了解IS-IS SRv6微环路规避行为

目录

简介

先决条件

- 要求
- 使用的组件

Microloop概述

通用配置

处于稳定状态

PE1上的接口TenGigE0/0/0/8（主路径）关闭时的事件顺序

时间表摘要

结论

命令

简介

本文档从技术角度概述了IS-IS微环路、产生微环路的条件以及用于防止微环路的机制。

先决条件

要求

思科建议您了解中间系统到中间系统(ISIS)网段路由(SR)版本6的基本知识。

使用的组件

本文档中的信息基于Device:网络融合系统(NCS)540、NCS 5500。

本文档中的信息都是基于特定实验室环境中的设备编写的。本文档中使用的所有设备最初均采用原始（默认）配置。如果您的网络处于活动状态，请确保您了解所有命令的潜在影响。

Microloop概述

IS-IS是一种广泛部署的链路状态路由协议，用于大规模服务提供商网络，因为它具有快速收敛性和可扩展性。但是，在链路或节点故障等拓扑更改期间，当路由器在不同时间更新其转发信息库(FIB)时，可能会出现临时转发一致性，通常称为微环路。这些微环路会导致暂时的数据包丢失、延迟增加或流量黑洞，从而对实时和延迟敏感型应用产生负面影响。

IS-IS路由协议利用分段路由（SR和SRv6）微环路避免机制，以防止在网络收敛期间出现此类瞬时转发环路。即使网络正在转换到新的稳定状态，这些机制也能确保无环转发。

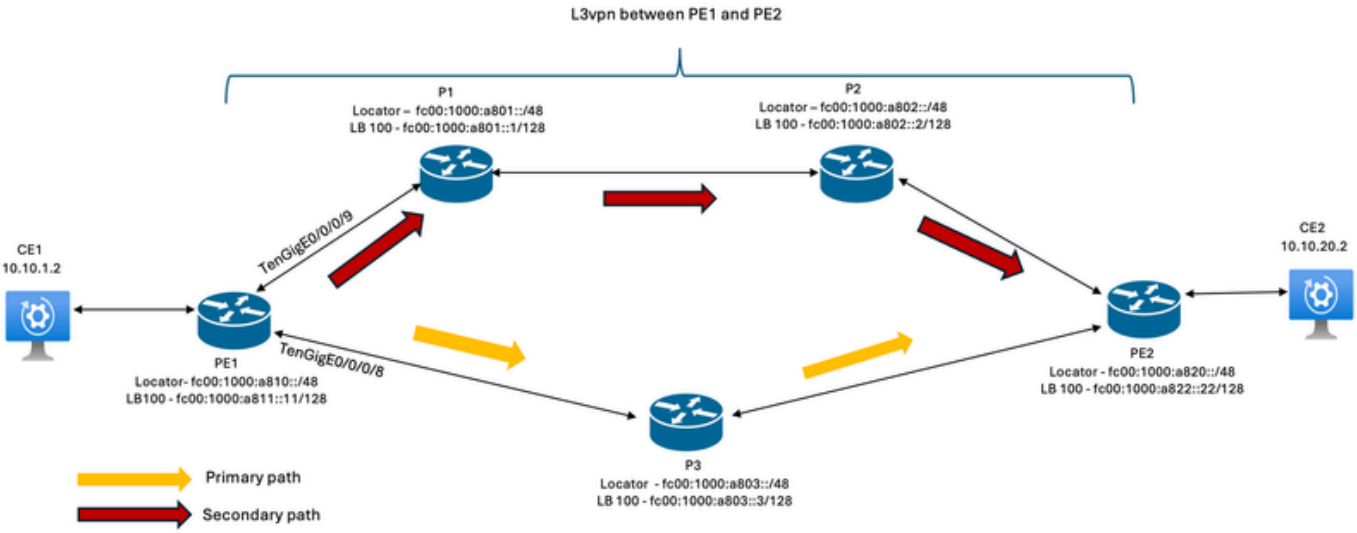


图1.网络拓扑图

通用配置

<#root>

```
interface Loopback100
ipv6 address <>

interface <>
ipv6 enable

router isis <>
 is-type level-2-only
 net <>
 address-family ipv6 unicast
  metric-style wide
  microloop avoidance segment-routing
```

enables Microloop avoidance mechanism

```
  microloop avoidance rib-update-delay <>
```

specify the time in ms

```
  router-id Loopback100
  segment-routing srv6
   locator <>

interface Loopback100
  address-family ipv6 unicast

 interface <>
    point-to-point
   address-family ipv6 unicast
   fast-reroute per-prefix
   fast-reroute per-prefix ti-lfa
```

enables topology-independent loop-free alternates (TI-LFA)

```
segment-routing
 srv6
  encapsulation
   source-address <>
  !
  locators
   locator <>
    micro-segment behavior unode psp-usd
```

enables SRv6 Micro-SIDs (uSIDs) the PSP-USD (Penultimate Segment Pop

- Ultimate Segment Pop

) flavor

```
    prefix <configure the locator >

router bgp <>
vrf <>
  address-family <> unicast
   segment-routing srv6
```

steering the packet using SRv6 uSID

```
    locator <>
```

处于稳定状态

当网络无变化时，提供商边缘1(PE1)通过边界网关协议(BGP)将10.10.1.0/24网络通告给提供商边缘2(PE2)，而PE2将10.10.20.0/24网络通告给PE1。

<#root>

```
RP/0/RP0/CPU0:PE1#show bgp vrf mobility 10.10.1.0/24 detail
BGP routing table entry for 10.10.1.0/24, Route Distinguisher: 10.10.11.11:0
```

**   SRv6-VPN SID: fc00:1000:a810:e003::/64**

```
Local
    0.0.0.0 from 0.0.0.0 (10.10.11.11), if-handle 0x3c000090
      Origin incomplete, metric 0, localpref 100, weight 32768, valid, redistributed, best, group-best,
      Received Path ID 0, Local Path ID 1, version 8
      Extended community:
```

注意：PE1的fc00:1000:a810:e003::/64 >> fc00:1000:a810 ##定位器，e003 ##函数。

<#root>

```
RP/0/RP0/CPU0: PE1#show bgp vrf mobility 10.10.20.0/24 detail
```

```
  Local
    fc00:1000:a822::22 (metric 2000) from fc00:1000:a822::22 (10.10.22.22), if-handle 0x00000000
      Received Label
```

0xe0030

```
      Origin incomplete, metric 0, localpref 100, valid, internal, best, group-best, import-candidate,
      Received Path ID 0, Local Path ID 1, version 714
```

Extended community:

PSID-Type:L3, SubTLV Count:1, R:0x00,

SubTLV:

T:1(Sid information), Sid:fc00:1000:a820::, F:0x00, R2:0x00, Behavior:63, R3:0x00, SS-TLV Count:

SubSubTLV:

T:1(Sid structure):

Length [Loc-blk,Loc-node,Func,Arg]:[32,16,16,0], Tpose-len:16, Tpose-offset:48

```
      Source AFI: VPNv4 Unicast, Source VRF: default, Source Route Distinguisher: 10.10.22.22:2
```

注意：此10.10.20.0/24是从PE1上的PE2收到的，定位器sid fc00:1000:a820::和函数e0030。

<#root>

```
RP/0/RP0/CPU0: PE1#show route vrf mobility 10.10.20.0/24 detail
  Known via "bgp 100", distance 200, metric 0, type internal
  Routing Descriptor Blocks
    fc00:1000:a822::22, from fc00:1000:a822::22
<snip>
```

SRv6 Headend: H.Encaps.Red [f3216], SID-list {fc00:1000:a820:e003::}

注意：在稳定状态下，将发往客户边缘2(CE2)的数据包发送到PE2，其中包含Internet协议

版本6(IPv6)报头目标地址fc00:1000:a820:e003::。

PE1稳态发送的数据包。

<#root>

```
Frame 2: 136 bytes on wire (1088 bits), 136 bytes captured (1088 bits)
Ethernet II, Src: Cisco_a7:8a:0d (c4:b2:39:a7:8a:0d), Dst: Cisco_ff:d4:16 (a0:b4:39:ff:d4:16)
    Destination: Cisco_ff:d4:16 (a0:b4:39:ff:d4:16)
    Source: Cisco_a7:8a:0d (c4:b2:39:a7:8a:0d)
    Type: IPv6 (0x86dd)

Internet Protocol Version 6, Src: fc00:1000:a811::11, Dst: fc00:1000:a820:e003::


    0110 .... = Version: 6
    <0110 .... = Version: 6 [This field makes the filter match on "ip.version == 6" possible]>
    .... 0000 0000 .... .... .... .... .... = Traffic Class: 0x00 (DSCP: CS0, ECN: Not-ECT)
    .... 0000 0000 0000 1110 1111 = Flow Label: 0x000ef
    Payload Length: 82
    Next Header: IPIP (4)
    Hop Limit: 254
    Source Address: fc00:1000:a811::11
    <Source or Destination Address: fc00:1000:a811::11>
    <[Source Host: fc00:1000:a811::11]>
    <[Source or Destination Host: fc00:1000:a811::11]>
    Destination Address: fc00:1000:a820:e003::
    <Source or Destination Address: fc00:1000:a820:e003::>
    <[Destination Host: fc00:1000:a820:e003::]>
    <[Source or Destination Host: fc00:1000:a820:e003::]>
Internet Protocol Version 4, Src: 10.10.1.2, Dst: 10.10.20.2
    0100 .... = Version: 4
    .... 0101 = Header Length: 20 bytes (5)
    Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)
```

注意：

- 源：用作源地址的PE1的fc00:1000:a811::11 >>环回100。
- Dst :fc00:1000:a820:e003::>> PE2定位器：功能.
- ipv6传输报头，接下来是IPv4负载。

通过独立于拓扑的无环路备用(TI-LFA)按计算的主路径和备用路径：

<#root>

```
RP/0/RP0/CPU0:PE1#show cef ipv6 fc00:1000:a820::
   via fe80::a2b4:39ff:feff:d416/128, TenGigE0/0/0/9, 11 dependencies, weight 0, class 0,

backup (TI-LFA
```

```
) [flags 0xb00]
    path-idx 0 NHID 0x0 [0x8ef0f2b0 0x0]
    next hop fe80::a2b4:39ff:feff:d416/128, Repair Node(s): fc00:1000:a802::2
    local adjacency
```

SRv6 H.Insert.Red SID-list {fc00:1000:a802::}

```
   via fe80::9ee1:76ff:feca:e8a8/128, TenGigE0/0/0/8, 4 dependencies, weight 0, class 0, protected [fla
    path-idx 1 bkup-idx 0 NHID 0x0 [0x8f2db710 0x0]
    next hop fe80::9ee1:76ff:feca:e8a8/128
```

PE1接口TenGigE0/0/0/8（主路径）关闭时的事件顺序

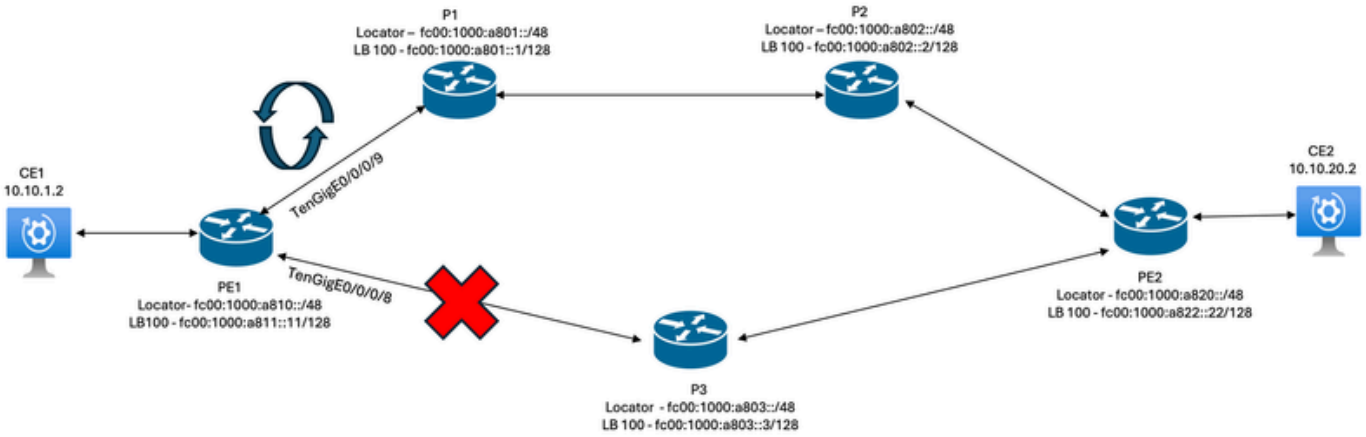


图2.微环路出现

在PE1上，当链路tengig 0/0/8断开时，在PE1和P1之间的备用路径上，会怀疑存在微环路，进而触发PE1上的微环规避(MLA)机制。

<#root>

```
RP/0/RP0/CPU0:PE1#show logging
RP/0/RP0/CPU0:
```

Mar 21 08:30:10.244

```
 UTC: ifmgr[307]: %PKT_INFRA-LINK-5-CHANGED : Interface TenGigE0/0/0/8, changed state to Administrative
```

当接口TenGigE0/0/0/8关闭时，首先出现快速重路由(FRR)事件，即流量在TI-LFA路径上发送。

<#root>

```
RP/0/RP0/CPU0:PE1#show cef trace
Mar 21 08:30:10.244 fib/common/frr 0/RP0/CPU0 43# t5991 Common: FRR-ITF-EVENT: proto=3 type=0 ifh=0x3c0
Mar 21 08:30:10.244 fib/common/frr 0/RP0/CPU0 13# t5991 IPv6: FRR-LOOKUP-DONE: evt=0, ifh=0x3c0000a0,
```

Mar 21 08:30:10.244 fib/common/frr 0/RP0/CPU0 12# t5991 IPv6: FRR-ITF-EVENT: Global Active; handle:0x3c0

Mar 21 08:30:10.244 fib/common/frr 0/RP0/CPU0 13# t5991 IPv6: FRR-ITF-EVENT: FRR Active; handle:0x3c000
Mar 21 08:30:10.244 fib/common/frr 0/RP0/CPU0 1# t5991 IPv6: FRR-EVENT: evt=0, notify protocol=1, ifh=0
Mar 21 08:30:10.244 fib/common/fast 0/RP0/CPU0 20# t5991 Common: PLAT-UPD-FAST: Proto=common, Obj[FIB_DA

在08:30:10:307，存在IS-IS删除邻接关系。

<#root>

RP/0/RP0/CPU0:PE1#show isis lsp last 20

08:30:10:307

    1 Te0/0/0/8   DELADJ

在08:30:10:358，收到标签交换路径(LSP)，计算小型封装热插拔(SFP)，并激活uloop。

<#root>

RP/0/RP0/CPU0:PE1#show isis spf-log detail

08:30:10:358 FSPF   2   5   2   PE1.00-00 DELADJ LINKBAD

  Delay: 50ms (since first trigger)
                        46257ms (since end of last calculation)
  Trigger Link: P3.00
  Trigger Next Hop: P3
  New LSP Arrivals: 0

  SR uloop: **Link Down**

在IS-IS跟踪中更详细地查看这些事件，在这种情况下，配置的路由信息库(RIB)更新延迟时间为：65535毫秒~65秒

<#root>

RP/0/RP0/CPU0:PE1#show isis trace all
Mar 21 08:30:10.308 isis/Mring_2801/std 0/RP0/CPU0 t8712 isis_roca_event_schedule_result_debug:329
Mar 21 08:30:10.308 isis/Mring_2801/spf 0/RP0/CPU0 t8712 isis_roca_spf_linkchanged_trigger:2609

Mar 21 08:30:10.358 isis/Mring_2801/std 0/RP0/CPU0 6669# t8712 isis_roca_event_start:1541

Mar 21 08:30:10.358 isis/Mring_2801/sr_ 0/RP0/CPU0 t8712 isis_roca_sr_uloop_prep:3069

Mar 21 08:30:10.358 isis/Mring_2801/sr_ 0/RP0/CPU0 8451# t8712 isis_roca_uloop_install_exp_path:3915

Mar 21 08:30:10.358 isis/Mring_2801/sr_ 0/RP0/CPU0 t8712   isis_roca_prefix_update_run:1040

```
Mar 21 08:30:10.864 isis/Mring_2801/std 0/RP0/CPU0 t8712 isis_roca_frr_run:1538
```

Mar 21 08:31:15.893 isis/Mring_2801/sr_ 0/RP0/CPU0 t8712 isis_ip_rib_worker_delayed_update_run:2344

MLA处于活动状态时的RIB状态。

<#root>

```
RP/0/RP0/CPU0:PE1#show route ipv6 fc00:1000:a820:: detail
Routing entry for fc00:1000:a820::/48
  Routing Descriptor Blocks
    fe80::a2b4:39ff:feff:d416, from fc00:1000:a822::22, via TenGigE0/0/0/9
      Route metric is 6000
<snip>
      SRv6 Headend: H.Insert.Red [f3216], SID-list {fc00:1000:a802::} ##
```

this locator of P2 is inserted before the SRH

MLA处于活动状态时的思科快速转发(CEF)状态。

<#root>

```
RP/0/RP0/CPU0: PE1#show cef ipv6 fc00:1000:a820:: detail
local adjacency to TenGigE0/0/0/9
<snip>
   via fe80::a2b4:39ff:feff:d416/128, TenGigE0/0/0/9, 10 dependencies, weight 0, class 0 [flags 0x0]
    SRv6 H.Insert.Red SID-list {fc00:1000:a802::}
```

P node (locator of P2)sid is inserted into the packet

```
    Load distribution: 0 (refcount 9)
    Hash  OK  Interface                 Address
    0     Y   TenGigE0/0/0/9            fe80::a2b4:39ff:feff:d416
```

从PE1发出的数据包通过P1转发，其中P2s微段标识符(uSID)作为第一个活动目标。当数据包到达P2时，与uSID关联的SRv6行为会触发分段路由报头(SRH)解封，之后原始数据包在MLA转发下转发到PE2。

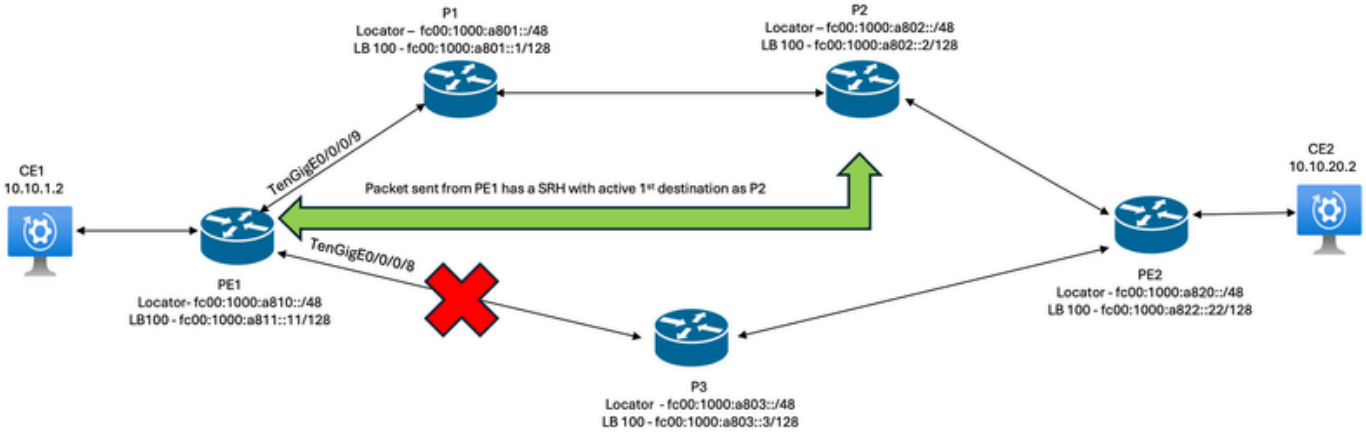


图3.在MLA期间采用的路径

PE1和P1在MLA期间转发的数据包。

<#root>

```
Frame 1: 160 bytes on wire (1280 bits), 160 bytes captured (1280 bits)
Ethernet II, Src: Cisco_a7:8a:0d (c4:b2:39:a7:8a:0d), Dst: Cisco_ff:d4:16 (a0:b4:39:ff:d4:16)

Internet Protocol Version 6, Src: fc00:1000:a811::11, Dst: fc00:1000:a802:: >> during MLA the 1st active


0110 .... = Version: 6
<0110 .... = Version: 6 [This field makes the filter match on "ip.version == 6" possible]>
.... 0000 0000 .... .... .... .... .... = Traffic Class: 0x00 (DSCP: CS0, ECN: Not-ECT)
.... 0000 0000 0000 1110 1111 = Flow Label: 0x000ef
Payload Length: 106

Next Header: Routing Header for IPv6 (43) >> indicates the next header is a SRH


Hop Limit: 254
Source Address: fc00:1000:a811::11
<Source or Destination Address: fc00:1000:a811::11>
<[Source Host: fc00:1000:a811::11]>
<[Source or Destination Host: fc00:1000:a811::11]>
Destination Address: fc00:1000:a802::
<Source or Destination Address: fc00:1000:a802::>
<[Destination Host: fc00:1000:a802::]>
<[Source or Destination Host: fc00:1000:a802::]>

Routing Header for IPv6 (Segment Routing) >>>>>>>>> SRH header which contains the orginal PE2 locator


Next Header: IPIP (4)
Length: 2
[Length: 24 bytes]
Type: Segment Routing (4)

Segments Left: 1


Last Entry: 0
Flags: 0x00
Tag: 0000

Address[0]: fc00:1000:a820:e003:: >>>>>>>>>>>>>>>> PE2 locator : function
```

```
Internet Protocol Version 4, Src: 10.10.1.2, Dst: 10.10.20.2
0100 .... = Version: 4
.... 0101 = Header Length: 20 bytes (5)
```

解封后，P2发送原始数据包，P2在MLA期间删除SRH。

<#root>

```
Frame 1: 136 bytes on wire (1088 bits), 136 bytes captured (1088 bits)
Ethernet II, Src: Cisco_87:d8:58 (b0:a6:51:87:d8:58), Dst: Cisco_af:48:01 (c8:47:09:af:48:01)

Internet Protocol Version 6, Src: fc00:1000:a811::11, Dst: fc00:1000:a820:e003::


0110 .... = Version: 6
<0110 .... = Version: 6 [This field makes the filter match on "ip.version == 6" possible]>
.... 0000 0000 .... .... .... .... .... = Traffic Class: 0x00 (DSCP: CS0, ECN: Not-ECT)
.... 0000 0000 0000 1110 1111 = Flow Label: 0x000ef
Payload Length: 82
Next Header: IPIP (4)
Hop Limit: 252
Source Address: fc00:1000:a811::11
<Source or Destination Address: fc00:1000:a811::11>
<[Source Host: fc00:1000:a811::11]>
<[Source or Destination Host: fc00:1000:a811::11]>
Destination Address: fc00:1000:a820:e003::
<Source or Destination Address: fc00:1000:a820:e003::>
<[Destination Host: fc00:1000:a820:e003::]>
<[Source or Destination Host: fc00:1000:a820:e003::]>
Internet Protocol Version 4, Src: 10.10.1.2, Dst: 10.10.20.2
Data (62 bytes)
```

收敛之后（在RIB更新延迟计时器之后），将删除插入的SRH，并在收敛的最佳内部网关协议(IGP)路径上发送数据包。

```
RP/0/RP0/CPU0: PE1#show cef ipv6 fc00:1000:a822::22/128
local adjacency to TenGigE0/0/0/9
 Prefix Len 128, traffic index 0, precedence n/a, priority 1
   via fe80::a2b4:39ff:feff:d416/128, TenGigE0/0/0/9, 9 dependencies, weight 0, class 0 [flags 0x0]
    path-idx 0 NHID 0x0 [0x8ef0f2b0 0x0]
    next hop fe80::a2b4:39ff:feff:d416/128
    local adjacency
```

时间表摘要

时间	操作	机制

08:30:10.244	链路故障，流量转移到备用路径	TI-LFA
08:30:10:307	已收到LSP，已计算新的最短路径优先(SPF)	MLA（触发）
08:30:10.358	RIB更新延迟，流量使用MLA隧道	MLA（活动）
08:31:15.893	延迟计时器到期，最终路径安装在FIB中	完全收敛

结论

本文档详细介绍IS-IS通告的SRv6 uSID路径在网络融合期间如何支持MLA。通过将有序的拓扑意图直接编码到uSID列表中，流量通过确定的节点序列进行引导，确保无环转发，即使IS-IS SPF计算在网络中暂时不一致。

在收敛期间，从入口PE出出的数据包采用预先计算的uSID序列，通过中间P节点，而不依赖瞬时IGP下一跳决策。指定uSID端点的解封行为可确保受保护网段完成后顺利转换回本机转发。IS-IS控制平面更新和SRv6 uSID数据平面行为之间的这种紧密结合的交互可实现快速且确定的重新路由。

基于IS-IS uSID的MLA为无微环收敛提供了可扩展、拓扑感知且操作简单的解决方案，使其非常适合支持SRv6的大型网络，在这些网络中，快速重新路由和确定性流量控制至关重要。

命令

- #show isis instance <> ipv6 microloop avoidance <prefix> detail
- #show isis

关于此翻译
思科采用人工翻译与机器翻译相结合的方式将此文档翻译成不同语言，希望全球的用户都能通过各自的语言得到支持性的内容。

请注意：即使是最好的机器翻译，其准确度也不及专业翻译人员的水平。

Cisco Systems, Inc. 对于翻译的准确性不承担任何责任，并建议您总是参考英文原始文档（已提供链接）。