



中华人民共和国通信行业标准

YD/T 2176-2010

公众 IP 网络可靠性 中间系统到中间系统路由交换协议 (IS-IS) 中平滑重新启动技术要求及测试方法

Reliability of public IP network – technical requirements and testing
methods for IS-IS graceful restart

2010-12-29 发布

2011-01-01 实施

中华人民共和国工业和信息化部 发布

目 次

前 言..... II

1 范围.....1

2 规范性引用文件.....1

3 术语、定义和缩略语.....1

4 概述.....2

5 平滑重启动机制.....2

 5.1 计时器.....2

 5.2 重启 TLV.....3

 5.3 重新建立邻接关系.....5

 5.4 数据库同步.....7

 5.5 LSP 生成、洪泛和 SPF 计算.....8

6 平滑重启动状态表.....9

 6.1 状态表.....9

 6.2 运行路由器.....9

 6.3 重启动路由器.....10

 6.4 启动路由器.....10

7 平滑重启动的安全问题.....11

8 IS-IS 平滑重启动测试.....11

 8.1 测试拓扑.....11

 8.2 IS-IS 平滑重启动功能测试.....12

 8.3 IS-IS 平滑重启动性能测试.....14

前 言

本标准是“公众IP网络可靠性”系列标准之一，该系列标准的预计结构及名称如下：

- 公众IP网络可靠性 总体技术要求
- YD/T 2175-2010 公众IP网络可靠性 标记分发协议（LDP）平滑重启动技术要求及测试方法
- YD/T 2176-2010 公众IP网络可靠性 中间系统到中间系统路由交换协议（IS-IS）中平滑重启动技术要求及测试方法

- 公众IP网络可靠性 RSVP-TE平滑重启动技术要求
- YD/T 1702-2007 公众IP网络可靠性 IP快速重路由技术框架
- 公众IP网络可靠性 双向转发检测（BFD）机制的技术要求
- 公众IP网络可靠性 虚拟路由冗余协议（Vrrp）技术要求

本标准由中国通信标准化协会提出并归口。

本标准起草单位：工业和信息化部电信研究院、上海贝尔股份有限公司。

本标准主要起草人：马军锋、顾方方。

广东省网络空间安全协会受控资料

公众 IP 网络可靠性
中间系统到中间系统路由交换协议（IS-IS）中平滑重启动
技术要求及测试方法

1 范围

本标准规定了公众IP网络可靠性中间系统—中间系统（IS-IS）平滑重启动机制及测试方法，包括重启动计时器、重启TLV报文格式、重启邻接关系建立过程、数据库同步以及LSP生成、洪泛和SPF计算。
本标准适用于所有支持中间系统到中间系统路由交换协议（IS-IS）的网络设备。

2 规范性引用文件

下列文件中的条款通过本标准的引用而成为本标准的条款。凡是注日期的引用文件，其随后所有的修改单（不包括勘误的内容）或修订版均不适用于本标准。然而，鼓励根据本标准达成协议的各方研究是否可使用这些文件的最新版本。凡是不注日期的引用文件，其最新版本适用于本标准。

- ISO/IEC 10589（2002） 用于无连接网络服务的中间系统到中间系统域内路由信息交换协议
- IETF RFC1195（1990） 适用于IP和双栈环境的IS-IS协议
- IETF RFC5303（2008） IS-IS点到点邻接三次握手

3 术语、定义和缩略语

3.1 术语和定义

下列术语和定义适用于本标准。

3.1.1

平滑重启动路由器 Graceful Restart Restarter
发生协议重启事件且具备平滑重启动能力的路由器。

3.1.2

平滑重启动辅助路由器 Graceful Restart Helper
和GR Restarter具有邻居关系，协助完成平滑重启动流程的路由器。

3.1.3

平滑重启动会话 Graceful Restart Session
在IS-IS邻居建立时进行GR能力的协商，一般把GR能力协商过程称为GR Session。协商的内容就是双方是否都具备GR能力。一旦GR能力协商通过，当协议重启时就可以进入GR流程。

3.2 缩略语

下列缩略语适用于本标准。

CSNP	Complete Sequence Numbers Protocol	完全序列号协议
DUT	Device Under Test	被测设备
FT	Fault Tolerant	容错
GR	Graceful Restart	平滑重启动

IGP	Interior Gateway Protocol	内部网关路由协议
IIH	IS-IS Hello	中间系统—中间系统 Hello 报文
IS-IS	Intermediate System-Intermediate System	中间系统—中间系统
LAN	Local Area Network	局域网
LSP	Link State Packet	链路状态报文
LSDB	Link State Database	链路状态数据库
RA	Restart Acknowledgement	重启确认
RR	Restart Request	重启请求
SA	Suppress adjacency	抑制邻接
SNP	Sequence Number Packet	序列号报文
SRM	Send Routing Message	发送路由消息
TLV	Type Length Value	类型、长度、值

4 概述

IS-IS路由协议是在IETF RFC1195和ISO/IEC 10589中定义的链路状态内部网关路由协议。通常，当一个运行IS-IS路由协议的路由器被重启时，由于重启路由器与其邻居之间的相关事件会导致路由的临时中断。

在重启路由器完成与其邻居数据库的同步之前，它会计算自己的路由，这样可能会导致它所计算的路由无法与域内的其他路由器计算的路由完全一致。

当重启路由器的邻居检测到重启事件，它会将邻居状态迁移到Down，并且重新与重启路由器建立邻接关系。邻居状态的振荡会引起邻居重新生成描述相关邻接关系的LSP，相应的会导致经过重启路由器的路由临时中断。

当与重启邻居重新建立邻接关系时，路由器将完成以下3个步骤：

- a) 重新生成描述自身的LSP，在整个域内触发SPF计算（或者是level2的域内）；
- b) 置位自身LSP数据库SRMflags标志；
- c) 对于点对点的链路，基于已建立的邻接关系传送一组CSNP摘要数据。

上述3个步骤，在路由器重启的过程中，第一步是最不希望发生的，而第二步为了确保LSP数据库的同步是必须的，第三步操作应尽量交互最少的LSP。如果确保交互是可靠的，那么应提供一种机制判定何时与邻居路由器的数据库实现同步。

本标准定义了一种用于重启路由器判定何时完成与其邻居数据库同步的机制，以及一种优化LSP的同步机制，采用这种机制将会使得由于路由器重启而造成的路由中断时间缩减到最短。

本标准中规定的重启动机制，在点到点链路上应遵循IETF RFC5303中定义的三次握手机制。

5 平滑重启动机制

5.1 计时器

在IS-IS路由平滑重启动机制中定义了3个计时器，T1、T2和T3。

T1计时器：每个接口应维护一个T1计时器实例，指示未确认的重启尝试时间，缺省值是3s。

T1计时器类似于IS-IS协议中的IIH定时器，每个接口都有一个，它定义了发送带RR标志位的Hello报文的重传时间。当设备重启时在每一个接口上都会创建T1定时器，并周期性地发送带RR标志位的Hello报文，直到在对应接口上收到带RA标志位的Hello确认报文以及全部CSNP报文后，该接口上的T1定时器才会被取消。当接口上没有邻居或者没有支持GR特性的邻居时，该接口就一直无法收到带RA标志位的Hello确认报文，因此T1计时器就会永远得不到撤销，为了避免这种情况发生，IS-IS GR限制了T1计时器最大超时次数，当T1定时器超时次数超过最大值时就自动撤销该计时器。

T2计时器：定义了设备重启后LSDB同步的最大等待时间。每个LSDB数据库都有一个计时器，例如对于Level1-2路由器来说，就需要有两个T2计时器，一个为Level-1 LSDB同步的最大等待时间，另外一个为Level-2 LSDB同步的最大等待时间。当某个Level的LSDB同步完成之后，相应Level的T2计时器就会被取消。如果T2计时器超时后LSDB还没有完成同步，则撤销T2计时器，GR失败。

T2计时器的缺省值是60s。

T3计时器：定义了设备重启过程的最大持续时间。整个系统维护一个T3计时器实例，其初始值为65535，但是在各个接口上收到的带RA标志位的Hello确认报文后，根据这些报文中的剩余时间域最小值重新设定T3计时器的超时时间。如果T3计时器超时后LSDB同步还没有结束，T3计时器被撤销，GR失败。

T3计时器仅用于重启路由器。

5.2 重启 TLV

5.2.1 报文格式

5.2.1.1 标志域

在IIH 协议数据报文中定义了一个新的TLV，类型为211。如果在IIH报文中出现该TLV，则表明发送者支持本标准中定义的功能。在TLV域中携带的标志用于在重启过程中传递信息。支持该功能的路由器，在其发送的所有IIH报文中必须包含该TLV。报文长度域根据TLV填写的具体情况而定，但必须在1到3+标识长度的范围内，值域的具体内容如图1所示。

No. of octets	
标志域	1
剩余时间	2
重启邻居标识	ID Length

图1 Restart TLV结构

标志域记录了一些必要的状态标志位，长度为1byte，具体格式如图2所示。

0	5	6	7
Reserved	SA	RA	RR

图2 Restart TLV中标志域字段结构

目前只使用了最后3位相应标志位。

RR：重启请求标志位，当RR位为1时表示该路由器刚刚发生重启；

RA：重启确认标志位，当RA位为1时表示是对重启路由器的应答；

当GR Restarter重启后，在每个接口上发的第一个Hello报文中将RR标志位置1，通知GR Helper本设备正在重启；当GR Helper收到GR Restarter发来带有RR标志位的Hello报文时，立即回复一个将RA标志位置1的Hello报文作为对对方请求的确认。

SA: 抑制邻接标志位, 是一个可选项, 其主要目的是为了避免出现路由黑洞。

在启动的情况下, 如果GR Helper将报文发送到本设备来进行转发将是一个黑洞, 会造成严重的丢包现象, 在这种情况下GR Restarter发送的Hello报文中必须将SA位置1, 而GR Helper接收到这种SA位被置1的Hello报文后就不会将发送该Hello报文的GR Restarter放入LSP扩散出去, 即GR Restarter将在网络上屏蔽一段时间, 所有设备都不会将报文送到GR Restarter上来进行转发, 这样就可以有效避免路由黑洞问题。

5.2.1.2 剩余时间

当RA比特置位时, 剩余时间域是要求的, 其长度为2byte。剩余时间标识邻居老化的剩余时间(s), 即邻居设备进入GR Helper处理流程的最长保持时间。如果超出这个时间, 则GR Restarter和GR Helper的邻居关系结束。当GR Helper收到GR Restarter发送的带有RR标志位的Hello报文后, 会立即回复一个将RA标志位置1的Hello报文作为确认, 在这个确认报文中, 需要将对应邻居(GR Restarter)离老化时间的剩余时间填入剩余时间字段。

5.2.1.3 重启邻居标识

重启邻居系统标识的内容是GR Restarter的系统标识。当GR Helper收到带有RR标志位的Hello报文时, 将报文中的系统标识字段值填入Hello确认报文的重启邻居标识字段中, 并发送此确认报文。这样就严格指定了该Hello确认报文的目的地。当多个GR Restarter同时收到该Hello确认报文时, 比较该报文中的重启邻居标识与本设备的系统标识是否相同, 就可以判断是否需要处理该确认Hello报文。

5.2.2 RR 和 RA 比特位的使用

重启(启动)路由器使用标志域中的RR比特向其邻居通告其正在重启(启动)的过程, 即使在邻居状态机要求重新初始化邻接关系, 请求一组CSNP以及设置SRMflags标志的情况下, 也应当维护已经建立的邻接关系。

重启(启动)路由器的邻居通过发送RA比特置位的消息来确认接收到一个RR比特置位的重启TLV。

当重启路由器邻居接收到一个包含RR比特置位的重启TLV的IIH报文时, 如果在该接口上已经存在一个具有相同系统标识且状态为UP的邻接关系, 那么在LAN电路的情况下, 具有相同源LAN地址, 不考虑中间系统邻居选项的其他内容(LAN电路), 或者是“点到点三方邻接关系选项”(点到点链路), 应按照如下方式进行处理:

a) 邻接关系状态不发生变化。如果这是从具有邻接关系的邻居接收到的第一个RR比特置位的IIH报文, 那么邻接关系被标志为重启模式, 并且刷新邻接关系保持计时器, 否则保持计时器不刷新。剩余传送时间根据b) 必须反应邻接关系超时的实际时间。当收到一个普通的RR比特重新置位的IIH报文, 将会清除重启模式状态。该过程允许重启路由器使得其邻居维持邻接关系足够长的时间以完成重启, 但是也要避免重复重启不确定地维护一个邻接关系。邻接关系状态是否标记为重启模式对于邻接关系状态的迁移是没有影响的。

b) 通过相应的接口立即发送一个包含重启TLV并且RR比特清0, RA比特置位的IIH报文。将点到点邻接关系更新到“点到点三方邻接关系”选项的情况来体现任何从重启路由器新接收到的值(这允许重启路由器快速得到正确的信息并将其放到hello报文中)。

在邻接关系保持计时器超时之前, 剩余时间必须被设置成当前时间(单位为s)。如果是LAN接口, 那么重启邻居系统标识应当被置为接收到RR比特置位的IIH报文的路由器标识。这就要求在LAN拓扑下

当多个系统同时重启时，能够正确地关联确认和保持计时器。IIH报文应当在任何LSP或者是SNP报文被转送之前传送，作为接收初始IIH报文的结果。

c) 如果相应的接口是点到点接口，或者如果接收路由器在该接口上的所有与其邻接状态为UP的路由器中具有最高的LnRouterPriority优先级（具有最高的源MAC地址），而且IIH报文中包含重启TLV，排除与那些认为是在重启模式路由器的邻接关系（需要指出的是在这个过程中实际的DIS并未发生改变），通过相应的接口开始发送一组CSNP报文，在本地LSP数据库中将所有与该接口相关的LSP置SRMflags标志位。否则（例如，与正在讨论的系统标识不存在邻接关系为UP状态），按照一般的方式处理IIH报文，并且重新初始化邻接关系，在返回的IIH报文中RA比特置位。

5.2.3 SA 比特位的使用

启动路由器使用SA比特请求它的邻居不要在邻居LSP中通告它们之间的邻接关系。

如果路由器不是第一次启动，那么启动路由器时先前已生成的LSP可能会在网络中其他路由器的LSDB数据库中已经存在。这些拷贝可能会让启动路由器初始生成的LSP更新，因为启动路由器LSP分片序列号重新初始化。这可能会导致瞬间的路由黑洞，直到正常的路由更新过程引起启动路由器重新生成和扩散具有更高序列号的LSP。如果启动路由器的邻居抑制向其通告邻接关系，直到启动路由器能够扩散最新版本LSP，那么可以避免瞬间的路由黑洞。

当路由器接收到包含重启TLV且SA比特置位的IIH报文时，如果在接收接口上已经存在一个状态为UP的邻接关系，并且邻居系统标识与报文中的相同，则在LAN电路的情况下，如果源LAN地址相同，那么路由器必须在自己生成的LSP中抑制通告邻居的邻接关系。邻居通告必须一直抑制，直到接收到一个SA比特清0的IIH报文。如果邻接关系状态迁移到UP，那么在接收到一个新的SA比特清0的IIH报文之前一定不能通告新的邻接关系。

需要注意的是向邻居抑制通告邻接关系的路由器在执行SPF计算时，一定不能使用与该邻接相关的LSP。

5.3 重新建立邻接关系

5.3.1 在重启过程中重新建立邻接关系

重新建立邻接关系的第一步就是要重新初始化状态。重启路由器将会在启动过程中使用重启TLV中的RR比特来抑制重新初始化状态。

在重启的过程中，重启路由器会明确地通告其邻居正在重新建立邻接关系，因此不应当重新初始化邻接关系。重启路由器会通过将重启TLV中的RR比特置位来抑制重新初始化邻接关系。当重启路由器的邻居接收到包含重启TLV且RR比特置位的IIH报文时，如果在接收接口上存在具有相同系统标识，且邻接关系状态为up，那么在LAN电路的情况下，源LAN地址相同，就按照5.2.2规定的过程处理。

如果路由器不支持重启功能，那么将忽略重启TLV并且按照一般过程重新初始化邻接关系，返回一个不包含重启TLV的IIH报文。

在重启时，路由器初始化T3计时器，同时为每个LSDB数据库启动T2计时器，为每个接口启动T1计时器，并且发送包含TLV且RR比特置位的IIH报文。

在点到点电路情况下，重启路由器应当设置邻接关系3次握手状态为“Init”，因为一旦接收到确认的IIH报文（RA比特置位）必然会立即引起邻接关系进入“Up”状态。

在LAN电路情况下，为电路分配的LAN标识应当与重启之前使用的相同。特别是重启路由器作为指派路由器的电路，使用不同的LAN标识必然要生成一组新的伪LSP，在LAN上所有其他的路由器都会参考伪节点LSP，相应的改变所有的LSP。通过在重启的过程中保持LAN标识，就可以避免这种不必要的扰动。为了使重启路由器学习在重启之前使用的LAN标识，在RR比特置位的IIH报文中指定的LAN标识必须被忽略。

应禁止传送普通的IIH报文，直到出现下面的情况（为了避免引起不必要的邻接关系重新初始化）。一旦T1计时器超时，应重新启动继续发送IIH报文。

当一个重启路由器接收到一个IIH时，按照一般过程建立邻接关系，如果IIH包含重启TLV，且RA比特置位（在LAN电路情况下，重启邻居系统ID与本地系统ID相匹配），通过该接口接收到确认应当被记录。当RA比特置位，并且远端邻接状态是UP，那么T3计时器设置成当前值和IIH报文中剩余时间域值的最小值。

对于点到点链路，接收到不包含重启TLV的IIH报文也认为是一个确认，因为它说明邻居不具有重启能力。但是，因为不能保证在该接口上能够接收到CSNP，所以T1计时器会立即取消而不用等待一组完整的CSNP。因此，即使有一些LSP仅保留在它的邻居数据库中，也可能会认为同步已经完成。在这种情况下，希望能够确定邻居将会重新初始化邻接关系，确保它的数据库中SRMflags被置位，从而保证最终的LSP数据库同步。这种情况必然会发生，除了下述这种情况，即在收到IIH报文时，邻接3次握手状态是UP，并且邻居扩展本地电路标识与重启路由器分配的扩展本地电路标识相匹配。在这种情况下，重启路由器必须强制邻接关系重新初始化，通过将本地邻接3次握手状态置为Down，并且发送普通的IIH报文。

在LAN接口的情况下，接收到不包含重启TLV的IIH报文是不寻常的，因为只要在LAN中至少存在一个非重启的邻居路由器支持重启功能，那么同步仍旧可能会发生。因此在这种情况下，T1计时器继续运行。如果在LAN中没有邻居具有重启能力，T1计时器在指定多次尝试后将会最终超时。

在点到点电路情况下，IIH报文中包含的LocalCircuitID和Extended Local Circuit ID信息能够被用于生成一个包含正确3次握手信息的IIH报文。如果出现Neighbor Extended Local Circuit ID信息与当前本地系统正在使用的值不匹配，应当忽略（因为IIH可能在邻居从重启路由器接收到新的值之前就已经被发送了）。但是邻接关系仍旧在初始化状态，直到接收到正确的信息。

在LAN电路的情况下，为了正常建立和维护邻接关系，应当记录源邻居的信息。

当一组完整的CSNP和确认消息在接口上接收到，那么T1计时器被取消。

一旦T1计时器被取消，后续的IIH根据正常的算法被发送，但是包含RR和RA比特位清0的重启TLV。如果LAN包含一个混合系统，其中只有一部分支持新的算法，那么数据库的同步仍旧能够保证，但是老的系统将重新初始化它们的邻接关系。

如果一个接口被激活，但是没有任何邻居路由器通过该接口可达，不会取消T1计时器，也不会执行SPF算法。但是在预先确定的超时周期之后，T1计时器会被取消。

5.3.2 在启动过程中建立邻接关系

如果邻居路由器与启动路由器有邻接关系，并且状态为UP，这个邻接关系应当重新初始化。启动路由器也希望邻居路由器抑制向启动路由器通告邻接关系，直到LSP数据库实现同步。通过发送包含重启TLV且RR比特清0，SA比特置位的IIH报文来实现数据库的同步。

RR比特仍旧保持清0, SA比特仍旧保持置位在后续发送的IIH报文中, 直到邻接关系达到UP状态, 初始T1计时器内部超时。

收到RR比特清0的IIH报文将会导致邻居路由器使用正常的邻接关系状态机进行操作。这将保证邻居路由器上所有旧的邻接关系重新初始化。

当接收到SA比特置位的IIH报文, 应当按照5.2.3节中规定的过程进行处理。

一旦启动, 路由器为每个LSDB数据库启动T2计时器。

针对每个接口(对每个level, 在LAN电路的情况下), 当邻接关系状态达到UP时, 启动路由器启动T1计时器, 同时发送包含重启TLV的IIH报文, 其中RR比特清0, SA比特置位。一旦T1计时器超时, 路由器被重启, IIH报文重新被发送, RR、SA比特均置位(仅仅RR比特发生状态改变)。

一旦接收到RR比特置位的IIH报文(无论SA是否置位), 都按照5.2.2节规定的过程进行处理。

当启动路由器接收到IIH报文, 并且IIH报文包含一个重启TLV, RA比特置位(在LAN电路重启邻居系统标识与本地系统标识匹配), 通过该接口收到确认报文应记录。

对于点到点链路, 接收到不包含重启TLV的IIH报文, 也同样作为一个确认, 因为它指示邻居不具有重启的能力。因为邻居将会重新初始化邻接关系, 这确保它的数据库SRMflags标识置位, 最终确保LSP数据库同步。然而, 因为无法保证在该接口上接收到CSNP报文, T1计时器立即被取消无需等待一组完整的CSNP。因此同步可能会被认为已经完成, 即使有些LSP还仅仅由邻居保存。

在LAN接口的情况下, 接收到不包含重启TLV的IIH报文是正常的, 因为同步可能仍旧在发生, 只要至少在LAN存在一个非重启的邻居路由器支持重启。因此, T1计时器需要在这种情况下继续运行。如果在LAN上没有邻居具有重启的能力, T1计时器将会超时。在本地定义的多次尝试之后, 数据库更新过程能够最终确保数据库实现同步。

当在接口上接收到一组完整的CSNP(对每个活跃的level, 在点到点链路的情况下)和一个确认, T1计时器被取消。后继的IIH报文由启动路由器发送, 包含重启TLV, 并且RR和RA比特清0, SA比特置位。

T1计时器取消, 在一些预先定义的超时次数以后(可能是1)。

当T2计时器超时或者取消, 将开始发送正常的IIH报文(RR、RA和SA比特均清0)。

5.3.3 多级域

当路由器在特定接口上同时运行Level1和Level2的路由进程, 必须在每一级上按照上面的过程来运行。

对于LAN接口, 它必须发送和接收Level1和level2的IIH, 在每级上执行完全独立的CSNP同步过程。

对于点到点接口, 只要求单一的IIH报文(指示支持两个level), 但是它必须在每级上单独执行CSNP同步。

5.4 数据库同步

5.4.1 数据库同步过程

当路由器启动/重启时, 它可能期望在每个接口上接收一组CSNP。因为一个RR比特置位的IIH报文在启动/重启的过程中会重传, 直到CSNP报文能够正确地接收, 所以CSNP的到达是有保证的。

CSNP规定了一组当前每个邻居所拥有的LSP。当接收到所有这些LSP时, 数据库的同步过程就会完成。

当启动/重启时，路由器为每个LSDB启动一个T2计时器实例。除了正常的CSNP过程，记录每个接口上接收到的第一个完全CSNP中的LSPID以及剩余时间。在LAN接口的情况下，一组完整的CSNP必须由从未重启的邻居接收的CSNP构成。如果在启动/重启的路由器上有多个接口，记录的LSPID是通过每个接口接收到的LSPID的集合。如果LSP的剩余时间为0将不被记录。

当通过任一接口收到LSP，相应的LSPID表项被删除（如果LSP在包含相关参考的CSNP前到达，也应被删除）。当LSPID被保存在剩余周期列表中时，也应该从该列表中删除。当LSPID的列表为空，并且在该级上有邻接关系的所有接口上的T1计时器被取消时，该接口上的T2计时器也被取消。

本地数据要保证包含所有的LSP（要么是相同的序列号，或者是最近的序列号），在启动/重启邻居数据库中出现。邻居数据库到达的LSP，在启动/重启之后可能出现，也可能没有出现，但是正常的更新操作过程将会保证它们最终被接收。

因为在CSNP中剩余时间为0的LSP不被记录，当生命周期超时，那些具有较短剩余周期的LSP将会从列表中删除，T2计时器取消也不会被阻止，因为是在等待一个从未到达的LSP。

5.5 LSP 生成、洪泛和 SPF 计算

5.5.1 重启动过程

为了避免造成其他路由器不必要的路由搅动，由重启系统生成的路由器自身的LSP保持与网络中先前通告的相同是完全必要的（假定其他的变化已经发生）。因此不要在所有的邻接关系再次建立之前产生和扩散LSP是非常重要的。理想的是，这类信息通过一种确定性的方式加载到LSP中，就好像这些信息发生在相同的地方相同的LSP（因此LSP与它们先前的版本一致）。如果能够做到这一点，那么新的版本不会在其他的系统中引起SPF计算。然而，只要相同的信息包含在一组LSP中，那么SPF运算的结果是相同的，不会引起转发表的搅动。

在重启路由器的情况下，路由器自己的LSP没有一条会被转发，当T3计时器在运行时，路由器自身的转发表也不会被更新。

在路由器的LSP向其他节点扩散前，必须重新生成Inter-level的重分布信息。因此在其他level的T2计时器超时及SPF运行之前，Level-n的non-pseudonode LSP一定不能被扩散。这将确保任何即将传播的inter-level信息能够包含在level-n的LSP中。

在LSP扩散期间，如果任一路由器自身的LSP（包括伪节点）被收到，并且其不包含在本地路由器的数据库中，那么它将被删除。在正常操作中，此类LSP会被删除，因为其显然不会出现在全局LSP数据库中。然而，在目前的情况下，这是不期望的，因为这会引起提前删除路由器自己的LSP，从而导致远端路由器搅动。即使本地系统有一个或者多个自身的LSP（已经生成，但尚未转发），但不能和接收到的LSP进行比较，因为它可能是在Level 1和Level2之间传播的结果。此外，当LSP数据库实现同步，需要重新生成一个路由器自身的LSP。

在路由器重启的过程中，应当按照正常的模式发送CSNP。包含描述路由器自身LSP的信息，但是必须基于接收到的LSP，而不能按照已经产生尚未转发的版本。为了防止提前从全局LSP数据库中删除一个LSP，这个限制是必须的。

当T2计时器超时，或者取消，表明Level同步已经完成，SPF计算正在运行为了生成需要转发到其他level的信息，但是转发表不会被更新。

一旦其他level的SPF计算已经运行，同时inter-level转播已经接收到，路由器自身的LSP可能被生成并且扩散。任何先前忽略的自身LSP，但它不是当前自身LSP的一部分（包括伪节点）必须被删除。需要指出的是，指派路由器的改变可能已经发生是可能的，因此路由器应当删除那些伪节点LSP先前所拥有的，但是目前不再是伪节点LSP的一部分。

当所有的T2计时器超时，或者是被取消时，T3计时器被取消，而且本地的转发表被更新。

如果在所有的T2计时器超时或者是被取消之前T3计时器超时，那么表明同步过程将花费比邻居 最小保持时间更长的时间。路由器自身的尚未完成第一次SPF计算的LSP将被洪泛，其中超载比特位置位，表明路由器的LSDB尚未同步（因此其他的路由器一定不能计算通过该路由器的路由）。更新过程的一般操作重新开始，本地转发表被更新。为避免邻居的邻接关系超时，包含保持时间普通接口值的IIH在所有接口转发，在重启TLV中邻居RR或者是RA均不置位。这会引起邻居刷新它们的邻接关系。路由器自身的LSP会继续将超载比特位置位直到T2计时器超时或者被取消。

5.5.2 启动过程

在启动路由器的情况下，一旦建立邻接关系，在交换CSNP之前，路由器自身第0个LSP被转发，并且超载比特位置位。这将避免其他的路由器计算通过该路由器的路由，直到可靠获得一组完整的LSP。当任何T2计时器运行时，超载比特保持置位在后续传送的zeroth LSP（这种情况将会发生，如果一个先前的路由器自身zeroth LSP的拷贝仍旧存在）。

当所有T2计时器已经被取消，路由器自身LSP（s）可能会重新生成，超载比特清零并且按照正常模式洪泛（假定路由器实际上没有过载，也没有其他别的原因使得过载比特位置位，例如不完全的BGP路由汇聚）。

路由器所拥有的其他LSP（包括伪节点）按照一般模式生成并且洪泛，与计时器T2无关。SPF按照一般模式进行运算，当路由可用时更新RIB和FIB。

为避免信息可能会出现临时黑洞，正在启动的路由器将它所发送的IIH消息中重启TLV的SA比特位置位。当所有的T2计时器已经被取消，正在启动的路由器必须发送SA比特清零的IIH报文。

6 平滑重启动状态表

6.1 状态表

在ISIS平滑重启动的过程中，运行路由器、重启动路由器和启动路由器的状态迁移应分别按照表1、表2和表3的规定。但是上述状态表仅总结了IS-IS GR重启动过程中路由器的行为，不包含路由器邻接关系的状态迁移以及LSDB数据库的同步过程。

6.2 运行路由器

运行路由器的状态迁移应遵循表1的规定。

表1 运行路由器状态迁移

事 件	状 态	
	Running	ADJ suppressed
接收到RR比特置位的Hello报文	保持邻接状态关系； 发送RA比特置位的Hello报文； SRMflags置位，发送CSNP； 更新保持计时器，设置Restart模式	

表1（续）

事 件	状 态	
	Running	ADJ suppressed
接收到RR比特清零的Hello报文	清除Restart模式	
接收到SA比特置位的Hello报文	在发送的LSP报文中抑制通告重启的邻居，并跳转到邻接关系抑制状态（ADJ suppressed）	
接收到SA比特清零的Hello报文		在发送的LSP报文中不抑制通告重启的邻居，并迁移到运行（Running）状态

6.3 重启动路由器

重启路由器的状态迁移应遵循表2的规定。

表2 重启动路由器状态迁移表

事 件	状 态			
	Restarting	ADJ Seen RA	ADJ Seen CSNP	SPF Wait
路由器重启	发送RR比特置位的IIH报文，邻接关系ADJ迁移到Init状态，启动T1、T2、T3计时器			
接收到RR比特置位的Hello报文	发送RA比特置位的Hello报文			
接收到RA比特置位的Hello报文	调整T3计时器，状态迁移到ADJ Seen RA		取消T1计时器，调整T3计时器	
接收到CSNP集	状态迁移到ADJ Seen CSNP	取消T1计时器		
接收到Hello报文（w/o）重启TLV	取消T1计时器（仅点到点链路）			
T1计时器超时	发送RR比特置位的IIH报文，重启T1计时器	发送RR比特置位的IIH报文，重启T1计时器	发送RR比特置位的IIH报文，重启T1计时器	
T1计时器第n次超时	发送普通的IIH报文	发送普通的IIH报文	发送普通的IIH报文	
T2计时器超时	触发SPF计算，状态迁移到SPF Wait			
T3计时器超时	设置0L，洪泛本地的LSP，更新转发平面数据			
LSP数据库同步	取消T2、T3计时器，触发SPF计算			
所有的SPF计算完成				Clear 0L，更新转发平面，洪泛本地LSP，迁移到运行状态

6.4 启动路由器

启动路由器的状态迁移应遵循表3的规定。

表3 启动路由器状态迁移表

事 件	状 态		
	Starting	ADJ Seen RA	ADJ Seen CSNP
路由器启动	发送SA比特置位的IIH报文，启动T1、T2计时器		
接收到RR比特置位的Hello报文	发送RA报文		
接收到RA比特置位的Hello报文	迁移到ADJ Seen RA状态		取消T1计时器
接收到CSNP集	迁移到ADJ Seen CSNP状态	取消T1计时器	
接收到Hello报文 w 不包含重启TLV	取消T1计时器（仅点到点链路）		
ADJ UP	启动T1计时器，发送本地LSP, w 0L		
T1计时器超时	发送RR、SA比特置位的IIH报文，重启T1计时器	发送RR、SA比特置位的IIH报文，重启T1计时器	发送RR、SA比特置位的IIH报文，重启T1计时器
T1计时器第n次超时	发送SA比特置位的IIH报文	发送SA比特置位的IIH报文	发送SA比特置位的IIH报文
T2计时器超时	Clear 0L，发送普通的IIH报文，迁移到运行状态		
LSP数据库同步	取消T2计时器，Clear 0L，发送普通的IIH报文		

7 平滑重启的安全问题

注入有错误但是有效的IIH。

如果在一个错误的IIH报文中RR比特置位，接收到该报文的邻居将继续维持一个已经存在的状态为UP的邻接关系，并且可能（重新）发送一组完整的CSNP。然而后者的动作是浪费的，任何动作在正确的协议操作中都会引起中断。

如果在一个错误的IIH报文中RA比特置位，启动或者是重启的路由器接收到该报文可能会错误的，相信在相应的接口上存在一个邻居，并且支持本标准规定的平滑重启机制。在这个接口上没有接收到一组完整的CSNP,通过要求T1计时器超时达到本地最大的尝试次数，可能会延迟启动或者重启过程的完成。

如果在一个错误的IIH报文中SA比特置位，这可能会引起一个IS邻居通告抑制。由于到网络中一些目的地的可达性丢失，可能会增加LSP洪泛和SPF计算的频度。

8 IS-IS 平滑重启测试

8.1 测试拓扑

测试仪表与被测设备以及辅助交换机如图3所示连接，分别在测试仪表端口和被测试设备上配置IS-IS路由协议，并且建立IS-IS level2的邻接关系；通过在辅助交换机上配置端口镜像功能，由协议分析仪抓包分析被测设备之间交互的IS-IS协议报文。

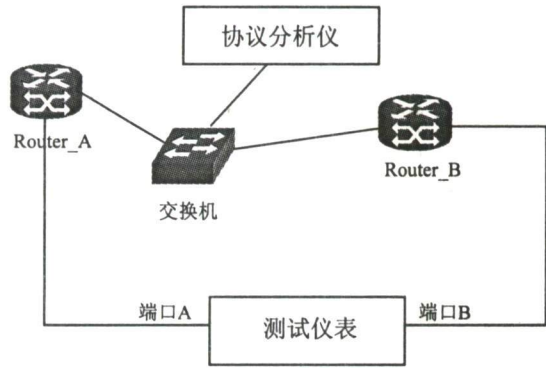


图3 测试拓扑

8.2 IS-IS 平滑重启功能测试

测试编号：1
测试项目：GR 能力协商
测试目的：验证被测系统能够与不具备 GR 能力的设备建立普通的 IS-IS 邻接关系
测试配置：如图 3
测试过程： 1) 按测试环境连接设备； 2) 被测设备 Router_A、Router_B 和测试仪表端口 A、B 分别配置 IP 地址，并且启用 IS-IS 路由协议； 3) 在被测设备 Router_A 和测试仪表端口 A 上分别配置 IS-IS GR 功能； 4) 由协议分析仪抓取 Router_A 和 Router_B 交互的控制报文
预期结果： 在步骤 3)，Router_A 与 Router_B 之间建立普通的 IS-IS 邻接关系； 在 GR 协商阶段，能够抓到 Router_A 发送的类型为 211 的 IIH 报文
判定原则： 应符合预期结果要求，否则不合格

测试编号：2
测试项目：GR 能力协商
测试目的：验证被测系统能够与具备 GR 能力的设备协商 GR 能力，并建立支持 GR 的 IS-IS 邻接关系
测试配置：如图 3
测试过程： 1) 按测试环境连接设备； 2) 被测设备 Router_A、Router_B 和测试仪表端口 A、B 分别配置 IP 地址，并且启用 IS-IS 路由协议； 3) 在被测设备 Router_A、Router_B 和测试仪表端口 A、B 上分别配置 IS-IS GR 功能； 4) 由协议分析仪抓取 Router_A 和 Router_B 交互的控制报文； 5) 由测试仪表分别从端口 A 和 B 通告 IS-IS 路由，并基于通告的路由构造双向测试流量； 6) 切换 Router_A 的路由控制引擎
预期结果： 在步骤 3)，Router_A 与 Router_B 之间建立具有 GR 能力的 IS-IS 邻接关系； 在 GR 协商阶段，能够抓到被测设备交互的类型为 211 的 IIH 报文； 在步骤 6)，切换 Router_A 的路由控制引擎不影响测试流量的转发（没有丢包）
判定原则： 应符合预期结果要求，否则不合格

测试编号：3
测试项目：T1 计时器
测试目的：验证被测系统能够正确处理 T1 计时器
测试配置：如图 3
测试过程： 1) 按测试环境连接设备； 2) 被测设备 Router_A、Router_B 和测试仪表端口 A、B 分别配置 IP 地址，并且启用 IS-IS 路由协议； 3) 在被测设备 Router_A 和测试仪表端口 A 上分别配置 IS-IS GR 功能； 4) 由协议分析仪表抓取 Router_A 和 Router_B 交互的控制报文
预期结果： 1) 通过命令查看被测设备 Router_A 上 T1 计时器的参数（发送周期，最大超时次数）； 2) 在步骤 4)，RR 置位的 IIH 控制报文的发送周期应与 Router_A 设备上 T1 计时器的发送周期相同，并且当 T1 计时器的超时次数大于最大超时次数时，T1 计时器撤销
判定原则： 应符合预期结果要求，否则不合格

测试编号：4
测试项目：T2 计时器
测试目的：验证被测系统能够正确处理 T2 计时器
测试配置：如图 3
测试过程： 1) 按测试环境连接设备； 2) 被测设备 Router_A、Router_B 和测试仪表端口 A、B 分别配置 IP 地址，并且启用 IS-IS 路由协议； 3) 在被测设备 Router_A、Router_B 和测试仪表端口 A、B 上分别配置 IS-IS GR 功能； 4) 在 Router_B 上配置 T2 计时器的超时间隔小于 Router_A 路由器的重启恢复时间； 5) 重启被测设备 Router_A
预期结果： 在步骤 5)，在被测设备 Router_B 上应可以看到 T2 计时器超时信息，并且 GR 失败
判定原则： 应符合预期结果要求，否则不合格

测试编号：5
测试项目：T3 计时器
测试目的：验证被测系统能够正确处理 T3 计时器
测试配置：如图 3
测试过程： 1) 按测试环境连接设备； 2) 被测设备 Router_A、Router_B 和测试仪表端口 A、B 分别配置 IP 地址，并且启用 IS-IS 路由协议； 3) 在被测设备 Router_A、Router_B 和测试仪表端口 A、B 上分别配置 IS-IS GR 功能； 4) 分别配置 Router_A、Router_B 和仪表接口 A、B 所仿真的路由设备重启过程的最大持续时间； 5) 重启被测设备 Router_A； 6) 在 Router_A 设备上查看 T3 计时器的值
预期结果： 在步骤 6)查看到的 T3 计时器的值,应为从各个接口上收到带 RA 标志位的 Hello 确认报文中 Remaining Time 域值的最小值
判定原则： 应符合预期结果要求，否则不合格

8.3 IS-IS 平滑重启动性能测试

测试编号：6
测试项目：IS-IS 平滑重启动性能
测试目的：验证被测系统 IS-IS 平滑重启动性能
测试配置：如图 3
测试过程： 1) 按测试环境连接设备； 2) 被测设备 Router_A、Router_B 和测试仪表端口 A、B 分别配置 IP 地址，并且启用 IS-IS 路由协议； 3) 在被测设备 Router_A、Router_B 和测试仪表端口 A、B 上分别配置 IS-IS GR 功能； 4) 由测试仪表从端口 A 和 B 分别通告 5000 条 IS-IS 路由，并基于通告的路由构造双向测试流量，发送速率为链路带宽的 80%； 5) 切换 Router_A 的路由控制引擎
预期结果： 步骤 5)，被测设备路由控制引擎正常切换；测试流量转发不受影响（无丢包）
判定原则： 应符合预期结果要求，否则不合格

广东省网络空间安全协会受控资料

中 华 人 民 共 和 国
通 信 行 业 标 准

公 众 IP 网 络 可 靠 性

中间系统到中间系统路由交换协议（IS-IS）中平滑重启技术要求及测试方法

YD/T 2176-2010

*

人民邮电出版社出版发行

北京市崇文区夕照寺街 14 号 A 座

邮政编码：100061

北京新瑞铭印刷有限公司印刷

*

开本：880×1230 1/16

2011 年 2 月第 1 版

印张：1.25

2011 年 2 月北京第 1 次印刷

字数：32 千字

ISBN 978 - 7 - 115 - 2195/ 11 - 146

定价：15 元