

CE 交换机常见故障 & 如何使用 Display 命令排查问题

文档版本

01

发布日期

2020-11-13



版权所有 © 华为技术有限公司 2020。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明



HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或默示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

华为技术有限公司

地址： 深圳市龙岗区坂田华为总部办公楼 邮编： 518129

网址： <https://www.huawei.com>

客户服务邮箱： support@huawei.com

客户服务电话： 4008302118

目 录

1 CE 交换机常见故障 & 如何使用 Display 命令排查问题	1
1.1 简介	1
1.2 CE 交换机常见的硬件故障有哪些	1
1.2.1 电源故障的原因有哪些	2
1.2.2 风扇故障的原因有哪些	5
1.2.3 单板故障有哪些	6
1.2.4 光模块故障的原因有哪些	10
1.2.5 接口故障有哪些	12
1.3 OSPF、BGP 故障可能的原因是什么	13
1.3.1 OSPF 状态异常的原因有哪些	14
1.3.2 BGP 状态异常的原因有哪些	19
1.4 堆叠故障可能的原因是什么	25
1.4.1 堆叠配置失败的原因有哪些	25
1.4.2 堆叠成员异常重启的原因有哪些	29
1.4.3 堆叠无法主备倒换的原因有哪些	32
1.5 VXLAN 故障可能的原因是什么	33
1.5.1 VXLAN 业务丢包的原因有哪些	33
1.5.2 VM 业务异常的原因有哪些	35
1.5.3 VXLAN 路由不通的原因有哪些	36
1.5.4 VM 迁移时间过长的原因有哪些	40
1.6 相关信息	41

1 CE 交换机常见故障 & 如何使用 Display 命令排查问题

1.1 简介

1.2 CE交换机常见的硬件故障有哪些

1.3 OSPF、BGP故障可能的原因是什么

1.4 堆叠故障可能的原因是什么

1.5 VXLAN故障可能的原因是什么

1.6 相关信息

1.1 简介

设备硬件故障、BGP状态异常、OSPF状态异常、堆叠配置与状态异常、VXLAN配置与状态异常是CE交换机较为常见的故障类型。本文汇总了可能造成上述常见故障的原因，提供了相关的Display命令，帮助实现故障原因的快速定位。

- [1.2 CE交换机常见的硬件故障有哪些](#)总结了常见的硬件故障类型，列举了可能的原因，提供了相应的定位方法。
- [1.3 OSPF、BGP故障可能的原因是什么](#)列举了OSPF配置、BGP配置的常见错误，总结了BGP路由、OSPF路由异常的可能原因，提供了相应的定位方法。
- [1.4 堆叠故障可能的原因是什么](#)总结了常见的堆叠故障类型，给出了可能的原因及相应的定位方法。
- [1.5 VXLAN故障可能的原因是什么](#)列举了VXLAN配置的常见错误，总结了VXLAN路由、VXLAN业务异常的可能原因，提供了相应的定位方法。

1.2 CE 交换机常见的硬件故障有哪些

CE交换机使用过程中可能会出现电源告警、风扇告警、单板无法上电、单板异常重启、光模块状态异常、接口状态异常等现象。以下内容可以帮助快速定位造成上述现象的原因。

1.2.1 电源故障的原因有哪些

1.2.2 风扇故障的原因有哪些

1.2.3 单板故障有哪些

1.2.4 光模块故障的原因有哪些

1.2.5 接口故障有哪些

1.2.1 电源故障的原因有哪些

如果设备存在电源告警，可以通过告警信息确定故障原因，并进行相应的操作。如果存在电源功率状态异常，请更换电源或者联系技术支持人员。

表 1-1 电源模块故障定位命令

定位方法	执行命令
根据电源告警信息查找故障原因	display device alarm hardware
确认电源状态是否正常	display device power system

根据电源告警信息查找故障原因

通过**display device alarm hardware**命令查询是否存在电源模块告警。如果存在，可按照**表1-2**操作处理。

```
<HUAWEI> display device alarm hardware
```

Index	Level	Date	Time	Info
1	Major	2016-04-18	18:19:29	The power completely failed.(Power=POWER 1/1, Reason=The actual power exceeded the rated power.)
2	Critical	2016-05-31	11:15:22	The board partially failed.(PowerID=POWER 1/1, Reason=Power IIC failed.)

表 1-2 电源告警说明

电源告警类型	告警说明	建议操作
Communication failed between power and CMU	电源与CMU通信故障	更换电源，如果仍存在问题，则尝试进行监控板的更换。
Fans For Cooling The Power Supply Module Failed	电源内部风扇故障	1. 将故障的电源断电。 2. 更换电源模块。
Power Partial Failure	电源模块内部故障	请联系技术支持人员。
Over-input-voltage occurred on the power module	电源模块输入过压	1. 检查环境的外部供电是否正常。 2. 若仍存在告警，则更换电源模块。

电源告警类型	告警说明	建议操作
Under-input-voltage occurred on the power module	电源模块输入欠压	<ol style="list-style-type: none"> 1. 检查环境的外部供电是否正常，不正常则需要修复外部电路。 2. 若仍存在告警，请与技术支持人员确认是否更换电源模块。
The power module was installed but not powered on	电源模块在位但未上电	<ol style="list-style-type: none"> 1. 检查环境的外部供电是否正常。 2. 若供电正常，则更换电源模块。
Over-output-voltage occurred on the power module	电源模块输出过压	<ol style="list-style-type: none"> 1. 检查环境的外部供电是否正常，不正常则需要修复外部电路。 2. 若仍存在告警，则更换电源模块。
The output of the power module failed	电源模块输出故障	<ol style="list-style-type: none"> 1. 检查环境的外部供电是否正常，不正常则需要修复外部电路。 2. 若仍存在告警，则更换电源模块。
The power module supply failed	电源模块供电故障	<ol style="list-style-type: none"> 1. 检查电源模块的风扇是否被防尘网堵住无法散热，如果防尘网被堵住则需要清洗防尘网。 2. 检查系统以及机房的通风是否正常，如果环境温度较高，则需要改善环境通风。 3. 检查环境的外部供电是否正常，不正常则需要修复外部电路。 4. 若仍存在告警，则更换电源模块。
The actual power exceeded the rated power	电源实时功率超过额定功率	<ol style="list-style-type: none"> 1. 查看是否使用标准电源，如果不是请执行步骤2，如果是请执行步骤3。 2. 更换标准电源。 3. 联系技术支持人员。

确认电源状态是否正常

使用命令 **display device power system** 查看设备的电源状态及功耗情况。

<HUAWEI> display device power system

Average power consumption: 2431 W

Current power consumption: 2442 W

Power manage cycle: 1 hour

Power manage mode: Standard

Power Redundant: Module Total

Power Capacity without Backup: 16200 W

Total Power Capacity with Backup: 18900 W

Total Power Allocated: 4866 W

Total Power Available: 11269 W

Total Power reserved for MPU,SFU,CMU and FAN Board(s): 65 W

Power Supply Module Information:

PowerNo	Present	Mode	State (Ampere)	Current (Volt)	Voltage (Watts)	ActualPower (Watts)	RatedPower
PWR1	YES	AC	Supply	8.6	53.6	465	2700
PWR2	YES	AC	Supply	7.7	53.6	417	2700
PWR3	YES	AC	Supply	6.2	53.6	333	2700
PWR4	YES	AC	Supply	5.8	53.6	312	2700
PWR5	YES	AC	Supply	5.5	53.6	296	2700
PWR6	YES	AC	Supply	6.0	53.6	326	2700
PWR7	YES	AC	Supply	5.4	53.6	292	2700
PWR8	YES	AC	NotSupply	N/A	N/A	N/A	N/A

Board Power Information:

Slot	BoardType	State	ActualPower (Watts)	RatedPower (Watts)
1	LPU	On	393	546
2	LPU	On	385	546
3	LPU	On	376	546
4	LPU	On	378	546
5	LPU	On	328	546
10	MPU	On	33	43
11	CMU	On	--	59
12	CMU	On	--	59
13	SFU	On	107	145
14	SFU	On	106	145
15	SFU	On	102	145
16	SFU	On	111	145
17	SFU	On	107	145
18	SFU	On	101	145
FAN1	FAN	On	--	85
FAN2	FAN	On	--	85
FAN3	FAN	On	--	85
FAN4	FAN	On	--	85
FAN5	FAN	On	--	85
FAN6	FAN	On	--	85
FAN7	FAN	On	--	85
FAN8	FAN	On	--	85
FAN9	FAN	On	--	85
FAN10	FAN	On	--	85
FAN11	FAN	On	--	85
FAN12	FAN	On	--	85
FAN13	FAN	On	--	85

N/A:Board or power not available

电源的“State”状态为“Supply”时表示正常供电，为“NotSupply”时表示供电异常，无电流输出，请更换电源。

1.2.2 风扇故障的原因有哪些

CE交换机运行过程中出现风扇噪音过大与风扇异常告警的情况，可以使用下面的方法来帮助查找原因。

表 1-3 风扇模块故障定位常用命令

可能的原因	执行命令
风扇转速设置不合理	<code>display device fan</code>
风扇损坏	<code>display device</code> <code>display device alarm hardware</code>

风扇转速设置不合理

- 执行命令**display device fan**，查看风扇转速。

如果风扇转速过高，执行命令**set fan-speed slot slot-id percent percent**，调整风扇的转速。该命令使用方法如下：

```
<HUAWEI> system-view
[~HUAWEI] diagnose
[~HUAWEI-diagnose] set fan-speed slot 1 percent 50
```

须知

执行命令前需确认设备温度是否正常。如果设备温度过高，需先排查是否设备本身存在故障或由外部环境导致的设备温度过高，进而导致风扇全速运转。

降低风扇的转速可能会导致设备温度过高，请谨慎使用此命令。

如果调整风扇转速无效，噪音大的现象依然未消除，请执行命令**upgrade fan**，升级风扇软件。该命令使用方法如下：

```
<HUAWEI> system-view
[~HUAWEI] diagnose
[~HUAWEI-diagnose] upgrade fan slot 1
```

说明

进行风扇升级时，风扇将在短时间内全速运转。

风扇损坏

- 使用命令**display device**查看设备的风扇模块状态是否正常。

以CE5800系列交换机的显示为例：

```
<HUAWEI> display device
Device status:
-----
Slot Card Type Online Power Register Alarm Primary
-----
1 - CE5810-48T4S-EI Present On Registered Normal Master
FAN1 FAN-40SB-B Present On Registered Normal NA
FAN2 FAN-40SB-B Present On Registered Normal NA
PWR2 - Present On Registered Normal NA
-----
```


风扇模块故障一般表现为两类：

- 风扇模块内部风扇故障：“Register”状态为“Registered”，但是“Alarm”状态为“Abnormal”。
- 风扇模块不注册：“Register”状态为“Unregistered”。

风扇模块状态正常的情况下，“Register”状态应该为“Registered”，“Alarm”状态应该为“Normal”。

- 执行命令**display device alarm hardware**查看是否存在风扇模块损坏告警。

```
<HUAWEI> display device alarm hardware
```

```
-----
Index  Level  Date      Time      Info
-----
1       critical 2013-10-29 19:18:06  The fan partially failed.(FanID=FAN 2,
Reason=The fan module was in the abnor
mal state.)
-----
```

如果出现风扇模块损坏告警，请更换风扇。

1.2.3 单板故障有哪些

CE交换机运行过程中如果发生单板重启、单板无法上电的情况，可以通过下面的方法来帮助查找问题。

表 1-4 单板故障定位常用命令

故障现象	执行命令
单板重启	display device board reset slot-id
单板无法上电	display device power system
接口板无法注册	display device alarm hardware display device power system display version

单板重启

在诊断视图下通过命令**display device board reset slot-id**查看单板重启原因。

```
<HUAWEI> system-view
[~HUAWEI] diagnose
[~HUAWEI-diagnose] display device board reset 4
Board 4 reset information:
-- 1. DATE:2013-10-28 TIME:16:11:34 BARCODE:NULL RESET Num:1
-- Reason:Product unknown reason(CPU Reset)
-- BootMode:NORMAL
-- BootCode:0x060100ff
```

根据查询到的重启原因，按照表1-5所示的建议操作进行处理。

表 1-5 单板重启原因及建议操作

类型	重启原因描述	重启原因说明	建议操作
用户相关	Reset board from command.	命令行或网管重启。	确认是否因命令行重启或下电单板。

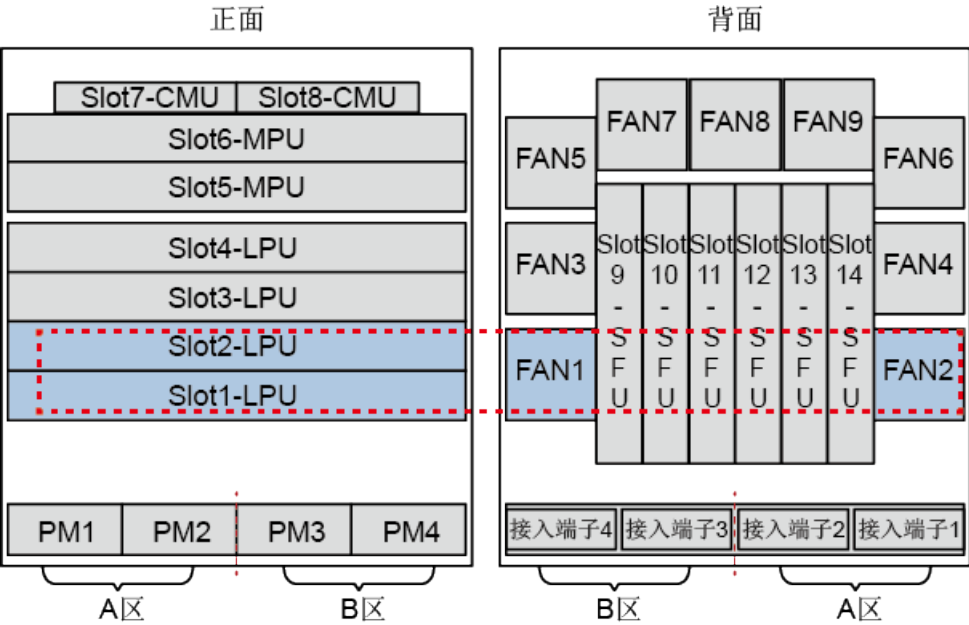
类型	重启原因描述	重启原因说明	建议操作
	Power off the board.		
	Reset board from PIC command.		
	Get pic offline message, and power off.		
	Canbus request to power off the board.		
	Power off board from command.		
环境相关	board cold reset(COLD Reset)	设备掉电。	检查设备的供电环境，确保供电正常。
加载相关	EPLD is upgrade,and reset board.	EPLD逻辑加载后重启。	正常。
	Board update by JTAG, and reset board.	通过监控系统 jtag通道升级后重启。	
	Board update mbus, and reset board.	升级mbus后重启。	
软件异常相关	Board task exception occurs and reset lpu.	检测到软件异常重启。	请收集日志、诊断日志和诊断命令 display reset-snapshot time 的信息，联系技术支持人员处理。
	Board task deadlock occurs and reset lpu.	检测到死循环重启。	请收集日志、诊断日志和诊断命令 display reset-snapshot time 的信息，联系技术支持人员处理。
	Board no enough Memory,and reset board.	内存耗尽重启。	1. 检查内存是否过高。 2. 收集内存信息、日志和诊断日志信息，联系技术支持人员处理。
	Board Memory overload,and reset board.		
	Component report failure.	组件上报失败。	通过诊断命令 display reportfailure number verbose slot slot-id 查看具体的原因，并请收集日志、诊断日志和诊断命令 display reset-snapshot time 的信息，联系技术支持人员处理。

类型	重启原因描述	重启原因说明	建议操作
	LAN Switch parity/ecc error, and reset board.	芯片发生了软失效。	请收集日志、诊断日志和诊断命令 display reset-snapshot time 的信息，联系技术支持人员处理。
设备管理相关	The heartbeat lost and reset lpu.	主控板未收到单板心跳报文。	请收集日志、诊断日志和诊断命令 display reset-snapshot time 的信息，联系技术支持人员处理。
	Semls register failed, and reset board.	检测到单板长时间未成功注册。	收集故障单板在启动过程中的串口打印信息，联系技术支持人员处理。
硬件器件相关	Board selftest error,and reset board.	自检错误。	器件自检失败，尝试拔插单板和更换槽位。如不能解决可判断为单板故障，请联系技术支持人员处理。

单板无法上电

1. 检查单板槽位对应的风扇是否插上，如果风扇不在位，请插上风扇。
- 一般情况下，建议设备的风扇插满，如果存在部分风扇缺失的情况，可能会导致对应槽位的单板无法上电。例如，图 CE12804的单板槽位与风扇的对应关系所示的CE12804交换机上，Slot1和Slot2对应风扇FAN1和FAN2，如果FAN1和FAN2缺失，Slot1和Slot2位置的单板将无法上电。

图 1-1 CE12804 的单板槽位与风扇的对应关系



2. 检查设备剩余功率是否不足，您可以：

- 在任意视图下执行命令**display device power system**，查看功率是否足够，如果功率不足，请增加电源模块。

```
<HUAWEI> display device power system
Average power consumption: 122 W
Current power consumption: 122 W
Power manage cycle: 1 hour
Power manage mode: Standard

Power Supply Module Information:
-----
Slot PowerNo Present Mode State      Current  Voltage  ActualPower
RatedPower
              (Ampere) (Volt)    (Watts)  (Watts)
-----
1  PWR1  YES  AC  Supply  10.2    12.0    122    600
   PWR2  NO   --  --      --      --      --      --
-----

Board Power Information:
-----
Slot Card  BoardType      State      ActualPower RatedPower
              (Watts)  (Watts)
-----
1  --  CE6863-48S6CQ  on        122      272
-----
N/A:Power not available
```

- 使用[硬件配置工具](#)，计算设备所需配置的电源数量。如果电源模块数量不足，请增加电源模块。

接口板无法注册

1. 执行**display device**命令查看接口板注册状态。

```
<HUAWEI> display device
CE12804's Device status:
-----
Slot Sub  Type      Online Power Register Alarm Primary
-----
1  -  -      Present On  Unregistered -  NA
3  -  -      Present On  Unregistered -  NA
5  -  CE-MPUA  Present On  Registered Normal Master
7  -  CE-CMUA  Present On  Registered Normal Master
13 -  CE-SFU04C Present On  Registered Normal NA
PWR1 -  -      Present On  Registered Normal NA
FAN3 -  -      Present On  Registered Normal NA
.....
```

如果接口板Register状态显示为Unregistered，说明接口板未注册。

2. 执行命令**display device alarm hardware**，查看该接口板是否存在告警，如果存在电压或器件告警，则可以判定为接口板故障，可以先更换接口板并联系技术支持人员进行处理。
3. 执行命令**display device power system**，查看功率是否足够。

```
<HUAWEI> display device power system
Average power consumption: 427 W
Current power consumption: 416 W
Power manage cycle: 1 hour
Power manage mode: Standard

Power Redundant: Non
Total Power Capacity without Backup: 2700 W
```

```
Total Power Capacity with Backup: 2700 W
Total Power Allocated: 1030 W
Total Power Available: 1206 W    //系统剩余可分配的功率
Total Power reserved for MPU,SFU,CMU and FAN Board(s): 464 W
.....
```

如果功率不足，请增加电源模块。

- 通过接口板拉手条右上角纸质标签或读取接口板的电子标签确认接口板的型号，然后再通过命令**display version**查询设备的软件版本，确认当前软件版本是否支持该接口板。

1.2.4 光模块故障的原因有哪些

CE交换机的光模块故障原因可根据接口下光模块相关的告警信息以及光模块的状态来进行判断。

表 1-6 光模块故障定位常用命令

定位方法	执行命令
根据光模块的告警信息查找故障原因	display interface transceiver
查看光模块光功率是否正常	display interface transceiver verbose

根据光模块的告警信息查找故障原因

执行命令**display interface transceiver**查看“Alarm information”下光模块是否有告警信息。

```
<HUAWEI> display interface 10ge 1/0/1 transceiver

10GE1/0/1 transceiver information:
-----
Common information:
Transceiver Type           :10GBASE_SR
Connector Type             :LC
Wavelength (nm)           :850
Transfer Distance (m)      :30(62.5um/125um OM1)
                           80(50um/125um OM2)
                           300(50um/125um OM3)
                           400(50um/125um OM4)
Digital Diagnostic Monitoring :YES
Vendor Name                :HUAWEI
Vendor Part Number         :02318169
Ordering Name              :
-----
Manufacture information:
Manu. Serial Number        :AQG269Y
Manufacturing Date         :2013-10-20
Vendor Name                :HUAWEI
-----
Alarm information:
-----
```

如果出现LOS Alarm告警，则说明对端没有信号发送过来，在接口模式下执行命令**display this**查看两端端口是否shutdown，如果端口shutdown了，则执行**undo shutdown**操作。

查看光模块光功率是否正常

执行**display interface transceiver verbose**命令查看光模块发送或接收光功率方面的信息，与光模块的标称光功率进行对比判断。光模块的诊断信息中，可以查看当前发送和接收的光功率值，以及默认的最高和最低功率值。

```
<HUAWEI> display interface 10ge 1/0/1 transceiver verbose
```

```
10GE1/0/1 transceiver information:
```

```
-----  
Common information:
```

```
Transceiver Type      :10GBASE_SR  
Connector Type       :LC  
Wavelength (nm)      :850  
Transfer Distance (m) :30(62.5um/125um OM1)  
                     :80(50um/125um OM2)  
                     :300(50um/125um OM3)  
                     :400(50um/125um OM4)  
Digital Diagnostic Monitoring :YES  
Vendor Name          :HUAWEI  
Vendor Part Number   :02318169  
Ordering Name        :
```

```
-----  
Manufacture information:
```

```
Manu. Serial Number   :AQG269Y  
Manufacturing Date    :2013-10-20  
Vendor Name           :HUAWEI
```

```
-----  
Alarm information:
```

```
-----  
Diagnostic information:
```

```
Temperature (Celsius) :33.68  
Voltage (V)           :3.29  
Bias Current (mA)     :7.97  
Bias High Threshold (mA) :13.20  
Bias Low Threshold (mA) :4.00  
Current RX Power (dBm) : -2.15  
Default RX Power High Threshold (dBm) :1.00  
Default RX Power Low Threshold (dBm) : -11.90  
Current TX Power (dBm) : -2.07  
Default TX Power High Threshold (dBm) :1.00  
Default TX Power Low Threshold (dBm) : -9.30  
-----
```

- 如果接收功率低（Current RX Power < Default RX Power Low Threshold），说明本端接收到的信号过低，则可能出现端口不UP或者UP后报文收发有丢弃，此时请先排查传输距离是否过远，超出了该光模块的传输距离，再排查光模块、光纤是否有损坏。
- 如果接收功率高（Current RX Power > Default RX Power High Threshold），说明本端接收到的信号过高，可能原因是该光模块为长距光模块，而实际传输距离太短，导致信号未衰减，此时应在光模块上增加光衰，以对光模块进行保护。
- 如果发送功率低（Current TX Power < Default TX Power Low Threshold），说明该光模块发送信号不好或光模块本身故障，可能会导致对端接收功率低，而造成端口不UP或者UP后报文收发有丢弃，请与技术支持人员联系。
- 如果发送功率高（Current TX Power > Default TX Power High Threshold），说明该光模块发送信号太强，可能会导致对端接收功率高，而造成对端光模块因接收功率持续过高而烧坏，可能原因是本端光模块故障，建议更换光模块。

因此，在端口插入光模块并对接成功后，要对发送或接收光功率方面的告警信息进行排查，避免因功率过低或者过高造成流量或者光模块不正常。

1.2.5 接口故障有哪些

CE交换机常见的接口故障有接口拆分与合并故障、接口不Up、接口丢包等，可以使用下面的方法帮助查找问题来源。

表 1-7 接口故障定位常用命令

故障现象	执行命令
接口拆分、合并故障	display current-configuration display port split
接口不UP	display interface <i>interface-type</i> <i>Interface-number</i> display interface <i>interface-type</i> <i>Interface-number</i> transceiver verbose
接口丢包故障	display interface [<i>interface-type</i> [<i>interface-number</i>]]

接口拆分、合并故障

- 40GE接口拆分为10GE接口失败。
 - 执行命令**display current-configuration**，查看是否有**port split**的配置信息。如果没有，请重新在系统视图下执行命令**port split dimension interface interface-type interface-number1 [to interface-type interface-number2]**，配置接口拆分。
 - 执行**display port split**命令，查看接口的当前状态。如果出现下述回显信息，请重启单板。

```
<HUAWEI> display port split
```

```
-----
Port          Status    Split-port
-----
40GE1/0/1      Enable    40GE1/0/1:1  40GE1/0/1:2
               40GE1/0/1:3  40GE1/0/1:4
40GE1/0/2      Enable    40GE1/0/2:1  40GE1/0/2:2
               40GE1/0/2:3  40GE1/0/2:4
-----
```

- 10GE接口合并成40GE接口失败。
 - 执行命令**display current-configuration**，查看是否有**port split**的配置信息。如果有，请重新在系统视图下执行命令**undo port split dimension interface interface-type interface-number1 [to interface-type interface-number2]**，配置接口合并。
 - 执行**display port split**命令，查看接口的当前状态。如果出现下述回显信息，请重启单板。

```
<HUAWEI> display port split
```

```
-----
Port          Status    Split-port
-----
40GE4/0/3      Disable
-----
```

接口不 UP

此部分可参考[接口不Up怎么办](#)。

- 执行 **display interface** *interface-type interface-number* 命令查看接口状态。
- 执行 **display interface** *interface-type Interface-number transceiver verbose* 命令查看光模块信息。

接口丢包故障

执行命令**display interface** [*interface-type* [*interface-number*]]，查看出现故障的接口报文统计信息，以确认丢包类型。

- 如果丢包类型为**crc**错误报文，且报文数量相对于报文统计总数较少，请检查物理连线接头是否松动或物理连线是否有损坏，发现异常请及时拧紧连线接头或更换传输介质（光纤、铜缆、光模块、网线）。最后在接口视图下执行命令**restart**，重启接口。可以参考[接口收到CRC错包怎么处理](#)进行处理。
- 如果丢包类型为**runts**报文，请检查本端接口接收的报文长度。
 - 如果接收报文长度小于64字节，请检查对端设备发送的报文长度是否正常。出现对端设备发送报文长度异常时，需要修改对端设备相应配置。
 - 如果报文正常（报文长度大于或等于64字节），请执行命令**restart**，重启本端设备接口。
- 如果丢包类型为**giants**报文，请检查本端接口接收的报文长度。执行**display interface** [*interface-type* [*interface-number*]]，可查看字段**The Maximum Frame Length**，该字段表示接口允许通过的最长帧。比较接收报文长度和**The Maximum Frame Length**字段的两者数值大小。
 - 如果接收报文长度大于**The Maximum Frame Length**数值，请在接口下执行**jumboframe enable value1**，配置接口允许通过的最长帧，使之大于或等于接收报文长度。
 - 如果接收的报文长度过大（即报文长度大于命令**jumboframe enable value1**可配置的最大值），请在对端发送报文的设备上执行**mtu mtu**，减小对端接口的最大传输单元数，使之小于或等于**The Maximum Frame Length**数值。
- 如果丢包类型为**discard**，请检查报文转发接口上是否配置了接口队列的队列整形功能，报文转发接口速率是否低于流量带宽。
 - 如果报文转发接口配置了指定接口队列的队列整形功能，则用户可执行**qos queue queue-index shaping cir cir-value [kbps | mbps | gbps] pir pir-value [kbps | mbps | gbps] [cbs cbs-value [bytes | kbytes | mbytes] pbs pbs-value [bytes | kbytes | mbytes]]**，调整队列整形参数。
 - 如果报文转发接口速率低于流量带宽，可通过以下几种方式提高接口速率或者增加可使用的带宽：
 - 当接口的当前配置或者协商的速率未达到接口可提供的最大速率时，使用**speed { 10 | 100 | 1000 }**或**speed { 100 | 1000 | 10000 }**命令配置更高的速率。
 - 更换更高速率的接口，比如将GE接口更换为10GE接口。
 - 使用链路聚合技术，将多个物理接口捆绑为一个逻辑接口以增加带宽。

1.3 OSPF、BGP 故障可能的原因是什么

OSPF路由故障、BGP路由故障的原因种类比较多。为了实现对BGP&OSPF故障原因的快速定位，本文总结了BGP&OSPF故障的常见原因，提供了相应的Display命令。

1.3.1 OSPF状态异常的原因有哪些

1.3.2 BGP状态异常的原因有哪些

1.3.1 OSPF 状态异常的原因有哪些

OSPF配置错误、设备硬件故障均有可能造成OSPF配置失败和OSPF状态异常。下面列举了常见的故障原因以及相应的定位命令，更详细的故障处理案例可以参考[维护宝典](#)相关内容。

表 1-8 OSPF 故障定位常用命令

可能原因	执行命令
OSPF邻居状态异常	<code>display logbuffer size buffersize</code>
链路故障	<code>display this interface</code>
CPU占用率过高	<code>display cpu</code>
接口没有Up	<code>display interface [interface-type [interface-number]]</code>
两端IP地址不在同一网段	<code>display interface interface-type [interface-number]</code>
各接口的MTU不一致	<code>display this interface</code>
各接口的优先级为零	<code>display ospf interface</code>
两端OSPF RouterID配置冲突	<code>display ospf brief</code>
两端OSPF Area配置不一致	<code>display ospf interface</code>
OSPF的其他配置错误	<code>display ospf error</code>
OSPF邻居震荡	<code>display ospf peer last-nbr-down</code>
OSPF路由震荡	<code>display ospf spf-statistics verbose</code>

OSPF 邻居状态异常

执行**display logbuffer size buffersize**命令，查看日志信息。如果出现如下的日志信息，说明故障是邻居状态改变导致。

```
CE12800 %%01 ospfv2comm/6/NBR_CHANGE(l):VS=0-CID=[UINT];Neighbor changes event: neighbor status changed. (ProcessId=[UINT], NbrIpAddr=[IPADDR], NbrEvent=[UINT], NbrPreviousState=[UINT], NbrCurrentState=[UINT])
```

重点关注关键字NbrEvent，此关键字记录的是OSPF邻居Down的原因。OSPF邻居Down的原因一般会有以下几种：

- Inactivity (NbrEvent=7)
发生邻居状态机InactivityTimer事件，表示在deadtime时间内没有收到Hello报文导致OSPF邻居Down，请检查链路故障。
- LLDn (NbrEvent=6)

发生邻居状态机LLDown事件，表示由下层协议通知邻居不可达到，请检查链路故障。

- 1-Way Received (NbrEvent=4)

发生邻居状态机1-Way Received事件，表示因为对端OSPF状态首先变成Down，从而向本端发送1-Way hello，导致本端OSPF状态也变成Down，出现这种情况请排查对端设备的故障。

- Kill Neighbor (NbrEvent=5)

表示接口Down或BFD Down。此时，可以执行**display interface** [*interface-type* [*interface-number*]]命令查看接口状态，排查接口故障。

链路故障

执行ping命令或在接口视图下执行**display this interface**命令。

- 如果状态异常，说明设备链路故障（包括传输设备故障）。
- 如果链路正常，请检查CPU、接口等其他问题。

CPU 占用率过高

执行**display cpu**命令检查故障设备的CPU利用率是否过高。如果CPU利用率过高会导致OSPF无法正常收发协议报文从而导致邻居振荡。

如果CPU利用率过高则先解决此问题，关闭一些不必要的功能。可以参考[CPU占用率过高怎么办](#)解决问题。

接口没有 Up

请执行**display interface** [*interface-type* [*interface-number*]]命令查看接口物理层状态。

- 如果接口物理层状态为Down请先处理接口故障问题，可参考[接口不Up怎么办](#)解决问题。
- 如果接口物理层状态是Up，请执行**display ospf interface**查看接口在OSPF协议下状态是否为正常状态（即接口状态为DR、BDR、DROther或P2P等正常状态）。

```
<HUAWEI> display ospf interface
      OSPF Process 1 with Router ID 10.1.1.1
      Interfaces
```

```
Area: 0.0.0.0      (MPLS TE not enabled)
Interface  IP Address  Type    State  Cost  Pri
Vlanif50   192.168.1.1    Broadcast DR     1     1
```

如果OSPF下的接口为Down，请执行命令**display ospf cumulative**检查OSPF进程下使能的接口数是否超出了当前版本OSPF的规格。

```
<HUAWEI> display ospf cumulative
      OSPF Process 1 with Router ID 10.1.1.1
      Cumulations
```

```
IO Statistics
  Type      Input  Output
  Hello      0       86
  DB Description  0       0
  Link-State Req  0       0
  Link-State Update  0       0
  Link-State Ack  0       0
  ASE: (Disabled)
  LSAs originated by this router
```

```

Router: 1
Network: 0
Sum-Net: 0
Sum-Asbr: 0
External: 0
NSSA: 0
Opq-Link: 0
Opq-Area: 0
Opq-As: 0
LSAs Originated: 1  LSAs Received: 0
Routing Table:
  Intra Area: 1  Inter Area: 0  ASE: 0
Up Interface Cumulate: 1
  Neighbor Cumulate:
=====
  Neighbor cumulative data. (Process 1)
-----
Down:    0 Init:    0 Attempt:  0 2-Way:  0
Exstart: 0 Exchange: 0 Loading:  0 Full:   1
Retransmit Count:1

  Neighbor cumulative data. (Total)
-----
Down:    0 Init:    0 Attempt:  0 2-Way:  0
Exstart: 0 Exchange: 0 Loading:  0 Full:   1
Retransmit Count:1

```

如果超出规格，请减少OSPF使能的接口数。

两端 IP 地址不在同一网段

执行**display interface *interface-type* [*interface-number*]**命令查看两端接口的IP地址。

如果IP地址不在同一网段，请执行**ip address**命令修改两端的IP地址，使其在同一网段。

各接口的 MTU 不一致

如果在接口上使能了**ospf mtu-enable**，则要求接口的MTU一致，否则OSPF邻居无法协商成功。请接口视图下执行**display this interface**命令查看接口MTU信息。

如果接口的MTU值配置不一致，请在接口视图下执行**mtu *mtu***，修改链路两端的MTU值为一致。

接口的优先级为零

对于Broadcast和NBMA类型的网段，各接口的优先级至少有一个是非零的，以确保能够正确的选举出DR，否则两边的邻居状态只能达到2-Way。

执行命令**display ospf interface**，查看接口的优先级。

```

<HUAWEI> display ospf interface
  OSPF Process 1 with Router ID 10.1.1.1
    Interfaces

Area: 0.0.0.0      (MPLS TE not enabled)
Interface  IP Address  Type      State  Cost  Pri
Vlanif50   192.168.1.1   Broadcast P-2-P  1     1

```

如果各接口优先级均为0，请修改接口的优先级。

两端 OSPF RouterID 配置冲突

执行命令**display ospf brief**查看Router ID。

```
<HUAWEI> display ospf brief
```

```
OSPF Process 1 with Router ID 10.1.1.1
OSPF Protocol Information
```

如果冲突则修改配置，否则检查两端OSPF Area配置是否一致。

两端 OSPF Area 配置不一致

执行命令**display ospf interface**检查两端OSPF Area配置是否一致。

```
<HUAWEI> display ospf interface
```

```
OSPF Process 1 with Router ID 10.1.1.1
Interfaces
```

```
Area: 0.0.0.0 (MPLS TE not enabled)
Interface IP Address Type State Cost Pri
Vlanif50 192.168.1.1 Broadcast BDR 1 1
```

如果OSPF Area不一致，请修改相关配置。

OSPF 的其他配置错误

每10秒钟执行一次命令**display ospf error**，持续5分钟。

```
<HUAWEI> display ospf error
```

```
OSPF Process 1 with Router ID 10.1.1.1
OSPF error statistics
```

General packet errors:

```
0 : IP: received my own packet 0 : Bad packet
0 : Bad version 0 : Bad checksum
0 : Bad area id 0 : Drop on unnumbered interface
0 : Bad virtual link 0 : Bad authentication type
0 : Bad authentication key 0 : Packet too small
0 : Packet size > ip length 0 : Transmit error
0 : Interface down 0 : Unknown neighbor
```

HELLO packet errors:

```
0 : Netmask mismatch 0 : Hello timer mismatch
0 : Dead timer mismatch 0 : Extern option mismatch
0 : Router id confusion 0 : Virtual neighbor unknown
0 : NBMA neighbor unknown 0 : Invalid Source Address
```

- 查看Bad authentication type字段，如果这个字段对应的计数值一直增长，表示建立邻居的两台设备配置的OSPF认证类型不一致，需要在两端设备上配置相同认证的类型。
- 查看Hello timer mismatch字段，如果这个字段对应的计数值一直在增长，表示接口上hello timer配置不一致，需要通过检查两端设备接口配置，将hello timer间隔配置一致。
- 查看Dead timer mismatch字段，如果这个字段对应的计数值一直在增长，表示接口的dead timer配置不一致，需要通过检查两端设备接口配置，将dead timer间隔配置一致。
- 查看Extern option mismatch字段，如果这个字段对应的计数值一直在增长，表示区域类型配置不一致（一端配置为普通区域，另一端配置为stub或nssa区域），需要将两端区域类型配置一致。

OSPF 邻居震荡

在故障设备的任意视图执行**display ospf peer last-nbr-down**命令，显示信息的**Immediate Reason**字段表明邻居down的直接原因，**Primary Reason**字段表明邻居down的根本原因，可以根据这两个字段确认OSPF邻居震荡的原因：

```
<HUAWEI> display ospf peer last-nbr-down

OSPF Process 1 with Router ID
192.168.2.200

Last Down OSPF Peer

...
Immediate Reason : Neighbor Down Due to Kill Neighbor
Primary Reason   : Link Fault or Interface Configuration Change
...
```

- 如果Immediate Reason显示为“Neighbor Down Due to LL Down”，则表明链路可能有问题。
- 如果Primary Reason显示为“BFD Session Down”，则表明BFD会话Down。
- 如果Immediate Reason显示为“Neighbor Down Due to Inactivity”，则表明接收Hello报文超时。
- 如果Immediate Reason显示为“Neighbor Down Due to 1-Wayhello”，则表明对端设备没有收到本端设备发送的Hello报文。
- 如果Immediate Reason显示为“Neighbor Down Due to Kill Neighbor”，有可能是本端配置变化引起邻居Down。

OSPF 路由震荡

登录到路由接收者设备，在任意视图下执行**display ospf spf-statistics verbose**命令，查看引起路由重新计算的原因。主要关注Type字段，该字段表示引起路由计算的LSA的类型。

```
<HUAWEI> display ospf spf-statistics verbose
OSPF Process 1 with Router ID 192.168.2.200
Routing table change statistics:
Index: 1
    Time   : 2015-07-15 11:48:46
    Intra  : 2   Added,0   Deleted
    Inter  : 0   Added,0   Deleted
    External: 0   Added,0   Deleted
    The reason of calculation is:Topo
    NO.   Type   LS ID      Adv Router
    1     Router  192.168.3.200  192.168.3.200
Index: 2
    Time   : 2015-07-15 11:48:42
    Intra  : 1   Added,0   Deleted
    Inter  : 0   Added,0   Deleted
    External: 0   Added,0   Deleted
    The reason of calculation is:Topo
    NO.   Type   LS ID      Adv Router
    1     Network 172.16.1.1   192.168.3.200
```

- 如果Type字段对应的LSA类型是Router，说明是Router LSA引起的路由震荡。
- 如果Type字段对应的LSA类型是Network，说明是Network LSA引起的路由震荡。
- 如果Type字段对应的LSA类型是Sum-Net，说明是Sum-Net LSA引起的路由震荡。
- 如果Type字段对应的LSA类型是External或NSSA，说明是External LSA或NSSA LSA引起的路由震荡。

1.3.2 BGP 状态异常的原因有哪些

BGP配置错误、路由不可达、路由策略配置错误、硬件故障均有可能造成BGP配置失败和BGP状态异常。下面列举了常见的故障原因以及相应的定位命令，更详细的故障处理案例可以参考[维护宝典](#)相关内容。

表 1-9 BGP 故障定位常用命令

可能原因	执行命令
BGP邻居未建立	<code>display bgp peer</code>
BGP邻居无法互通	<code>ping</code>
ACL过滤了TCP的179端口	<code>display acl all</code>
邻居的Router ID冲突	<code>display bgp peer</code>
配置的邻居的AS号错误	<code>display bgp peer</code>
用Loopback口建立邻居时没有配置peer connect-interface	<code>peer connect-interface</code>
用Loopback口建立EBGP邻居未配置peer ebgp-max-hop	<code>peer ebgp-max-hop</code>
Peer valid-ttl-hops配置错误	<code>peer valid-ttl-hops hops</code>
对端配置了peer ignore	<code>undo peer ignore</code>
地址族能力	<code>peer enable</code>
路由下一跳不可达导致路由不活跃	<code>display bgp routing-table network { mask mask-length }</code>
路由策略配置不当导致路由无法发布/接收	<code>display current-configuration configuration bgp</code>
路由数量超限导致收到的路由被丢弃	<code>display current-configuration configuration bgp include peer destination-address</code> <code>display current-configuration configuration bgp include peer group-name</code>
标签超限导致私网路由无法发布	<code>display bgp vpnv4 all routing-table ipv4-address [mask mask-length]</code>
ERT/IRT不匹配导致路由无法交叉到私网路由表中	<code>display current-configuration configuration vpn-instance</code>
对端（路由接收者）没有配置peer上的ip-prefix入口策略	<code>display current-configuration configuration bgp</code>
对端（路由接收者）没有配置peer上的ip-prefix入口策略对应的前缀列表	<code>display ip ip-prefix ip-prefix-name</code>

BGP 邻居未建立

执行**display bgp peer**命令，查看BGP邻居状态是否是Established状态。

- 如果不是Established状态，请检查AS号、Router-ID、loopback口等BGP配置。
- 如果是Established状态，说明BGP邻居已成功建立。

BGP 邻居无法互通

使用**ping**命令检测BGP邻居之间是否可以Ping通。

- 如果可以Ping通，则说明BGP邻居之间有可达的路由并且链路传输也没有问题。

说明

请使用命令**ping -a source-ip-address -s packetsize host**或**ping ipv6 -a source-ipv6-address -s packetsize destination-ipv6-address**来检测两端的互通性，因为带源地址可以同时检测两端路由是否正常，指定ping的字节可以检查大包在链路上传输是否正常。

- 如果不能Ping通，请参考[维护宝典](#)Ping不通部分处理Ping不通问题，排除链路传输的故障问题。

ACL 过滤了 TCP 的 179 端口

在两端执行**display acl all**命令查看是否禁止TCP的179端口。

```
<HUAWEI> display acl all
Advanced ACL 3001, 2 rules
ACL's step is 5
ACL's match-order is config
rule 5 deny tcp source-port eq bgp
rule 10 deny tcp destination-port eq bgp
```

如果有禁止TCP的179端口的ACL，请执行**undo rule rule-id**命令取消配置。

邻居的 Router ID 冲突

在两端分别查看无法建立的BGP邻居的情况，例如ipv4单播邻居无法建立可以执行**display bgp peer**命令，查看Router ID是否冲突。显示Router ID信息的命令行示例如下：

```
<HUAWEI> display bgp peer
BGP local router ID : 1.1.1.1
Local AS number : 65001
Total number of peers : 12          Peers in established state : 4

Peer      V      AS  MsgRcvd  MsgSent  OutQ  Up/Down    State PrefRcv
10.9.0.8   4      100    1601    1443    0 23:21:56 Established 10000
10.10.0.10 4      200    1565    1799    0 23:15:30 Established 9999
```

说明

查看其他地址族的邻居可以使用如下命令：

- **display bgp vpnv4 all peer**查看所有VPNv4的对等体信息。
- **display bgp ipv6 peer**查看IPv6的对等体信息。
- **display bgp vpnv6 all peer**查看所有VPNv6的对等体信息。

如果Router ID冲突，请在BGP视图下运行命令**router id**将Router ID修改为不同（一般会用Loopback口的地址作为本端的Router ID）。

配置的邻居的 AS 号错误

在邻居两端分别执行**display bgp peer**，检查邻居的AS号是否是对端的AS号。

```
<HUAWEI> display bgp peer
BGP local router ID : 223.5.0.109
Local AS number : 41976
Total number of peers : 12                Peers in established state : 4

Peer      V      AS  MsgRcvd  MsgSent  OutQ  Up/Down    State PrefRcv
10.9.0.8   4      100    1601    1443    0 23:21:56 Established 10000
10.10.0.10 4      200    1565    1799    0 23:15:30 Established  9999
```

说明

查看其他地址族的邻居可以使用如下命令：

- **display bgp vpnv4 all peer**查看所有VPNv4的对等体信息。
- **display bgp ipv6 peer**用来查看IPv6的对等体信息。
- **display bgp vpnv6 all peer**查看所有VPNv6的对等体信息。

如果AS号配置错误，请将AS号配置为对端的AS。

用 Loopback 口建立邻居时没有配置 peer connect-interface

如果邻居两端使用Loopback口建立邻居，则需要使用命令**peer connect-interface**指定相应的Loopback口为发送BGP报文的源接口或者指定相应的ip地址为相应的源地址。

用 Loopback 口建立 EBGP 邻居未配置 peer ebgp-max-hop

如果直连设备用Loopback口建立EBGP邻居，或者非直连多跳设备建立EBGP邻居，则需要配置命令**peer ebgp-max-hop**指定允许的最大跳数 *hop-count*。

- 直连设备使用Loopback口建立连接时，*hop-count*只要大于1即可。
- 非直连设备建立连接时需要指定 *hop-count* 为相应的跳数。

说明

命令**peer ebgp-max-hop**的配置是对称的，即需要在BGP会话两端同时使能该命令。

Peer valid-ttl-hops 配置错误

如果有该配置，请确认**peer valid-ttl-hops hops**是否正确：如果配置为 *hops*，则被检测的报文的TTL值有效范围为[255-*hops*+1, 255]。

其中 *hops* 是BGP会话两端之间的跳数值，直连设备之间的 *hops* 为1。

说明

命令**peer valid-ttl-hops**的配置是对称的，即需要在BGP会话两端同时配置该命令。

对端配置了 peer ignore

如果对端配置了**peer ignore**，说明由于某种原因对端暂时不想和本端建立邻居。如果想建立邻居时，将对端的**undo peer ignore**配置去使能即可。

地址族能力

请检查BGP会话两端的地址族能力是否匹配。例如，建立BGP VPNv4邻居时，需要两端都要在BGP-VPNv4地址族下配置命令**peer enable**。

如果一端已配置而另一端没有配置时，配置的一端BGP邻居状态为“No neg”，表示对端BGP对等体没有使能该地址族。

路由下一跳不可达

在路由的发送端执行**display bgp routing-table network { mask | mask-length }**命令查看目标路由（*network*表示目标路由前缀），确认路由是否活跃，并且查看此路由是否已经被发送给路由接收端。

```
<HUAWEI> display bgp routing-table 10.0.0.0 8
```

```
BGP local router ID : 10.1.1.2
Local AS number : 100
Paths: 1 available, 1 best, 1 select
BGP routing table entry information of 10.0.0.0/8:
From: 10.1.1.1 (121.1.1.1)
Route Duration: 4d21h29m39s
Relay IP Nexthop: 172.16.1.1
Relay IP Out-Interface: Vlanif50
Original nexthop: 10.1.1.1
Qos information : 0x0
AS-path Nil, origin incomplete, localpref 100, pref-val 0, valid, internal, best, select, active, pre 255
Aggregator: AS 100, Aggregator ID 121.1.1.1
Advertised to such 1 peers:
10.3.3.3
```

- 如果目标路由不活跃，请确认IP路由表中是否存在到BGP下一跳（Original nexthop）的路由，如果不存在说明BGP路由不发布是由于路由下一跳不可达导致，请确认为何没有到BGP下一跳（Original nexthop）的路由（一般属于IGP或静态路由问题）。
- 如果目标路由活跃且被优选，但没有显示发送给路由接收端，请检查路由发送端的出口策略。
- 在路由接收端执行**display bgp routing-table network { mask | mask-length }**查看是否收到目标路由。
 - 如果收到目标路由，请重复执行上述步骤判断路由下一跳是否可达并且是否被优选。
 - 如果没有收到目标路由，请检查路由策略配置是否正确。

说明

在BGP4+的组网环境中，请使用**display bgp routing-table ipv6-address prefix-length**查看是否收到目标路由。

路由策略配置不当导致路由无法发布/接收

在路由的发送端/接收端执行**display current-configuration configuration bgp**命令查看BGP配置，确认是否配置邻居的出口/入口策略。

```
<HUAWEI> display current-configuration configuration bgp
#
bgp 100
peer 10.1.1.1 as-number 100
#
ipv4-family unicast
filter-policy ip-prefix aaa import
```

```
filter-policy ip-prefix aaa export
peer 10.1.1.1 enable
peer 10.1.1.1 filter-policy acl-name acl-name import
peer 10.1.1.1 filter-policy acl-name acl-name export
peer 10.1.1.1 as-path-filter 1 import
peer 10.1.1.1 as-path-filter 1 export
peer 10.1.1.1 ip-prefix prefix-name import
peer 10.1.1.1 ip-prefix prefix-name export
peer 10.1.1.1 route-policy policy-name import
peer 10.1.1.1 route-policy policy-name export
#
ipv4-family vpnv4
policy vpn-target
peer 10.1.1.1 enable
#
return
```

- 如果两端配置了出口/入口策略，则需要确认这些策略是否会把目标路由过滤掉，导致该路由无法正常收发。
- 如果两端没有配置相应的出口/入口策略，请检查路由超限问题。

路由数量超限导致收到的路由被丢弃

在路由接收端执行**display current-configuration configuration bgp | include peer destination-address**和**display current-configuration configuration bgp | include peer group-name**（如果Peer被加入到对等体组中）命令查看BGP配置，确认是否配置邻居路由限制。

例如，限制只能从邻居10.1.1.1收5条路由，超限之后将丢弃路由并记录日志。

```
<HUAWEI> display current-configuration configuration bgp | include peer 10.1.1.1
peer 10.1.1.1 as-number 100
peer 10.1.1.1 route-limit 5 alert-only
peer 10.1.1.1 enable
```

如果BGP邻居被加入到组中，显示信息中有可能没有route-limit的配置。

```
<HUAWEI> display current-configuration configuration bgp | include peer 10.1.1.1
peer 10.1.1.1 as-number 100
peer 10.1.1.1 group IBGP
peer 10.1.1.1 enable
peer 10.1.1.1 group IBGP
```

这种情况下，需要使用**display current-configuration configuration bgp | include peer group-name**来查看该对等体组的配置。

```
<HUAWEI> display current-configuration configuration bgp | include peer IBGP
peer IBGP route-limit 5 alert-only
peer IBGP enable
```

如果流量中断时，产生了路由超限告警BGP_1.3.6.1.4.1.2011.5.25.177.1.3.6 hwBgpPeerRouteExceed，表示路由超限导致目标路由被丢弃，则需要扩大本端的路由限制数值。

说明

修改BGP邻居限制的最大路由数量时会中断邻居，建议在路由发送端通过路由聚合以减少路由数量来解决。

标签超限导致私网路由无法发布

首先在路由发送端（本端PE）确认是否使能了mpls。然后，使用**display bgp vpnv4 all routing-table ipv4-address [mask | mask-length]**查看目标路由，确定该目标路由

由是否分到私网标签。如果显示信息中没有Label information字段，则可能是标签资源不足，导致无法为该路由申请到标签而不会给其它对等体。

```
<HUAWEI> display bgp vpnv4 all routing-table 10.10.1.1

BGP local router ID : 10.1.1.2
Local AS number : 100

Total routes of Route Distinguisher(1:1): 1
BGP routing table entry information of 10.10.1.0/24:
Imported route.
Label information (Received/Applied): NULL/13312
From: 0.0.0.0 (0.0.0.0)
Route Duration: 00h21m24s
Direct Out-interface: NULL0
Original nexthop: 0.0.0.0
Qos information : 0x0
Ext-Community:RT <1 : 1>
AS-path Nil, origin incomplete, MED 0, pref-val 0, valid, local, best, select, pre 255
Advertised to such 1 peers:
    10.1.1.1

Total routes of vpn-instance vpna: 1
BGP routing table entry information of 10.10.1.0/24:
Imported route.
From: 0.0.0.0 (0.0.0.0)
Route Duration: 00h21m24s
Direct Out-interface: NULL0
Original nexthop: 0.0.0.0
Qos information : 0x0
AS-path Nil, origin incomplete, MED 0, pref-val 0, valid, local, best, select, pre 60
Not advertised to any peer yet
```

- 如果是标签不足，可在VPN实例视图下通过命令**apply-label per-instance**配置每实例每标签，来减少标签的使用量。也可以通过路由聚合来减少路由数量。
- 如果标签没有超限，请检查路由是否超限。

ERT/IRT 不匹配导致路由无法交叉到私网路由表中

在路由的发送端（本端PE）/接收端（远端PE）执行**display current-configuration configuration vpn-instance**命令查看是否本端VPN实例的ERT与远端VPN实例的IRT不匹配，导致路由发送到远端PE后无法交叉到远端VPN实例中。

export-extcommunity表示ERT， **import-extcommunity**表示IRT。

```
<HUAWEI> display current-configuration configuration vpn-instance
#
ip vpn-instance vpna
 route-distinguisher 1:1
 apply-label per-instance
 vpn-target 1:1 export-extcommunity
 vpn-target 1:1 import-extcommunity
ip vpn-instance vpnb
 route-distinguisher 1:2
 vpn-target 1:1 export-extcommunity
 vpn-target 1:1 import-extcommunity
#
return
```

- 如果ERT和IRT不匹配，请在VPN实例下配置匹配的vpn-target。
- 如果ERT和IRT匹配，请检查标签是否超限。

对端（路由接收者）没有配置 peer 上的 ip-prefix 入口策略

在对端执行**display current-configuration configuration bgp**命令，检查显示信息中IPv4单播地址族中是否有**peer ipv4-address ip-prefix ip-prefix-name import**配置。

```
<HUAWEI> display current-configuration configuration bgp
#
bgp 100
peer 7.1.1.1 as-number 100
#
ipv4-family unicast
  peer 7.1.1.1 ip-prefix in import
peer 7.1.1.1 capability-advertise orf ip-prefix both
#
```

- 如果对端没有配置peer上的ip-prefix入口策略，则进入BGP IPv4单播地址族视图，执行**peer ipv4-address ip-prefix ip-prefix-name import**命令，在对端配置peer上的ip-prefix入口策略。

```
<HUAWEI> system-view
[~HUAWEI] bgp 100
[~HUAWEI-bgp] ipv4-family unicast
[~HUAWEI-bgp-af-ipv4] peer 7.1.1.1 ip-prefix in import
```

- 如果对端已经配置peer上的ip-prefix入口策略，但是本端仍然不能收到对端的ORF前缀信息，请检查对端入口策略前缀列表。

对端（路由接收者）没有配置 peer 上的 ip-prefix 入口策略对应的前缀列表

在对端执行**display ip ip-prefix ip-prefix-name**命令，检查对应BGP邻居ip-prefix入口策略的前缀列表是否配置。

```
<HUAWEI> display ip ip-prefix in
Info: The specified filter list does not exist.
```

出现上述提示信息，表示前缀列表in没有配置成功。请进入系统视图，执行**ip ip-prefix ip-prefix-name index index-number permit ipv4-address mask-length**命令配置前缀列表。

```
<HUAWEI> system-view
[~HUAWEI] ip ip-prefix in index 10 permit 10.1.1.0 24
```

配置完成后，在对端执行**display ip ip-prefix ip-prefix-name**命令，检查对应BGP邻居ip-prefix入口策略的前缀列表是否配置成功。

1.4 堆叠故障可能的原因是什么

堆叠配置失败、堆叠系统不能正常运行、设备的异常重启的原因通常涉及设备支持情况、设备连接与配置错误等。下面的内容有助于快速查找堆叠的故障原因。

1.4.1 堆叠配置失败的原因有哪些

1.4.2 堆叠成员异常重启的原因有哪些

1.4.3 堆叠无法主备倒换的原因有哪些

1.4.1 堆叠配置失败的原因有哪些

设备不支持堆叠、堆叠配置错误、硬件故障等均有可能造成堆叠配置失败和堆叠状态异常。下面列举了常见的故障原因以及相应的定位命令，更详细的故障处理案例可以参考[维护宝典](#)相关内容。

表 1-10 堆叠无法建立定位常用命令

可能原因	执行命令
设备不支持堆叠	display device
堆叠配置不正确	display stack configuration all
堆叠连线不正确	-
叠连接端口不Up	display interface brief
有堆叠故障事件信息	display stack troubleshooting display interface <i>interface-type interface-number</i> display stack statistics

设备不支持堆叠

使用命令**display device**查看设备型号，然后确认这些设备彼此之间是否支持组建堆叠。如果不支持，需要更换设备。

堆叠配置不正确

使用命令**display stack configuration all**检查堆叠配置是否符合要求。

```
<HUAWEI> display stack configuration all
Oper : Operation
Conf : Configuration
* : Offline configuration
Isolated Port: The port is in stack mode, but does not belong to any Stack-Port

Attribute Configuration:
-----
MemberID   Domain      Priority    Mode   Enable
Oper(Conf) Oper(Conf) Oper(Conf) Oper(Conf) Oper
-----
1(1)       10(10)      150(150)   MB(MB) Enable

Stack-Port Configuration:
-----
Stack-Port      Member Ports
-----
Stack-Port1/1   10GE1/1/0/1  10GE1/1/0/2
-----
```

显示信息“Oper(Conf)”里的“Oper”表示当前生效的配置，“Conf”表示下次启动生效的配置。

- 检查堆叠域（Domain）是否一致。组建堆叠的成员设备的堆叠域必须一致，否则无法组建堆叠。

可使用命令**stack member { *member-id* | all } domain *domain-id***修改堆叠域。

- （仅CE12800, CE12800E涉及）检查堆叠模式（Mode）是否一致。组建堆叠的成员设备的堆叠模式必须一致，否则无法组建堆叠。显示“MB”表示为默认的主控板直连方式，显示为“LC”表示为业务板直连方式。

可使用命令**stack member { *member-id* | all } link-type { *mainboard-direct* | *linecard-direct* }**修改堆叠模式。

- （仅CE12800, CE12800E, CE16800涉及）检查堆叠功能是否使能（Enable）。必须使能设备的堆叠功能才能组建堆叠。显示“Enable”表示堆叠已使能，显示“Disable”表示堆叠未使能。

可使用命令**stack enable**使能设备的堆叠功能。

- 检查是否存在离线配置（带“*”号的配置为离线配置），如果存在，则删除离线配置。离线配置可能会导致堆叠配置冲突，从而导致堆叠无法建立。

堆叠连线不正确

根据前期的规划和配置，检查堆叠连线是否与规划一致。

叠连接端口不 Up

使用命令**display interface brief**检查用于堆叠连接的端口的物理状态是否Up，包括堆叠物理成员端口、SIP口（仅CE12800, CE12800E, CE16800涉及）。如果物理状态为Down，则检查光模块、光纤是否有问题。可参考[接口不Up](#)解决相关问题。

有堆叠故障事件信息

使用命令**display stack troubleshooting**检查是否有堆叠故障事件信息。该命令可以记录堆叠组建过程中发生的部分故障，包括配置错误、连线错误等。可以根据故障事件的详细描述进行相应的故障处理。

```
<HUAWEI> display stack troubleshooting current
Total :1
-----
Seq  Time                Event Description
-----
1   2012-11-23 19:28:23.889  The devices belong to different stack domains,
                                and stack cannot be established. (MemberID = 1,
                                DomainID = 10, PeerMemberID = 2, PeerDomainID = 20)
```

说明

在V200R019C00版本及之前版本，查看堆叠系统当前产生的故障，请执行命令**display stack troubleshooting**。

在V200R005C20版本、V200R019C10版本及之后版本，查看堆叠系统当前产生的故障，请执行命令**display stack troubleshooting current**。

表 1-11 常见堆叠故障及处理步骤

描述	含义	处理步骤
The devices belong to different stack domains.	堆叠成员设备间堆叠域（Domain ID）不一致。	1. 使用命令 display stack configuration 检查各成员设备的堆叠域ID。 2. 使用命令 stack member member-id domain domain-id 修改堆叠域ID，使各成员设备的堆叠域ID一致。
The type of devices is different.	堆叠成员设备类型不同。	使用同一类型的设备来组建堆叠。

描述	含义	处理步骤
Stack-Port link invalid.	堆叠逻辑端口有错误连线。	<ol style="list-style-type: none"> 1. 检查堆叠逻辑端口是否存在错误连线，例如本端一个逻辑端口连接对端两个逻辑端口，或者逻辑端口内有环路。 2. 若有错误连线，则拆除错误连线并重新正确连接。
Switches working in different forward modes cannot set up a CSS.	两台设备的单板互通模式不同，不能建立堆叠。	<ol style="list-style-type: none"> 1. 执行命令display forward capability查看设备当前单板互通模式。 2. 如果不同，执行命令[undo] set forward capability enhanced设置两台设备为相同模式。设置后需要重启设备。
The interface Stack-Port is down.	堆叠逻辑端口协议 Down。	<ol style="list-style-type: none"> 1. 使用命令display interface brief检查堆叠逻辑端口物理是否 Down，如果是则检查连线是否松动或掉落。 2. 检查堆叠逻辑端口下是否有shutdown配置。
The physical status of the stack member port is up, but the protocol status is down.	堆叠物理成员端口协议 Down，但物理Up。	<ol style="list-style-type: none"> 1. 检查堆叠连线是否符合预期。如果不是则修改堆叠连线。 2. 使用命令display interface interface-type interface-number检查端口是否存在错包。如果是则检查两端的光模块是否插紧或者光模块是否有故障。 3. 在诊断视图下多次执行命令display stack statistics查看堆叠协议收、发报文是否异常。如果是，则排查堆叠配置或连线。

描述	含义	处理步骤
Configuration conflict.	堆叠成员交换机与主交换机间配置有冲突。	该故障通常是由于主交换机上有其他成员交换机的离线堆叠端口配置而导致的。 可以在主交换机上执行命令 display current-configuration all 检查离线配置，然后删除冲突的离线配置。
The port on CE-FWA board did not support configured as stack port.	单板上的端口不能用来堆叠。	使用支持堆叠的单板来组建堆叠。

1.4.2 堆叠成员异常重启的原因有哪些

堆叠竞争、软硬件故障均有可能造成堆叠成员异常重启。下面列出了快速查找命令，更详细的故障处理案例可以参考[维护宝典](#)相关内容。

表 1-12 堆叠成员异常重启定位常用命令

可能的问题	执行命令
堆叠竞争导致设备复位	display trapbuffer
软、硬件故障导致异常重启	display device board reset slot-id

堆叠竞争导致设备复位

除去初始组建堆叠过程的竞争场景，堆叠竞争一般是由于堆叠先分裂，然后再合并导致的。堆叠分裂通常会产生分裂告警DAD_1.3.6.1.4.1.2011.5.25.246.1.1 hwDadConflictDetect，或者成员设备离开告警STACKMNG_1.3.6.1.4.1.2011.5.25.183.1.22.24 hwStackMemberLeave。告警可以通过命令**display trapbuffer**查看。

```
<HUAWEI> display trapbuffer
Trapping buffer configuration and contents : enabled
Allowed max buffer size : 1024
Actual buffer size : 256
Channel number : 3, Channel name : trapbuffer
Dropped messages : 0
Overwritten messages : 0
Current messages : 129

Aug 15 2011 14:32:35 HUAWEI %%01DAD/1/hwDadConflictDetect(t):CID=0x807f0419-
OID=1.3.6.1.4.1.2011.5.25.246.1.1;Dual-active scenario is
detected.
```

如果是堆叠先分裂后合并导致的复位，需要检查分裂原因，一般是由于端口故障导致。确认堆叠分裂时间，然后查看该时间是否存在端口Down告警，如果存在端口Down告警，可以参考[接口不Up](#)定位并处理端口故障。

软、硬件故障导致异常重启

在诊断视图下通过命令**display device board reset slot-id**查看设备重启原因。

```
<HUAWEI> system-view
[~HUAWEI] diagnose
[~HUAWEI-diagnose] display device board reset 4
Board 4 reset information:
-- 1. DATE:2013-10-28 TIME:16:11:34 BARCODE:NULL RESET Num:1
-- Reason:Product unknown reason(CPU Reset)
-- BootMode:NORMAL
-- BootCode:0x060100ff
```

与堆叠相关的复位码（BootCode）范围是：0x80000020～0x80000038。

根据查询到的重启原因，参照表 [单板重启原因及建议操作](#) 所示的建议操作进行处理。

表 1-13 单板重启原因及建议操作

类型	重启原因描述	重启原因说明	建议操作
用户相关	Reset board from command.	命令行或网管重启。	确认是否因命令行重启或下电单板。
	Power off the board.		
	Reset board from PIC command.		
	Get pic offline message, and power off.		
	Canbus request to power off the board.		
	Power off board from command.		
环境相关	board cold reset(COLD Reset)	设备掉电。	检查设备的供电环境，确保供电正常。
加载相关	EPLD is upgrade,and reset board.	EPLD逻辑加载后重启。	正常。
	Board update by JTAG, and reset board.	通过监控系统 jtag通道升级后重启。	
	Board update mbus, and reset board.	升级mbus后重启。	
软件异常相关	Board task exception occurs and reset lpu.	检测到软件异常重启。	请收集日志、诊断日志和诊断命令 display reset-snapshot time 的信息，联系技术支持人员处理。

类型	重启原因描述	重启原因说明	建议操作
	Board task deadlock occurs and reset lpu.	检测到死循环重启。	请收集日志、诊断日志和诊断命令 display reset-snapshot time 的信息，联系技术支持人员处理。
	Board no enough Memory,and reset board.	内存耗尽重启。	1. 检查内存是否过高。 2. 收集内存信息、日志和诊断日志信息，联系技术支持人员处理。
	Board Memory overload,and reset board.		
	Component report failure.	组件上报失败。	通过诊断命令 display reportfailure number verbose slot slot-id 查看具体的原因，并请收集日志、诊断日志和诊断命令 display reset-snapshot time 的信息，联系技术支持人员处理。
	LAN Switch parity/ecc error, and reset board.	芯片发生了软失效。	请收集日志、诊断日志和诊断命令 display reset-snapshot time 的信息，联系技术支持人员处理。
设备管理相关	The heartbeat lost and reset lpu.	主控板未收到单板心跳报文。	请收集日志、诊断日志和诊断命令 display reset-snapshot time 的信息，联系技术支持人员处理。
	Semls register failed, and reset board.	检测到单板长时间未成功注册。	收集故障单板在启动过程中的串口打印信息，联系技术支持人员处理。
硬件器件相关	Board selftest error,and reset board.	自检错误。	器件自检失败，尝试拔插单板和更换槽位。如不能解决可判断为单板故障，请联系技术支持人员处理。

在任意视图通过命令**display device alarm hardware**、**display alarm active**查看有无硬件异常，如果有，则排除相应的硬件故障。

1.4.3 堆叠无法主备倒换的原因有哪些

堆叠无法主备倒换常常与交换机的状态有关。下面列出了可能的问题及对应的快速查找命令，更详细的故障处理案例可以参考[维护宝典](#)相关内容。

表 1-14 堆叠无法主备倒换定位常用命令

可能的问题	执行命令
框式设备上没有两块已注册成功的主控板	display device
主、备交换机之间批量备份是否已完成	display switchover state

框式设备上没有两块已注册成功的主控板

执行命令**display device**，查看主交换机上是否有两块已注册成功的主控板。如果主交换机只有一块注册成功的主控板，则主备倒换无法执行（仅针对CE12800, CE12800E, CE16800）。

```
<HUAWEI> display device
CE12804's Device status:
```

Slot	Card	Type	Online	Power	Register	Alarm	

Primary							

1	-	CE-L24LQ-EA	Present	On	Registered	Normal	NA
4	-	CE-L48XS-EF	Present	On	Registered	Normal	NA
6	-	CE-MPUA	Present	On	Registered	Normal	
7	-	CE-CMUA	Present	On	Registered	Normal	
8	-	CE-CMUA	Present	On	Registered	Normal	
9	-	CE-SFU04C	Present	On	Registered	Normal	NA
10	-	CE-SFU04C	Present	On	Registered	Normal	NA
11	-	CE-SFU04B	Present	On	Registered	Normal	NA
PWR1	-	PAC-2700WA	Present	On	Registered	Normal	NA
PWR3	-	PAC-2700WA	Present	Off	Registered	Abnormal	NA
FAN1	-	FAN-12C	Present	On	Registered	Normal	NA
FAN2	-	FAN-12C	Present	On	Registered	Normal	NA
FAN3	-	FAN-12C	Present	On	Registered	Normal	NA
FAN4	-	FAN-12C	Present	On	Registered	Normal	NA
FAN5	-	FAN-12C	Present	On	Registered	Normal	NA
FAN6	-	FAN-12C	Present	On	Registered	Normal	NA
FAN7	-	FAN-12C	Present	On	Registered	Normal	NA
FAN8	-	FAN-12C	Present	On	Registered	Normal	NA

```
FAN9 - FAN-12C Present On Registered Normal NA
```

- 如果主交换机上只有一块主控板，请再插入一块主控板，以保证主交换机有两块主控板。
- 如果主交换机上有两块主控板，但是备用主控板没有注册，请重新拔插备用主控板，再次查看备用主控板是否注册。如果仍未注册，请更换新的备用主控板。
- 如果主交换机上已有两块已注册成功的主控板，请检查主备间的备份是否已经完成。

主、备交换机之间批量备份是否已完成

执行命令**display switchover state**，查看主、备交换机之间的备份状态。

```
<HUAWEI> display switchover state
Switchover State : No slave
Switchover Policy : Board Switchover
MainBoard       : 5
SlaveBoard      : NULL
```

只有当“Switchover State”显示为**Ready**时，系统才满足主备倒换的条件。

- 如果批量备份没有完成，请等待一段时间至批量备份完成后，再进行主备倒换。
- 如果批量备份已经完成，但是无法执行主备倒换，请联系技术支持人员。

1.5 VXLAN 故障可能的原因是什么

VXLAN丢包、VXLAN网络速度慢、EVPN流量不通是常见的VXLAN故障类型。我们总结了常见的VXLAN配置错误、VXLAN故障原因，并提供了相应的Display命令，帮助实现VXLAN的故障原因的快速查找。

1.5.1 VXLAN业务丢包的原因有哪些

1.5.2 VM业务异常的原因有哪些

1.5.3 VXLAN路由不通的原因有哪些

1.5.4 VM迁移时间过长的原因有哪些

1.5.1 VXLAN 业务丢包的原因有哪些

VXLAN业务丢包常常与路由配置错误、流量抑制等有关，下面列出了可能的问题及对应的快速查找命令，更详细的故障处理案例可以参考[维护宝典](#)相关内容。

表 1-15 VXLAN 业务丢包定位常用命令

可能原因	执行命令
路由配置错误	display ip routing-table display current-configuration interface Nve <i>nve-number</i>
存在MAC HASH冲突	display mac-address hash-conflict

可能原因	执行命令
BD下配置了未知单播抑制	display current-configuration configuration bridge-domain <i>bd-id</i>

路由配置错误

1. 登录业务受损设备，通过命令**display ip routing-table**查看设备路由表中目的地址和下一跳等信息是否符合预期，确认Overlay的路由信息正确。如果不正确，请修改相关配置。
2. 查看各VTEP的NVE MAC地址是否存在冲突。

通过命令**display current-configuration interface Nve *nve-number***查看各VTEP的NVE MAC地址。

```
<HUAWEI> display current-configuration interface Nve 1
#
interface Nve1
 source 1.1.1.1
 vni 5010 head-end peer-list protocol bgp
 mac-address 0000-5e00-0106
#
return
```

如果NVE MAC地址存在相同的情况，请修改相关配置。

存在 MAC HASH 冲突

通过命令**display mac-address hash-conflict**查看由于哈希冲突而无法使用的MAC地址。通过**Conflicting MAC Address**字段可以看到有冲突的MAC地址，其中带“*”号表示当前时刻冲突的MAC地址，不带“*”号表示历史冲突的MAC地址。

```
<HUAWEI> display mac-address hash-conflict 0010-1100-3710 bridge-domain 10
Flags: * - Current MAC address in the hash bucket of the chip
      _ - Internal bridge domain resource
BD : bridge-domain
```

```
-----
Slot: 1      Chip: 0
-----
MAC Address  VLAN/BID      Conflicting MAC Address  Time
-----
0010-1100-3710  -/10      0010-1100-7bd2*        -
                0010-1100-6bd7*        -
                0010-1100-08a7*        -
                0010-1100-20c9*        -
                0010-1100-7bd2        2018-02-01 14:32:23
                0010-1100-6bd7        2018-02-01 14:32:23
                0010-1100-08a7        2018-02-01 14:32:23
                0010-1100-20c9        2018-02-01 14:32:23
-----
```

如果使用的MAC地址存在HASH冲突，请更换相关的MAC地址。

BD 下配置了未知单播抑制

通过命令**display current-configuration configuration bridge-domain *bd-id***查看BD下的配置，存在未知单播抑制的配置**storm suppression unknown-unicast cir**。

```
<HUAWEI> display current-configuration configuration bridge-domain 10
#
bridge-domain 10
```

```
... ..
storm suppression unknown-unicast cir 1000 kbps cbs 1048576 bytes
... ..
#
return
```

当存在MAC HASH冲突，同时又配置了未知单播限速，MAC学习失败后业务流量被当做未知单播，超过限制的部分被丢弃，从而出现丢包。

1.5.2 VM 业务异常的原因有哪些

VM业务异常通常与MAC地址有关，下面给出了快速查找命令，更详细的故障处理案例可以参考[维护宝典](#)相关内容。

表 1-16 VM 业务异常定位常用命令

可能的原因	执行命令
MAC地址漂移	<code>display mac-address</code> <code>display mac-address flapping</code>

MAC 地址漂移

1. 通过命令**display mac-address**查看MAC地址表，确认VM的MAC地址。

```
<HUAWEI> display mac-address
Flags: * - Backup
      # - forwarding logical interface, operations cannot be performed based
      on the interface.
BD   : bridge-domain   Age : dynamic MAC learned time in seconds
-----
MAC Address  VLAN/VSI/BD  Learned-From  Type      Age
-----
00e0-fc00-0001 -/-/100    10.1.1.1      evn        -
0000-0000-0002 -/-/100    10.1.1.1      evn        -
-----
Total items: 2
```

2. 通过命令**display mac-address flapping**查看是否存在对应的MAC漂移记录。

```
<HUAWEI> display mac-address flapping
MAC Address Flapping Configurations :
-----
Flapping detection      : Enable
Aging time(s)           : 300
Quit-VLAN Recover time(m) : --
Exclude VLAN-list       : --
Security level           : Middle
Exclude BD-list          : --
-----
S: start time   E: end time   (D): error down
-----
Time           : S:2018-10-24 15:46:04      E:2018-10-24 15:47:32
VLAN/BD        : -/100
MAC Address    : 00e0-fc00-0001
Original-Port  : Eth-Trunk1.3
Move-Ports     : Eth-Trunk10.2
MoveNum        : 1348
-----
Total items on slot 1: 1
```

说明

如果大量MAC发生漂移，考虑为环路问题，请移除新增连线或者关闭对应接口。

- 通过命令**display interface vbdif 100**查看VBDIF100的MAC地址。通过以下显示信息可以看出，VBDIF的MAC地址也为**00e0-fc00-0001**。

```
<HUAWEI> display interface vbdif 100
Vbdif100 current state : UP (ifindex: 149)
Line protocol current state : UP
Last line protocol up time : 2018-10-24 15:55:30
Description:
Route Port,The Maximum Transmit Unit is 1500
Internet Address is 1.1.1.1/24
IP Sending Frames' Format is PKTFMT_ETHNT_2, Hardware address is 00e0-fc00-0001
Current system time: 2018-10-24 15:55:32
```

问题确认，VM的MAC与网关MAC发生冲突，导致服务器业务访问异常。

- 修改VM或网关的MAC地址，保证网络中没有MAC地址冲突。如果网关不使用，建议删除对应的VBDIF接口。

```
<HUAWEI> system-view
[~HUAWEI] undo interface vbdif 100
[*HUAWEI] commit
```

1.5.3 VXLAN 路由不通的原因有哪些

License失效、隧道未建立成功、转发表项错误等均有可能造成VXLAN路由不通，下面给出了快速查找命令，更详细的故障处理案例可以参考[维护宝典](#)相关内容。

表 1-17 VXLAN 流量不通定位常用命令

可能的问题	执行命令
License是否有效	display license
隧道模式是否为VXLAN	display fei frame resource
VNI是否Up	display vxlan vni
隧道是否UP	display vxlan tunnel
隧道SIP的路由是否存在	display vxlan tunnel display ip routing-table
隧道DIP的路由是否存在	display vxlan tunnel display ip routing-table
接口板FIB表是否存在	display vxlan tunnel display ip fib display alarm active
接口板ARP表是否存在	display fei ipv4 arp display alarm active

License 是否有效

（仅CE8800&7800&6800&5800系列交换机需要执行此步骤）任意视图下，执行命令**display license**，查看设备上的License信息。

```
<HUAWEI> display license
MainBoard:
Active License : CloudEngine7800V200R001_20160621145931511.dat
```

```
License state   : Trial
Revoke ticket  : No ticket

RD of Huawei Technologies Co., Ltd.
Product name   : CloudEngine 7800
Product version : V200R001
License Serial No : LIC20160621660650
Creator        : Huawei Technologies Co., Ltd.
Created Time   : 2016-06-21 14:59:30
Feature name   : CELIC
Authorize type  : demo
Expired date    : 2016-08-05
Trial days     : 60 Item name

Item type      Value      Description
-----
CE-LIC-VXLAN   Function   YES      CE-LIC-VXLAN
CE-LIC-FCF-ALL Function   YES      CE-LIC-FCF-ALL
CE-LIC-NPV     Resource  YES      CE-LIC-NPV
CE-LIC-FCF-PORT Resource  16      CE-LIC-FCF-PORT
License state: Demo. The license for the current configuration will expire in 3 day(s).
Apply for authentic license before the current license expires.
```

检查当前设备上的License是否有效。如果未加载License、License已过期或者License不支持VXLAN，则需要安装包含VXLAN功能的License。

隧道模式是否为 VXLAN

（仅CE12800系列交换机需要执行此步骤）仅CE12800系列交换机需要执行此步骤。诊断视图下，执行命令**display fei frame resource**，查看隧道模式。

```
<HUAWEI> system-view
[~HUAWEI] diagnose
[~HUAWEI-diagnose] display fei frame resource

Slot          Unit CurrentMode  NextMode
-----
1              0  VXLAN        VXLAN
4              0  VXLAN        VXLAN
```

- 如果CurrentMod（当前资源模式）不是VXLAN模式，NextMode（下次启动生效的资源模式）是VXLAN模式，则需要重启设备使当前隧道模式切换为VXLAN模式。
- 如果CurrentMode不是VXLAN模式，NextMode模式也不是VXLAN模式，则需要通过命令行**ip tunnel mode vxlan**设置隧道模式为VXLAN模式，并重启设备生效。

VNI 是否 Up

任意视图下，执行命令**display vxlan vni**，查看VNI状态。

```
<HUAWEI> display vxlan vni
Number of vxlan vni : 3
VNI      BD-ID      State
-----
20        20        down
10000     100        up
200000    200        down
```

如果VNI状态为down，则需要检查当前设备的隧道相关配置。

隧道是否 UP

任意视图下，执行命令**display vxlan tunnel**，查看隧道状态。


```
<HUAWEI> display vxlan tunnel
Number of vxlan tunnel : 1
Tunnel ID   Source      Destination   State Type
-----
4026531846  1.1.1.1      2.2.2.2      up    static
```

如果VXLAN隧道状态为UP，则需要进一步查询接口板FIB表内容。

隧道 SIP 的路由是否存在

- 任意视图下，执行命令**display vxlan tunnel**，查看隧道对应的SIP。

```
<HUAWEI> display vxlan tunnel
Number of vxlan tunnel : 1
Tunnel ID   Source      Destination   State Type
-----
4026531846  1.1.1.1      2.2.2.2      up    static
```

其中Source即为VXLAN隧道的SIP。

- 任意视图下，执行命令**display ip routing-table**，查看隧道SIP的路由。

```
<HUAWEI> display ip routing-table 1.1.1.1
Proto: Protocol      Pre: Preference
Route Flags: R - relay, D - download to fib, T - to vpn-instance, B - black hole route
-----
Routing Table : _public_
Summary Count : 1

Destination/Mask  Proto  Pre  Cost    Flags NextHop      Interface
1.1.1.1/32      Direct  0    0        D  127.0.0.1    LoopBack1
```

查到隧道SIP的路由下一跳，一般为Loopback接口。如果查不到隧道SIP路由信息，则需要尝试ping隧道SIP。如果无法ping通，则需要查看链路问题；如果可以ping通，请检查隧道DIP的路由是否存在。

隧道 DIP 的路由是否存在

- 任意视图下，执行命令**display vxlan tunnel**，查看隧道对应的DIP。

```
<HUAWEI> display vxlan tunnel
Number of vxlan tunnel : 1
Tunnel ID   Source      Destination   State Type
-----
4026531846  1.1.1.1      2.2.2.2      up    static
```

其中Destination即为VXLAN隧道的DIP。

- 任意视图下，执行命令**display ip routing-table**，查看隧道DIP的路由。

```
<HUAWEI> display ip routing-table 2.2.2.2
Proto: Protocol      Pre: Preference
Route Flags: R - relay, D - download to fib, T - to vpn-instance, B - black hole route
-----
Routing Table : _public_
Summary Count : 1

Destination/Mask  Proto  Pre  Cost    Flags NextHop      Interface
2.2.2.2/32      Direct  0    0        D  192.168.1.1  10GE4/0/5
```

如果查不到隧道DIP路由信息，则需要尝试ping隧道DIP。如果无法ping通，则需要查看链路问题；如果可以ping通，请联系技术支持人员。

接口板 FIB 表是否存在

- 任意视图下，执行命令**display vxlan tunnel**，查看隧道对应的DIP。

```
<HUAWEI> display vxlan tunnel
Number of vxlan tunnel : 1
Tunnel ID   Source      Destination   State Type
-----
4026531846  1.1.1.1      2.2.2.2      up    static
```

其中Destination即为VXLAN隧道的DIP。

- 任意视图下，执行命令**display ip fib**，查看接口板上DIP对应的FIB表。

```
<HUAWEI> display ip fib slot 4 2.2.2.2
Route Flags: G - Gateway Route, H - Host Route, U - Up Route
              S - Static Route, D - Dynamic Route, B - Black Hole Route

-----
FIB Table: _public_
Total number of Routes: 3

Destination/Mask  Nexthop      Flag  Interface      TunnelID
2.2.2.2/32 192.168.1.1  DGU   10GE4/0/5      -
```

接口板上的FIB表中保存了VXLAN隧道的下一跳IP和出接口信息。如果查不到接口板FIB表内容，则需要进一步查询设备上是否存在资源不足的告警信息。

- 任意视图下，执行命令**display alarm active**，查看是否存在资源不足的告警信息。

```
<HUAWEI> display alarm active

-----
Sequence  AlarmId  Severity Date Time  Description
-----
77        0x95E2022  Major   2012-09-15 20:16:37 The number of forwarding resources rea
              ches the alarm threshold. (Slot = 4, T
              hreshold = 90, Reason = 174, Descripti
              on : The number of ARP exceeded the wa
              rning threshold.)
-----
```

如果存在资源不足的告警内容，请联系技术支持人员。

接口板 ARP 表是否存在

- 诊断视图下，执行命令**display fei ipv4 arp**，查看接口板上的ARP表。其中slot为当前设备在转发过程中使用的接口板，next-ip为接口板FIB表中的Nexthop。FIB表查看方法参考接口板FIB表是否存在。

```
<HUAWEI> system-view
[~HUAWEI] diagnose
[~HUAWEI-diagnose] display fei ipv4 arp slot 4 next-ip 192.168.1.1
ARP Table:
Total number: 1

-----
IPAddr: 192.168.1.1 IfIndex: 41981
VrIndex: 8 VrflIndex: 0 NhplIndex: 1676 EepIndex: 36940
MACAddr: 2013-2428-2650 FVrf: 0 PeVid: 0
CeVID: 0 IfType: 8 TargetBlade: 0
TargetPort: 0 TrunkId: 25
```

如果查不到接口板ARP表内容，则需要进一步查询设备上是否存在资源不足的告警信息。

- 任意视图下，执行命令**display alarm active**，查看是否存在资源不足的告警信息。

```
<HUAWEI> display alarm active

-----
Sequence  AlarmId  Severity Date Time  Description
-----
77        0x95E2022  Major   2012-09-15 20:16:37 The number of forwarding resources rea
              ches the alarm threshold. (Slot = 4, T
              hreshold = 90, Reason = 174, Descripti
              on : The number of ARP exceeded the wa
              rning threshold.)
-----
```

如果存在ARP资源不足的告警内容，请联系技术支持人员。

1.5.4 VM 迁移时间过长的原因有哪些

VM迁移时间过长主要是由相关参数配置不合理导致的，下面给出了快速查找命令，更详细的故障处理案例可以参考[维护宝典](#)相关内容。

表 1-18 VM 迁移时间过长定位常用命令

可能的问题	执行命令
EVPN的BGP打包组信息有误	<code>display bgp evpn update-peer-group</code>
查看对应组内EVPN路由的发送情况	<code>display bgp evpn update-peer-group index <i>number</i> verbose</code>
Server Leaf和Border Leaf的MAC地址老化时间不合理	<code>mac-address aging-time <i>aging-time</i></code>

EVPN 的 BGP 打包组信息有误

登录RR反射器（Spine），查看EVPN的BGP打包组信息。

```
<HUAWEI> display bgp evpn update-peer-group
```

```
The EVPN instance's update peer group number : 1
Keep buffer update peer group number : 0
BGP Version : 4
```

```
Group ID : 4
Group Type : internal
Addr Family : EVPN
AdvMinTimeVal : 15
Total Peers : 2
Leader Peer : 10.1.1.1
Peers List : 10.1.1.1
```

```
10.1.1.2
```

如果BGP组配置与规划不一致，请修改相关配置。

查看对应组内 EVPN 路由的发送情况

执行诊断命令**display bgp evpn update-peer-group index *index* verbose**。

```
<HUAWEI> system-view
```

```
[~HUAWEI] diagnose
```

```
[~HUAWEI-diagnose] display bgp evpn update-peer-group index 4 verbose
```

```
... ..
```

```
BufPacketCnt      : 6153
BufInUsingPeerCnt  : 0
BufferFlag         : 0   SequenceID      : 0
UptPeerGrp PktBuffer count : 100
realtime-advlist    : 0
batch-update-advlist : 0
batch-withdraw-advlist : 0
PassiveCleanupPeerCount : 0
---- More ----
```

UptPeerGrp PktBuffer count字段数据非0，表示RR路由发送未完成，从而可以判断路由频繁更新导致RR反射器发送队列出现拥塞。

Server Leaf 和 Border Leaf 的 MAC 地址老化时间不合理

登录Server Leaf和Border Leaf，在系统视图下，通过命令**mac-address aging-time aging-time**配置动态MAC地址表项的老化时间为30分钟（1800秒）。缺省情况下，动态MAC地址表项的老化时间为300秒。

```
<HUAWEI> system-view
[~HUAWEI] mac-address aging-time 1800
[*HUAWEI] commit
```

修改后再次查看VM迁移速度，如果问题仍未解决可参考[维护宝典](#)VXLAN故障相关内容。

1.6 相关信息

您可以根据查找到的故障原因，参考[维护宝典](#)、[故障启示录](#)、[产品手册](#)解决问题。

您可以使用[智能问答](#)服务来咨询相关问题的解决办法。