

包装技术说明的OSPF、MTU和LSA

目录

[简介](#)

[OSPF数据包大小](#)

[在DBD信息包的MTU](#)

[OSPF行为和包装LSA到LS更新数据包里](#)

[在Cisco Bug ID CSCse01519前](#)

[在Cisco Bug ID CSCse01519以后](#)

[Cisco Bug ID CSCse01519](#)

[概述](#)

[方案](#)

简介

本文描述开放最短路径优先(OSPF)数据包、最大转换单元(MTU)， Link State Advertisement (LSA)和林克状态(LS)更新数据包的交互作用在Cisco Bug ID [CSCse01519](#)中。

OSPF数据包大小

在路由器的链路有MTU。输出数据包，例如OSPF信息包，大于接口MTU不可以。

[请求注释\(RFC\) 2328](#) OSPF协议的文档版本2。附录RFC 2328 A.1如此描述OSPF信息包的封装：

OSPF直接地在互联网协议的网络层运行。因此OSPF信息包由IP和本地数据链路报头独自地封装。

OSPF大于网络MTU不定义了方式分段其协议信息包，并且依靠IP分段，当传送信息包。如果需要，长度OSPF信息包可以是65,535个字节(包括IP报头)。可能大的OSPF数据包类型(数据库描述信息包，连接状态请求，连接状态更新和林克状态确认数据包)可能通常拆分到几个独立的协议信息包，没有功能损耗。推荐这;应该避免IP分段若情况许可。

可以有在LS更新数据包的一个或更多LSA。在一LS更新数据包的许多LSA叫作包装LSA到LS更新数据包。

在DBD信息包的MTU

数据库描述(DBD)数据包，也指定在RFC 2328，描述OSPF链路状态数据库的内容：


```
OSPF: Retransmitting DBD to 10.100.1.2 on GigabitEthernet0/1 [11]
%OSPF-5-ADJCHG: Process 1, Nbr 10.100.1.2 on GigabitEthernet0/1 from EXSTART to
DOWN, Neighbor Down: Too many retransmissions
```

OSPF行为和包装LSA到LS更新数据包里

在Cisco Bug ID CSCse01519前

不管接口MTU，在Cisco Bug ID [CSCse01519前](#)，OSPF在Cisco IOS软件里构件OSPF信息包没有更加大的than1500字节。因此，如果接口MTU大于1500个字节，OSPF只仍然包装了1500个字节到OSPF数据包。因为OSPF可能发送在链路的更加大的数据包和达到更加了不起的吞吐量，这是有些效率低的。

注意：有一例外对此方案。如果一个LSA保持超过1500个字节，OSPF构件该数据包，没有问题大小，因为OSPF不能分段一个LSA。路由器的IP协议栈然后分段数据包为了适合流出接口的MTU。这典型地发生，当OSPF路由器有许多链路，并且路由器LSA大于链路MTU变得。

同样地，如果流出接口的MTU小于1500个字节，OSPF程序仍然构件或包装了OSPF信息包1500个字节，并且路由器的IP协议栈分段数据包到更加小的IP信息包为了适合流出的链路的MTU。这典型地发生在运行OSPF的两路由器之间的一个IPSec隧道。通道的封装字节的已添加开销导致了小于1500个字节的MTU。OSPF构件OSPF信息包至1500个字节，并且数据包在路由器前然后被分段传送他们。这是另外的无效用。

在Cisco Bug ID CSCse01519以后

在Cisco Bug ID [CSCse01519以后](#)，在IOS的OSPF大于1500个字节能包装OSPF信息包。如果流出接口的MTU大于1500个字节，这发生。因为更多信息可以被包装到一更加大的数据包，发射是更有效的。换句话说，如果一个OSPF路由器需要传达许多外部LSA给OSPF邻居，它能包装更多外部LSA到一LS更新数据包，如果该路由器运行IOS以CSCse01519实现的Cisco Bug ID。

Cisco Bug ID CSCse01519也允许OSPF建立数据包小于1500个字节。在某些情况下，在两个OSPF邻居之间的MTU小于1500个字节。在前一个示例中用IPSec隧道，OSPF传送小于的OSPF信息包1500个字节并且避免IP分段;再次，例外是大于接口MTU LSA的事例。

Cisco Bug ID CSCse01519

当您升级OSPF路由器时，您可以发现Cisco Bug ID导致的OSPF MTU问题[CSCse01519](#)。

概述

许多网络有通过第2层网络连接的OSPF邻居，或者传输网络，由L2VPN服务或同步数字分级体系/Synchronous光纤网络(SDH/SONET)网络组成。这些传输网络比运行OSPF的路由器能有不同的MTU设置。

虽然MTU设置应该是正确在所有路由器，并且应该反射真的MTU，经常有未被注意的错误。

这是一个示例网络用运行OSPF的两路由器。路由器1 (R1)和路由器2 (R2)通过L2交换机连接。

在本例中，路由器有千兆以太网接口MTU设置到2000年。L2交换机的MTU只是1500个字节。

如果数据流的大小大于1500个字节从未，您能使用IOS，不用Cisco Bug ID [CSCse01519](#)，因为OSPF信息包大于1500个字节从未。然而，如果有是1800个字节，例如，在R1或R2的OSPF程序建立LS更新数据包大于1500个字节并且传送它的LSA，但是数据包由L2交换机丢弃在路由器之间。

如果在R2的OSPF数据库有足够的网络，本地产生的LSA很大LS更新数据包大于接口MTU也许。

- 如果这些网络由覆盖物网络命令产生，网络在R2路由器LSA出现。R2建立大于2000个字节并且传送它的路由器LSA，但是IP段它对2000个字节，接口MTU。然而L2交换机丢弃这些数据包。OSPF不断地然后重传此数据包，并且OSPF邻接状态从未全双工。因此，既使当您运行IOS，不用Cisco Bug ID CSCse01519，问题立即发现。
- 如果这些网络由**redistribute connected**命令产生，网络在外部LSA出现。OSPF设法包装外部LSA到大小上是1500个字节的一LS更新数据包。在这种情况下，因为接口MTU是2000个字节，OSPF邻接到达‘FULL’状态。不适于的基础MTU的问题不是立即发现。当一个路由器升级对与Cisco Bug ID CSCse01519的IOS问题将发现。

方案

假设，两路由器运行IOS版本，不用Cisco Bug ID [CSCse01519](#)。

当OSFP邻接修造，注意R1从未接收OSPF数据包大于1500个字节，虽然接口的MTU是2000年。

启用**packets**命令**debug ip的ospf**。

```
OSPF: rcv. v:2 t:1 l:48 rid:10.100.1.2
      aid:0.0.0.0 chk:72CF aut:0 auk: from GigabitEthernet0/1
...
OSPF: rcv. v:2 t:4 l:1468 rid:10.100.1.2
      aid:0.0.0.0 chk:8389 aut:0 auk: from GigabitEthernet0/1
OSPF: rcv. v:2 t:4 l:136 rid:10.100.1.2
...
```

在此debug输出中，'l:1468是OSPF数据包的长度，因此您能看到最大的OSPF数据包是1468个字节。't:4表明OSPF数据包是类型4，是林克状态更新信息包。从RFC 2328的第4.3部分的此表定义了不同的OSPF数据包类型：

类型	数据包名称	协议功能
1	Hello	发现/保养邻居
2	数据库描述	汇总数据库内容
3	林克状态请求	数据库下载
4	林克状态更新	数据库更新
5	林克状态Ack	充斥确认

OSPF邻接到达‘FULL’状态。

```
R1#show ip ospf neighbor gigabitEthernet 0/1
```

Neighbor ID	Pri	State	Dead Time	Address	Interface
10.100.1.2	0	FULL/ -	00:00:34	10.1.1.2	GigabitEthernet0/1

```
R2#show ip ospf neighbor gigabitEthernet 0/1
```

Neighbor ID	Pri	State	Dead Time	Address	Interface
10.100.100.1	0	FULL/ -	00:00:34	10.1.1.1	GigabitEthernet0/1

其次，在R2to的升级IOS与Cisco Bug ID CSCse01519的一个IOS版本。

```
R2#show ip ospf neighbor gigabitEthernet 0/1
```

Neighbor ID	Pri	State	Dead Time	Address	Interface
10.100.100.1	0	LOADING/ -	00:00:33	10.1.1.1	GigabitEthernet0/1

```
R2#show ip ospf neighbor gigabitEthernet 0/1 detail
```

```
Neighbor 10.100.100.1, interface address 10.1.1.1
  In the area 0 via interface GigabitEthernet0/1
  Neighbor priority is 0, State is LOADING, 5 state changes
  DR is 0.0.0.0 BDR is 0.0.0.0
  Options is 0x12 in Hello (E-bit L-bit )
  Options is 0x52 in DBD (E-bit L-bit O-bit)
  LLS Options is 0x1 (LR)
  Dead timer due in 00:00:39
  Neighbor is up for 00:00:49
  Index 1/1, retransmission queue length 0, number of retransmission 0
  First 0x0(0)/0x0(0) Next 0x0(0)/0x0(0)
  Last retransmission scan length is 0, maximum is 0
  Last retransmission scan time is 0 msec, maximum is 0 msec
Number of retransmissions for last link state request packet 9
  Poll due in 00:00:00
```

```
R2#show ip ospf neighbor gigabitEthernet 0/1 detail
```

```
Neighbor 10.100.100.1, interface address 10.1.1.1
  In the area 0 via interface GigabitEthernet0/1
  Neighbor priority is 0, State is LOADING, 5 state changes
  DR is 0.0.0.0 BDR is 0.0.0.0
  Options is 0x12 in Hello (E-bit L-bit )
  Options is 0x52 in DBD (E-bit L-bit O-bit)
  LLS Options is 0x1 (LR)
  Dead timer due in 00:00:33
  Neighbor is up for 00:02:06
  Index 1/1, retransmission queue length 0, number of retransmission 0
  First 0x0(0)/0x0(0) Next 0x0(0)/0x0(0)
  Last retransmission scan length is 0, maximum is 0
  Last retransmission scan time is 0 msec, maximum is 0 msec
Number of retransmissions for last link state request packet 25
  Poll due in 00:00:03
```

```
%OSPF-5-ADJCHG: Process 1, Nbr 10.100.100.1 on GigabitEthernet0/1 from LOADING
to DOWN, Neighbor Down: Too many retransmissions
```

OSPF邻接在‘加载’状态被滞留，并且不到达‘FULL’状态。重新传输发生，直到OSPF达到25重新传输其限制。OSPF设法再设立邻接，同一个问题再发生，并且环路不断地继续。

因此，在R2的升级找到一个以前隐藏的问题：基础MTU小于OSPF路由器使用的那个。

当交换机更改MTU到2000年时，OSPF数据包大于1500个字节(!:1980)传送没有问题。

```
R1#
```

```
OSPF: rcv. v:2 t:3 1:1980 rid:10.100.1.2
  aid:0.0.0.0 chk:AC5B aut:0 auk: from GigabitEthernet0/1
```

为了检查强调的MTU问题，总是请ping与大小的OSPF邻居IP地址相等与MTU和DF (不要分段)位集

。

为了发现基础MTU的值，执行ping和清扫大小。计数感叹号数量(!)在输出中为了确定正确MTU。在本例中，从ping命令的最后echo replies有大小1500字节。

```
R2#ping
Protocol [ip]:
Target IP address: 10.1.1.1
Repeat count [5]: 1
Datagram size [100]:
Timeout in seconds [2]:
Extended commands [n]: yes
Source address or interface:
Type of service [0]:
Set DF bit in IP header? [no]: yes
Validate reply data? [no]:
Data pattern [0xABCD]:
Loose, Strict, Record, Timestamp, Verbose[none]:
Sweep range of sizes [n]: yes
Sweep min size [36]: 1460
Sweep max size [18024]: 1540
Sweep interval [1]:
Type escape sequence to abort.
Sending 81, [1460..1540]-byte ICMP Echos to 10.1.1.1, timeout is 2 seconds:
Packet sent with the DF bit set
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
.....
Success rate is 49 percent (40/81), round-trip min/avg/max = 1/1/4 ms
```