地路由表无需默认路由，即可产生一条 IS-IS 默认路由并且传递到 backbone，因此 R2 能学习到。这个时候突发奇想，在 R2 上做 level2 路由到 level1 路由的重发布，发现 10.1.23.0 也就是 R2-R3 的直连链路都被导入进了 49.0001，但是 R3 产生的这条默认路由却没有一同进来。

### 3. Level2 area 默认路由（条件通告）



使用 default-information originate 关联一个 route-map，来进行条件通告，只有当 route-map 匹配成立的时候，才会下发默认路由，可以在 route-map 中关联前缀列表匹配特定的路由前缀，只有当 R3 路由表里有这条路由时，才通告默认路由。

在 CISCO IOS 12.4 上 Acl 用标准和扩展的访问控制列表试过了 不成功 就算是 acl permit any 都不行。用前缀列表匹配路由就成功了。

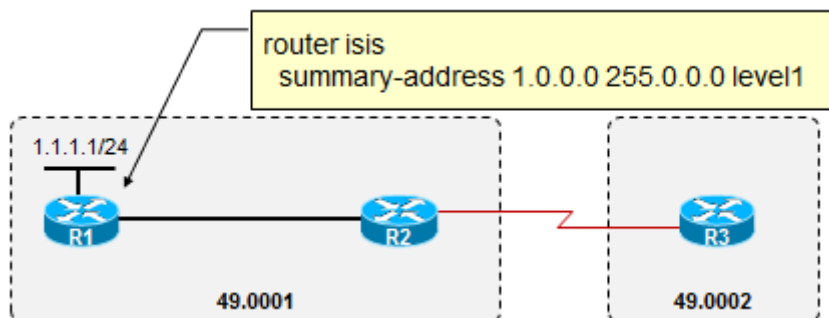
## 5.9 路由汇总

An IS-IS router can be configured to summarize IP routes into Level 1, Level 2, or both, at the same time,

with the following router-level configuration command: **summary-address <prefix> [level-1|level-2|level-1-2]**.

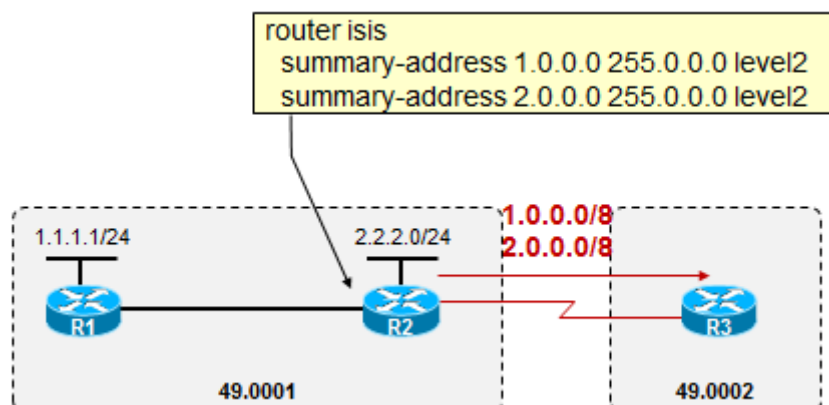
By default, summaries go into Level 2 if no routing level option is indicated.

### 1. 在 L1 router 上能否进行路由汇总？



在 R1 这台 L1 router 上企图进行路由汇总，结果是失败的，这点和 OSPF 很像，区域内的路由器是无法做路由汇总的，因为我们要泛洪 LSP，区域内的路由器必须知道区域里拓扑的详细情况，如果允许 L1 路由器进行路由汇总，岂不就失去了链路状态路由协议的天性？

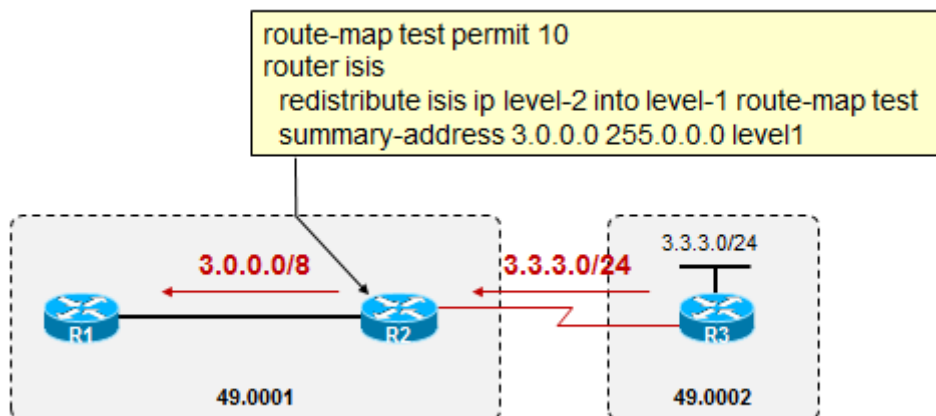
### 2. 在 L1/L2 router 上进行路由汇总 ( to level2 )



在 L1/L2 router 上，对其下属 Level1 area 的路由进行汇总，使得 backbone 或 level2 area 内学习到汇总路由，配置如上。

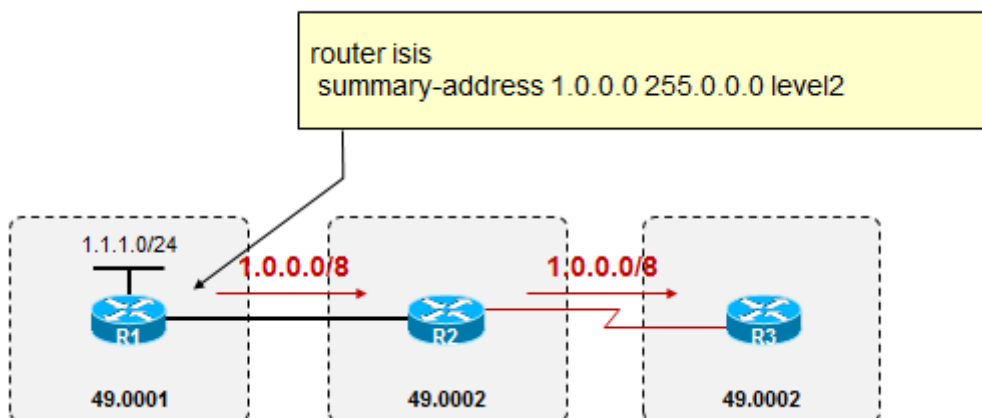
### 3. 在 L1/L2 router 上进行路由汇总 ( to level1 )





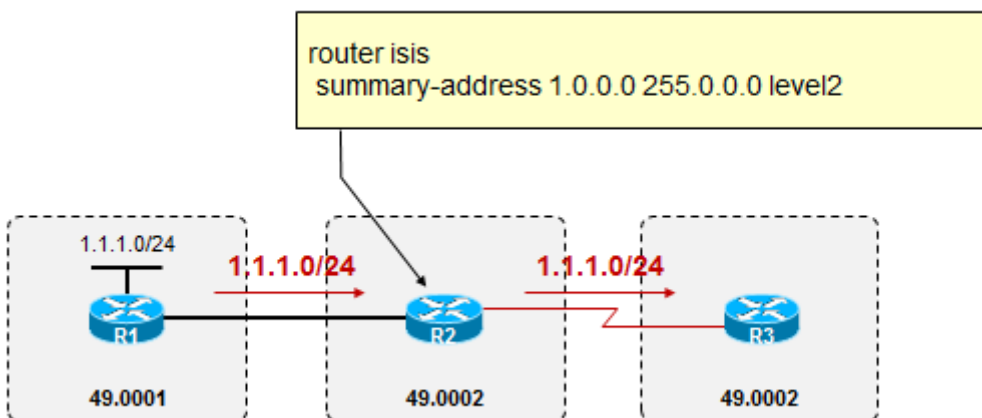
将 level2 路由注入 level1 area , 同时在 L1/L2 router 上可以对这些 level2 路由进行汇总, 配置如上。

#### 4. 在 L2 router 上进行路由汇总 例 1



在 L2 router 上进行汇总, 对本地始发的路由做汇总, 配置如上

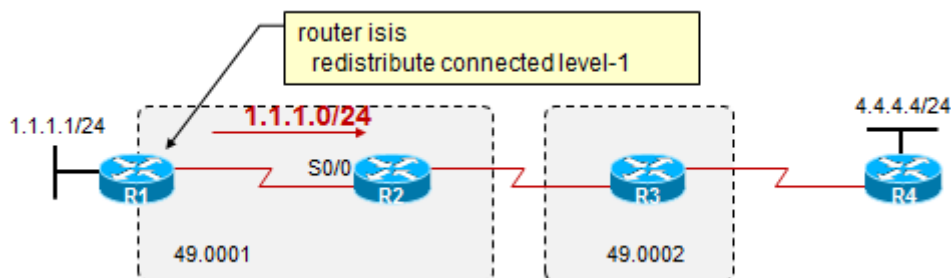
#### 5. 在 L2 router 上进行路由汇总 例 2



必须在 L2 路由的始发路由器上进行路由汇总，否则无效

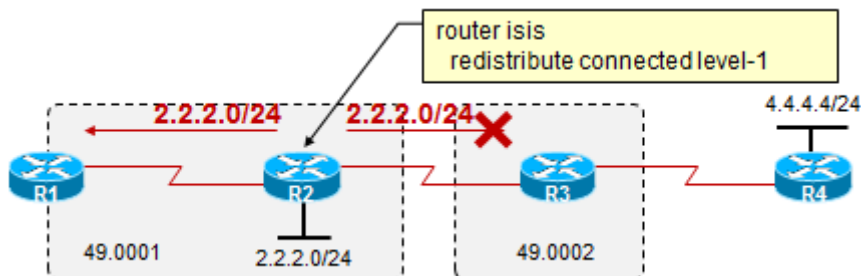
## 5.10 路由重发布

### 1. L1 router 重发布外部路由进 IS-IS



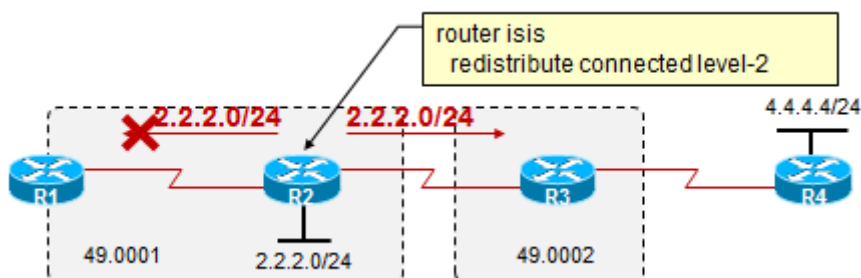
- 在 isis 进程模式中，redistribute 命令直接敲不加 level 参数的话，默认是重发布到 level-2
- 如果不手工设置 metric，重发布进来后默认 metric 为 0，也就是在 R2 上路由的 metric 为 R2 的 S0/0 口的 level1 metric

### 2. L1/L2 router 重发布外部路由进 IS-IS ( to level1 )



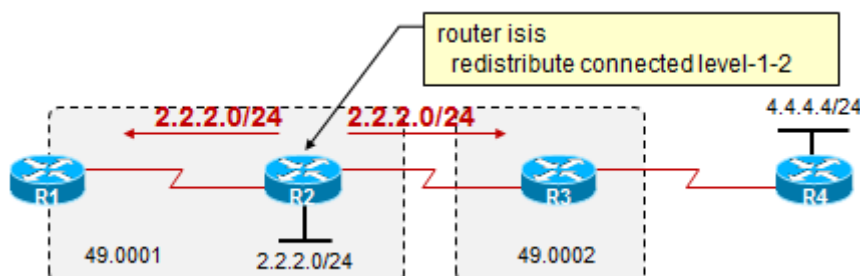
如果在 L1/L2 router 上重发布外部路由到 Level1，则 Level1 area 中能学习到这条路由，但是 R2 不会将路由注入 Level2 area 也就是 Backbone 中

### 3. L1/L2 router 重发布外部路由进 IS-IS ( to level2 )



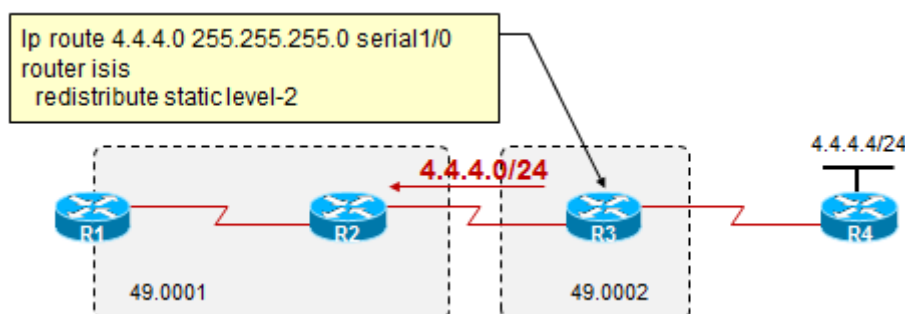
如果重发布命令不加 level 参数, 则默认是如图所示的 level-2 参数, 也就是将外部路由注入到 level2 中, 这样一来 level2 area 都能学习到这条外部路由, 但是 level1 area 中是无法学习到的, 因为 R2 在其产生的 level1 LSP 中不会携带该外部路由。

#### 4. L1/L2 router 重发布外部路由进 IS-IS ( to level-1-2 )

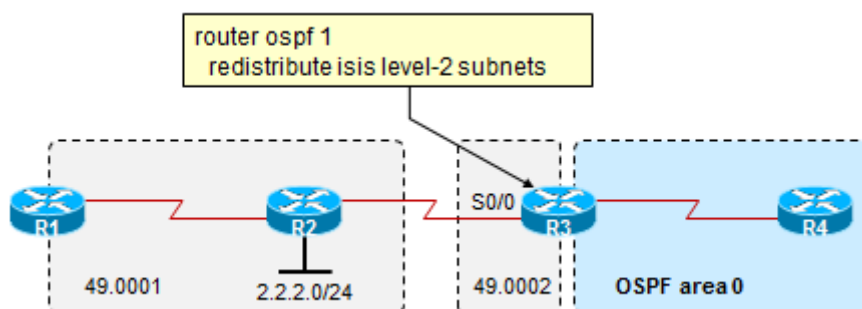


L1\L2 router 在 Level1 area 及 Bacbbone 都会注入这条外部路由

#### 5. L2 router 重发布外部路由进 IS-IS



#### 6. IS-IS level2 to OSPF



这条命令只将 R3 上的 L2 路由重发布进 OSPF, 但是 R3 激活 IS-IS 的 S0/0 直连口却不会被重发布进 OSPF, 这点与其他 IGP 协议不太一样。因此, 如果要让路由学全, 还需要在 R3 上增加一条重发布直连进 OSPF

## 5.11 安全及认证

IS-IS 支持明文及密文 MD5 认证。下面做个小小的汇总：

- **明文接口认证**

我们可以为 IS-IS 的 level1 或 level2 分配一个接口认证密码。如果不指定 level2，则默认的 level1 的认证。如果接口开启了明文认证，那么密码会被插入到所有的 IS-IS 报文中：IIH、LSP、CSNP、以及 PSNP（基于特定的 level）。实际在 CISCO 设备上抓包的时候，发现只有 IIH 中携带了明文的密码信息（放在用于认证的 TLV 中），其他报文没有看到。

命令如下（接口级别）：

```
Isis password <string>
```

- **明文 area 认证**

区域验证，其实就是 Level1 area 的验证。在一台 IS-IS 路由器上增加了区域验证后，其将会检查自己收到的 IS-IS 报文，如果报文中没有验证信息或者错误，则忽略。而其自身在产生的 IS-IS 报文则会携带验证信息，如果接收方没有配置区域认证，在接收到这些携带了验证信息的报文后，仍然可以正常的学习。Area 认证的认证信息只会包含在 LSPs 报文中，其他报文不包含。命令如下（路由进程中）：

```
area-password <string>
```

- **明文 domain 认证**

是在 Level2 LSP、CSNP、PSNP 中插入密码。命令如下（路由进程中）：

```
domain-password <string>
```

在 CISCO IOS 上，我经过抓包验证，发现使用 domain 认证，IIH 报文、CSNP、PSNP 是不携带认证信息的，LSP 则携带用于认证的 TLV

- **密文接口认证**

原理与明文接口认证差不多，只不过可以给予 HMAC-MD5 认证

- **密文 LSPs 认证**

在 LSP 中携带密文的认证信息，可以指定 level1 或 level2 的密文 lsp 认证

下面一个个的实验一下：

### 1. 明文接口认证



我们在 R1 及 R2 之间开启接口认证：

**R1 的配置如下（其他基本的配置就直接省略了）：**

```
Interface fa0/0
  isis password cisco
```

在完成这个配置后，来看一下现象：

**R1#show isis neighbors**

System Id	Type	Interface	IP Address	State	Holdtime	Circuit Id
-----------	------	-----------	------------	-------	----------	------------

R1 的 isis 邻居表直接是空的，这是因为它忽略了这个接口上收到的 R2 的 hello 包。

**R1#debug isis adj-packets**

IS-IS Adjacency related packets debugging is on

R1#

```
*Mar 1 00:09:39.243: ISIS-Adj: Rec L1 IIH from cc01.0a44.0000 (FastEthernet0/0), cir type L1, cir id
0000.0000.0002.02, length 1497
```

```
*Mar 1 00:09:39.243: ISIS-Adj: Authentication failed
```

而在 R2 上

**R2#sh isis neighbors**

System Id	Type	Interface	IP Address	State	Holdtime	Circuit Id
R3	L2	Se1/0	10.1.23.3	UP	28	00
R1	L1	Fa0/0	10.1.12.1	INIT	22	R2.02

R2 已经发送了 IIH，但是 R1 不承认，因为 R2 没有配置认证，这将导致 R1 不会把 R2 放入自己的 IIH 中发回给 R2（包含在 TLV 的 IS neighbors 中）。在 R2fa0/0 口上配上密码，邻接关系就正常了。通过抓包可以看到在 IIH 中已经包含了用于认证的 TLV，认证密码就在 TLV 中以明文的形式存在。在 LSP、CSNP、PSNP 报文中未看到认证的 TLV。

## 2. 明文 area 认证



**R1 的配置如下（R2 暂时不开启 area password）：**

```
router isis
  is-type level-1
  area-password cisco
```

配置完成后，我们发现 R1、R2 的邻接关系是 UP 的，也就是说区域认证，IIH 报文中并没有携带认证信息。不仅如此，R1 发出的 PSNP 也没有认证信息，但是发出的 LSP 中，是有明文的认证信息的。实验的现象还是挺有意思的，对于 R1 而言，收到 R2 的所有 IS-IS 消息都是没有认证信息的，R1 主要关注 LSP，发现 R2 发来的 LSP 没有用于认证的 TLV，因此 R1 忽略 R2 发来的 LSP，这就导致 R1 的路由表是空的；然而虽然 R1 发出来的 LSPs 都是有认证信息的，但是 R2 由于没有开启 area 认证，因此 R2 收到这个包，也可以读取，并且还能用这些 LSP，所以 R2 能学习到 R1 的路由。

这个时候，我们再做一下测试，R1 配置的 area-password 与 R2 配置的不一样。这样一来，R1、R2 虽然邻居关系能建立，但是都无法学习到对方的路由，因为 LSPs 双方都不认啊。

### 3. 明文 domain 认证



首先配置 R3：

```
Router isis
```

```
Domain-password cisco
```

完成如上配置后，我们发现 R2、R3 以及 R3、R4 的 is-is 邻居关系都是 UP 的。这是因为 IIH 包并没有携带用于验证的 TLV 信息。另外，R3 的路由表是空的，

**R2 的路由表如下（省略直连路由）：**

```
i L1    1.1.1.0 [115/20] via 10.1.12.1, FastEthernet0/0
i L2    3.3.3.0 [115/20] via 10.1.23.3, Serial1/0
i L2    10.1.34.0 [115/20] via 10.1.23.3, Serial1/0
```

**R3 的路由表如下（省略直连路由）：**

```
i L2    3.3.3.0 [115/20] via 10.1.34.3, Serial0/0
i L2    10.1.23.0 [115/20] via 10.1.34.3, Serial0/0
```

我们发现 R2 能学习到 R1 的 loopback、R3 的 loopback、以及 R3 的直连接口路由 34.0，这是因为 R4 传给 R3 以及 R2 传给 R3 的 LSP 都没有携带认证信息，那么这些 LSPs 到了 R3 之后，R3 是肯定报错的，因此 R3 不会将 R4 的发出的原始 LSP 泛洪给 R2 的，同时 R3 也是不会将 R2 发出来的原始 LSP 泛洪给 R4 的。但是，R3 自己发出来的 LSP，里头有 R3 的直连网段 3.3.3.0 及 10.1.23.0、10.1.34.0，这些 LSP 泛洪到了 R2 及 R4，虽然 LSPs 携带了用于认证的 TLV 信息，但是，R2 和 R4 并没有开启 domain 认证啊，因此他们直接无视这些认证信息，然后开始读 LSP，这就是我们实验的结果的解释。当然如果 R2 和 R4 配置了 domain-password 但是密码和其他设备的不一样，那就不用说了，直接丢弃密码不同的 LSP，压根是不会读

取的。

完成实验我们只要在 R2 和 R4 上配置 domain-password 即可，R1 是无需配置的（domain 认证实际上是在 level2 LSP 中嵌入用于认证的 TLV）。

使用 domain 认证，IIH 报文、CSNP、PSNP 是不携带认证信息的，LSP 则携带用于认证的 TLV。

#### 4. 密文接口认证



例如要在 R1、R2 之间进行接口密文验证：

**R1 的配置如下：**

```
R1(config)#key chain test
R1(config-keychain)#key 1
R1(config-keychain-key)#key-string cisco
R1(config-keychain-key)#exit
R1(config-keychain)#exit
!!!
```

```
R1(config)#interface fa0/0
```

```
R1(config-if)#isis authentication key-chain test level-1
```

```
R1(config-if)#isis authentication mode md5 level-1
```

R2 的配置类似，这样一来在 IIH 中用于认证的 TLV 中，装载的就是经过 hash 后的密文，不会暴露密码。

#### 5. 密文 LSPs 认证 ( Level1 )



**R1 的配置如下：**

```
R1(config)#key chain test
R1(config-keychain)#key 1
```



```
R1(config-keychain-key)#key-string cisco
```

```
R1(config-keychain-key)#exit
```

```
R1(config-keychain)#exit
```

```
!!!
```

```
R1(config)#router isis
```

```
R1(config-router)#authentication key-chain test level-1
```

```
R1(config-router)# authentication mode md5 level-1
```

这样一来，R1 发送出来的 Level1 LSP 都会携带密文认证的 TLV，然而由于目前为止 R2 并没有开启认证，因此 R1 发送的这些 LSP 到了 R2，R2 直接忽略 LSP 中的认证信息，直接去读取，因此 R2 能学习到 R1 通告的路由。当然 R1 啥路由也没有。

那么只要在 R2 上补全配置就行了。这里不再赘述。

## 6. 密文 LSPs 认证 (Level2)



**R2 的配置如下：**

```
R2(config)#key chain test
```

```
R2 (config-keychain)#key 1
```

```
R2 (config-keychain-key)#key-string cisco
```

```
R2 (config-keychain-key)#exit
```

```
R2 (config-keychain)#exit
```

```
!!!
```

```
R2 (config)#router isis
```

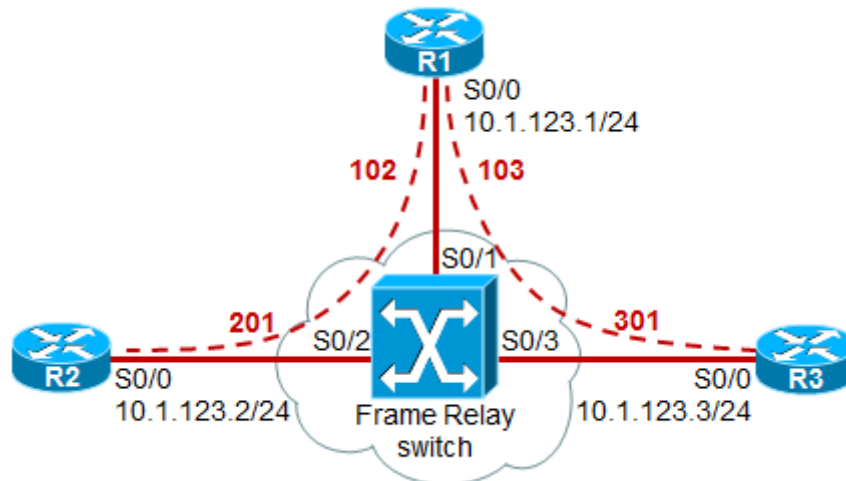
```
R2 (config-router)#authentication key-chain test level-2
```

```
R2 (config-router)#authentication mode md5 level-2
```

完成如上配置后，R2 发出来的 level2 LSP 就会携带上密文认证的 TLV 信息（level1 LSP 是不会携带的），如果保持这样的配置，我们会发现 R2 只能学习到 R1 发布出来的路由，这是因为 level1 LSP 不用认证。完成这个实验，R3、R4 补齐配置就成了。

## 5.12 NBMA 环境中的 IS-IS

我们知道 IS-IS 其实真正识别的网络类型就两个，一个是点对点，一个是广播多路访问网络。对于 NBMA，例如帧中继网络来说，IS-IS 的部署要谨慎，来看一下会有什么问题：



上面的网络是一个典型的 Hub&Spoke 网络，R1-R2、R1-R3 之间各维护一条 PVC，R2,R3 之间没有 PVC。

R1 的配置如下：

```
interface Loopback0
ip address 1.1.1.1 255.255.255.255
ip router isis
interface Serial0/0
ip address 10.1.123.1 255.255.255.0
ip router isis
encapsulation frame-relay
isis priority 100
frame-relay map ip 10.1.123.2 102 broadcast
frame-relay map ip 10.1.123.3 103 broadcast
frame-relay map clns 102 broadcast
frame-relay map clns 103 broadcast
no frame-relay inverse-arp
router isis
net 49.0001.0000.0000.0001.00
```

!! 让 R1 成为 DIS

!! 这是 IP 的映射

!! 务必不要忘了 CLNS 的映射

R2 的配置如下：

```
interface Loopback0
```

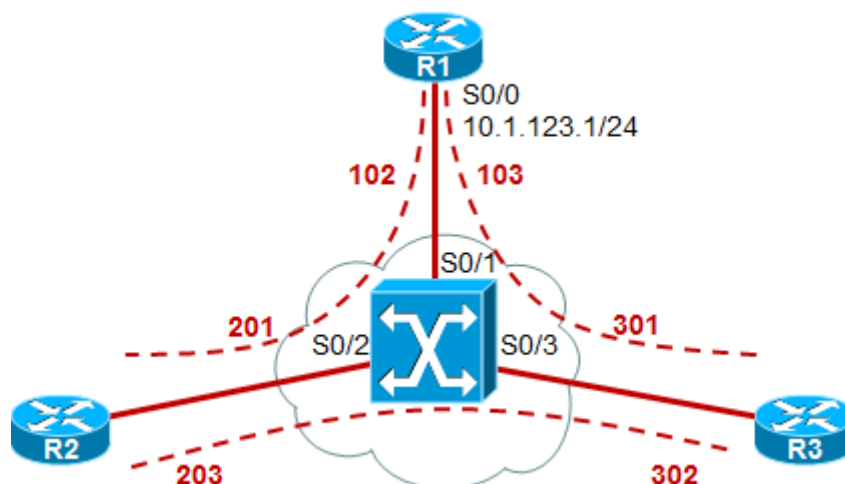
```
ip address 2.2.2.2 255.255.255.255
ip router isis
interface Serial0/0
ip address 10.1.123.2 255.255.255.0
ip router isis
encapsulation frame-relay
frame-relay map ip 10.1.123.1 201 broadcast
frame-relay map ip 10.1.123.2 201 broadcast
frame-relay map ip 10.1.123.3 201 broadcast
frame-relay map clns 201 broadcast
no frame-relay inverse-arp
!
router isis
net 49.0001.0000.0000.0002.00
is-type level-1
```

### R3 的配置如下：

```
interface Loopback0
ip address 3.3.3.3 255.255.255.255
ip router isis
interface Serial0/0
ip address 10.1.123.3 255.255.255.0
ip router isis
encapsulation frame-relay
frame-relay map ip 10.1.123.1 301 broadcast
frame-relay map ip 10.1.123.2 301 broadcast
frame-relay map ip 10.1.123.3 301 broadcast
frame-relay map clns 301 broadcast
no frame-relay inverse-arp
!
router isis
net 49.0001.0000.0000.0003.00
is-type level-1
```

在完成上述配置后，我们会发现，R1 能学习到 R2、R3 的路由，但是 R2、R3 之间却无法学习到对端的路由。

接下去，在 R2、R3 之间增加一条 PVC：



路由才算正常，三台路由器都能学习到全网的路由。

所以：

- **ISIS 对于 NBMA 网络的支持还是存在缺陷的，像前面实验演示的那样，如果是在帧中继主接口上跑 ISIS，或者在 P2MP 子接口上跑 ISIS，那么就要求 PVC 全互联，否则路由学习不正常，但是即使是 PVC 全互联，也存在隐患，如果某条 PVC 故障了，那么路由还是会出现问题。**
- **所以在 NBMA 网络中，譬如帧中继环境，强烈建议使用 P2P 子接口来跑 ISIS。**

## 6 IS-IS 对 IPv6 的支持

IETF 的 draft-ietf-isis-ipv6-05.txt 中规定了 IS-IS 为支持 IPv6 所新增的内容。主要是新添加的支持 IPv6 路由信息的两个 TLVs ( Type-Length-Values ) 和一个新的 NLPID ( Network Layer Protocol Identifier )。IS-IS 不像 RIP 和 OSPF，这两个协议有单独的版本 RIPng 和 OSPFv3 支持 IPv6。

**新增的两个 TLV 分别是：**

- IPv6 Reachability：类型值为 236 ( 0xEC )，通过定义路由信息前缀、度量值等信息来说明网络的可达性。
- IPv6 Interface Address：类型值为 232 ( 0xE8 )，它相当于 IPv4 中的 “IP Interface Address” TLV，只不过把原来的 32 比特的 IPv4 地址改为 128 比特的 IPv6 地址。

NLPID 是标识网络层协议报文的一个 8 比特字段，IPv6 的 NLPID 值为 142 ( 0x8E )。如果 IS-IS 路由器支持 IPv6，那么它必须以这个 NLPID 值向外发布路由信息。

## 7 IS-IS 与 OSPF

集成型 IS-IS 和 OSPF 都是 20 世纪 80 年代后期定义的, 大约是 1988 年。OSPF 实际上是从 IS-IS 的早期版本进化而来的, 不过它采用 IP 作为前提。OSPF 的核心概念, 例如链路状态信息的扩散、SPF 算法以及在广播链路中使用指定路由器都是从 IS-IS 早起版本中借用过来的。

OSPF version1 在 1989 年 10 月作为 RFC1131 发布。而 1990 年 12 月具有 IP 路由选择扩展功能的集成 IS-IS 协议作为 RFC1195 被发布。

CISCO 实现的集成 IS-IS 协议最早是在 1991/1992 这个时间段发布的。此后进行了一系列的改进和扩展。在主要的 ISP 网络中大规模开发 IS-IS 协议开始于 1995 年, 其中另一个重要的原因就是美国政府对 ISO CLNS 的兴趣。虽然集成 IS-IS 具有双重路由选择功能, 但是有趣的是当今大多数全球最大的 ISP 网络中使用 IS-IS 协议都只是为了实现 IP 路由选择。

由于 OSPF 天生被设计用来支持 IP 路由选择, 而 IS-IS 仍然有部分厂商的设备不能完美支持, 因此对于相当一部分的网络来说, OSPF 仍然是首选。然而将集成型 IS-IS 用于 IP 路由选择的兴趣还在增加, 欧洲的大多数 ISP 用的都是 IS-IS。许多新的 ISP 也在考虑部署 IS-IS。

当然, 在集成 IS-IS 及 OSPF 都在测试期间获得成功并且都已经获得了 ISP 网络内部网 IP 路由选择协议的参选资格。两种协议在过去一段时间都取得了一些新的扩展。例如对 MPLS TE 的支持、对 IPv6 的支持等等。另一个方面, 我们能够观察到的一个比较直观的现象是, 相比于 IS-IS, OSPF 的技术文档要多得多。

### IS-IS 与 OSPF 的术语比较:

IS-IS	OSPF	注释
终端系统 ES	主机	
中间系统 IS	路由器	
电路 circuit	链路 Link	
SNPA	数据链路层地址	
PDU	Packet	
DIS	DR	
-	BDR	
IIH	Hello	
LSP	LSA	
CSNP	DBD	
PSNP	LSR 或 LSAck	
路由 domain	AS	
Level2 area	Backbone area	

Level1 area	常规 area ( 非骨干区域 )	
虚链路	虚链路	两者存在区别
系统 ID	routerID	
LSP ID	链路状态 ID	

### 共同之处：

- 都是链路状态路由协议，都要求区域内的路由器交换链路状态信息，链路状态信息被收集到链路状态数据库中
- 都使用了一种实现路由选择信息交换的相似机制，即扩散
- 都在广播网络中选择指定路由器来控制扩散并降低这类介质中多对多邻接的系统资源需求
- 都是基于链路状态数据库中的信息，采用几乎相同的算法—SPF 算法来计算最佳路由
- 都支持两个 level 的分层路由选择
- 都支持 IP 前缀的无类路由选择（支持 VLSM）
- 都是公有协议（对比于 CISCO 私有的 EIGRP）

### 重要区别：

	IS-IS	OSPF
1	IS-IS 支持 ISO CLNP 和 IP 两种环境	仅支持 IP
2	IS-IS 报文封装在数据链路层帧中	OSPF 报文封装在 IP 包中，因此在网络层传输
3	IS-IS 支持 ISO 无连接网络环境，注意数据链路是 ISO 协议（在以太网上数据链路类型为 FEFE），在 ISO 协议栈中 IS-IS 网络层协议 ID 是 0x83	OSPF 封装在 IP 报文中，协议号为 89
4	IS-IS 路由器通告包含直连邻居及路由信息 TLV 的 LSP，使用 LSP 承载所有的路由选择信息	OSPF 使用不同类型的 LSA 承载不同的路由信息，LSA 被封装进 LSU 通告给邻居
5	IS-IS 数据包利用 TLV 字段承载所有易于扩展的信息	OSPF 只有 LSA 可扩展，而 LSA 扩展性太差
6	IS-IS 可以忽略它不支持的 TLV	网络中的路由器为了进行适当的操作必须识别所有 LSA
7	IS-IS 数据包可以承载多个 TLV，只有一个包头，节省带宽	1 类、2 类 LSA 可以承载多个 IP 前缀；3 类、4 类、5 类 LSA 只能承载单个 IP 前缀，如果需要发送多个 IP 前缀信息，需要多个 LSA
8	对于所有实际应用，IS-IS 仅支持广播和点对点链路。不支持 NBMA 链路。在 NBMA 环境下，可配置为 P2P 子接口或者广播链路（如果是全互联的	OSPF 支持如下网络类型：P2P、广播、NBMA、点到多点和按需电路

	连接方式 )	
9	仅仅在广播链路实现 3 步邻接关系, IETF 正在努力制定点到点链路的 3 步进程	OSPF 邻接关系的建立涉及到一个更加复杂的过程
10	最初数据库同步在邻接关系建立后进行	最初数据库同步在邻接关系形成之前进行
11	IS-IS 路由器只属于一个特定区域	OSPF 基于接口划分区域, 路由器可属于不同区域
12	区域的边界在链路上	区域的边界在路由器上
13	默认情况下 IS-IS 区域是 stub 区域, 最新发布的 RFC2966 标准化了从 Level2 到 Level1 的路由泄露	默认情况下, OSPF 区域不是 stub, 可以配置为 stub
14	IS-IS 仅支持在点对点链路上可靠扩散, 广播链路上的扩散是不可靠的。然而通过 DIS 周期性的广播实现可靠性	OSPF 确保所有链路上扩散的可靠性
15	DIS 无备份 DIS, DIS 可以被抢占, DIS 以 3 倍的频率发送 Hello PDU	有 BDR, DR 不能被抢占, DR 以正常的频率发送 Hello 报文
16	默认情况下, IS-IS 的 LSP 最大生存时间为 1200s, 刷新间隔为 900s, 而且定时器的值可调	OSPF 的 LSA 的老化时间为 3600s, 刷新间隔为 1800s, 而且是固定值
17	默认情况下, IS-IS 的接口 cost 值为 10	默认情况下, OSPF 的接口 cost 值根据带宽进行计算
18	默认情况下, IS-IS 保持时间( holding-time )为 30s, 而且在建立邻接关系时不需要双方的保持时间匹配	默认情况下, OSPF 的保持时间 ( dead-interval ) 为 40s, 而且为了建立邻接关系, 必须使双方的保持时间一致
19	IS-IS 通过将 Hello PDU 的大小填充至接口 MTU 大小来检查双方的 MTU 是否匹配	OSPF 通过在 DBD 报文中嵌入接口 MTU 字段来检查双方的 MTU 是否匹配
16	由于 IS-IS 区域中 IP 前缀是 SPF 树的叶子, 故部分路由计算 ( PRC ) 较多, 通常这就意味着在一个大的区域中路由处理器的负载较低	部分 SPF 被限制用于域间和外部路由, 任何要求较小区域和分层拓扑扩展引起的域间链路动荡导致完全 SPF 计算
17	没有对 Ip 组播路由选择的支持	MOSPF 扩展提供对 IP 组播路由选择的支持

## 8 参考书目

Cisco Press - IS-IS Network Design Solutions

IS-IS 网络设计解决方案

相关标准：

文档编号	描述
ISO 10589	ISO IS-IS Routing Protocol
ISO 9542	ES-IS Routing Protocol
ISO 8348/Ad2	Network Services Access Points
RFC 1195	Use of OSI IS-IS for Routing in TCP/IP and Dual Environments
RFC 2763	Dynamic Hostname Exchange Mechanism for IS-IS
RFC 2966	Domain-wide Prefix Distribution with Two-Level IS-IS
RFC 2973	Domain-wide Prefix Distribution with Two-Level IS-IS
RFC 3277	IS-IS Transient Blackhole Avoidance
RFC 3358	Optional Checksums in IS-IS
RFC 3373	Three-Way Handshake for IS-IS Point-to-Point Adjacencies
RFC 3567	Intermediate System to Intermediate System (IS-IS) Cryptographic Authentication
RFC 3719	Recommendations for Interoperable Networks using IS-IS
RFC 3786	Extending the Number of IS-IS LSP Fragments Beyond the 256 Limit
RFC 3787	Recommendations for Interoperable IP Networks using IS-IS
RFC 3784	IS-IS extensions for Traffic Engineering
RFC 3847	Restart signaling for IS-IS