Nota técnica de paquetes OSPF, MTU y LSA

Contenido

Introducción

Este documento describe la interacción de paquetes Open Shortest Path First (OSPF), unidad de transición máxima (MTU), anuncios de estado de link (LSA) y paquetes de actualización de estado de link (LS) en el contexto de la ID de bug de Cisco CSCse01519.

Tamaño de paquete OSPF

Los links en los routers tienen una MTU. Los paquetes salientes, como los paquetes OSPF, no pueden ser mayores que la MTU de interfaz.

La solicitud de comentarios (RFC) 2328 documenta la versión 2 del protocolo OSPF. El Apéndice A.1 de RFC 2328 describe la Encapsulación de los paquetes OSPF de esta manera:

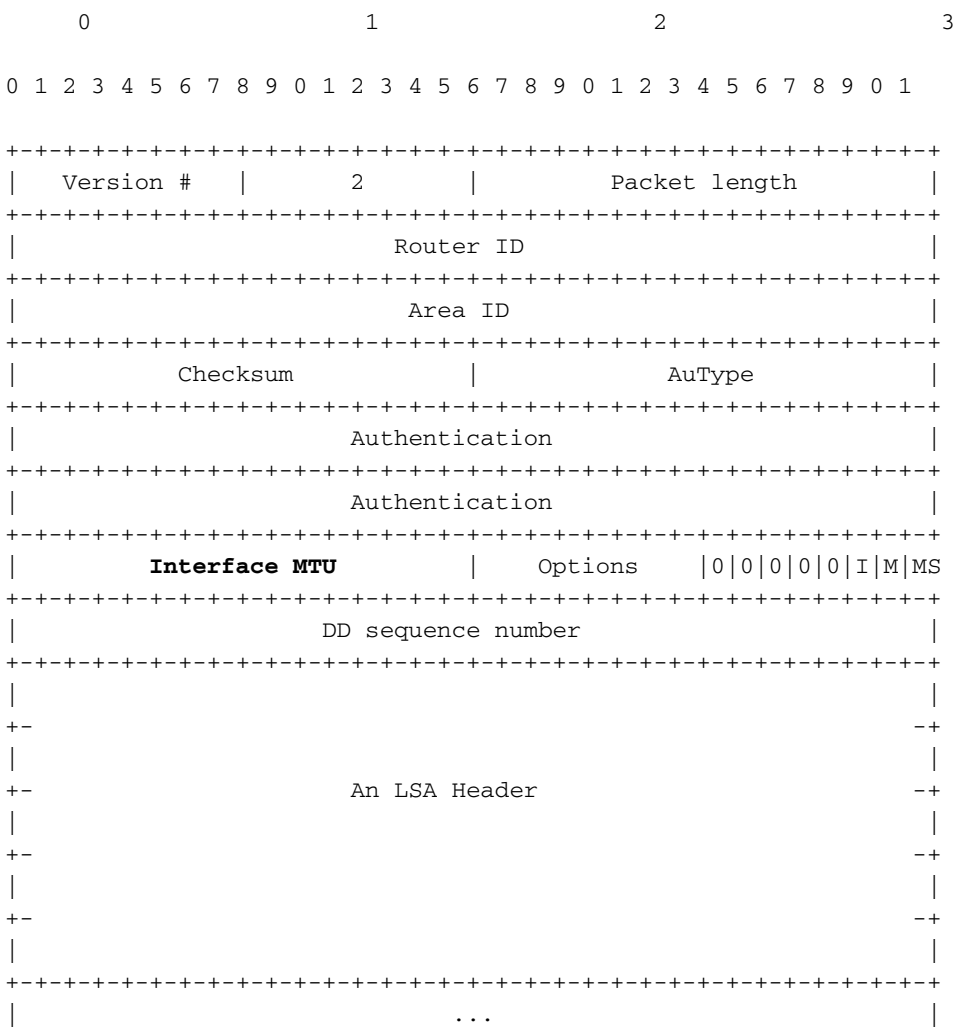
OSPF se ejecuta directamente en la capa de red del Protocolo de Internet. Por lo tanto, los paquetes OSPF son encapsulados únicamente por los encabezados de link de datos local e IP.

OSPF no define una manera de fragmentar sus paquetes de protocolo y depende de la fragmentación IP cuando se transmiten paquetes mayores que la MTU de red. Si es necesario, la longitud de los paquetes OSPF puede ser de hasta 65.535 bytes (incluido el encabezado IP). Los tipos de paquetes OSPF que probablemente sean grandes (paquetes de descripción de la base de datos, solicitud de estado de link, actualización de estado de link y paquetes de reconocimiento de estado de link) normalmente se pueden dividir en varios paquetes de protocolo independientes, sin pérdida de funcionalidad. Esto se recomienda; Siempre que sea posible, debe evitarse la fragmentación IP.

Puede haber uno o más LSA en un paquete LS Update. Muchos LSA en un paquete LS Update se conocen como paquetes LSA en un paquete LS Update.

MTU en paquete DBD

El paquete Descripción de la Base de Datos (DBD), también especificado en RFC 2328, describe el contenido de la base de datos de estado de link OSPF:



Apéndice A.3.3. de RFC 2328 describe la interfaz MTU como:

El tamaño en bytes del datagrama IP más grande que se puede enviar fuera de la interfaz asociada, sin fragmentación.

Los routers que están conectados a un link intercambian su valor de MTU de interfaz en paquetes DBD cuando se inicializa la adyacencia OSPF.

La sección 10.6 de RFC 2328 establece:

Si el campo MTU de la interfaz en el paquete Descripción de la base de datos indica un tamaño de datagrama IP mayor que el que puede aceptar el router en la interfaz de recepción sin fragmentación, se rechaza el paquete Descripción de la base de datos.

Cuando se utiliza el comando **debug ip ospf adj**, puede ver la llegada de estos paquetes DBD.

En este ejemplo, hay una discordancia en los valores de MTU entre dos vecinos OSPF. Este router tiene MTU 1600:

```
OSPF: Rcv DBD from 10.100.1.2 on GigabitEthernet0/1 seq 0x2124 opt 0x52 flag 0x2
len 1452 mtu 2000 state EXSTART
OSPF: Nbr 10.100.1.2 has larger interface MTU
```

El otro router OSPF tiene la interfaz MTU 2000:

```
OSPF: Rcv DBD from 10.100.100.1 on GigabitEthernet0/1 seq 0x89E opt 0x52 flag 0x7
len 32 mtu 1600 state EXCHANGE
OSPF: Nbr 10.100.100.1 has smaller interface MTU
```

Los paquetes DBD se retransmiten de forma continua hasta que la adyacencia OSPF finalmente se derribe.

```
OSPF: Send DBD to 10.100.1.2 on GigabitEthernet0/1 seq 0x9E6 opt 0x52 flag 0x7
len 32
OSPF: Retransmitting DBD to 10.100.1.2 on GigabitEthernet0/1 [10]
OSPF: Send DBD to 10.100.1.2 on GigabitEthernet0/1 seq 0x9E6 opt 0x52 flag 0x7
len 32
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
```

Comportamiento OSPF y empaquetado de LSA en un Paquete de Actualización LS

Antes del ID de bug de Cisco CSCse01519

Antes del Id. de bug Cisco CSCse01519, OSPF en el software Cisco IOS® generó paquetes OSPF no mayores de 1500 bytes, independientemente de la MTU de la interfaz. Por lo tanto, si la MTU de la interfaz era mayor que 1500 bytes, OSPF todavía empaquetaba sólo hasta 1500 bytes en un paquete OSPF. Esto fue algo ineficiente porque OSPF podía enviar paquetes más grandes en el link y lograr un mayor rendimiento.

> **Nota:** Hubo una excepción a este escenario. Si un LSA tenía más de 1500 bytes, OSPF construyó ese paquete, sin importar el tamaño, porque OSPF no puede fragmentar un LSA. A continuación, la pila IP del router fragmentó el paquete para ajustarse a la MTU de la interfaz saliente. Esto ocurrió típicamente cuando un router OSPF tenía muchos links, y el LSA del router se volvió más grande que la MTU del link.

De manera similar, si la MTU de la interfaz saliente era menor a 1500 bytes, el proceso OSPF aún generó o empaquetó paquetes OSPF de hasta 1500 bytes, y la pila IP del router fragmentó el paquete en paquetes IP más pequeños para ajustar la MTU del link saliente. Esto normalmente ocurrió con un túnel IPSec entre dos routers que ejecutaban OSPF. La sobrecarga agregada de los bytes de encapsulación del túnel llevó a una MTU menor a 1500 bytes. OSPF generó paquetes OSPF de hasta 1500 bytes y luego los paquetes se fragmentaron antes de que el router los transmitiera. Esta fue una ineficiencia adicional.

Después del ID de bug de Cisco CSCse01519

Después del Id. de bug Cisco CSCse01519, OSPF en IOS puede empaquetar paquetes OSPF

para que sean mayores de 1500 bytes. Esto ocurre si la MTU de la interfaz saliente es mayor que 1500 bytes. Las transmisiones son más eficientes porque se puede empaquetar más información en un paquete más grande. En otras palabras, si un router OSPF necesita transmitir muchos LSA externos a un vecino OSPF, puede empaquetar más LSA externos en un paquete de actualización de LS si ese router ejecuta IOS con el ID de error de Cisco CSCse01519 implementado.

El ID de bug Cisco CSCse01519 también permite que OSPF genere paquetes menores a 1500 bytes. En algunos escenarios, la MTU entre dos vecinos OSPF es menor que 1500 bytes. En el ejemplo anterior con un túnel IPSec, OSPF transmite paquetes OSPF que son menores a 1500 bytes y evita la fragmentación de IP; nuevamente, la excepción es el caso de un LSA que es mayor que la MTU de la interfaz.

Id. de error de Cisco CSCse01519

Cuando actualiza un router OSPF, puede detectar un problema de MTU OSPF causado por el ID de bug de Cisco CSCse01519.

Overview

Muchas redes tienen vecinos OSPF que se conectan a través de una red conmutada de capa 2 (L2) o de una red de transporte, que consta de un servicio VPN de capa 2 o una red de red de red SDH/SONET (Synchronous Digital Hierarchy/Synchronous Optical Network). Estas redes de transporte pueden tener diferentes configuraciones de MTU que los routers que ejecutan OSPF.

Aunque la configuración de MTU debe ser correcta en todos los routers y reflejar la MTU verdadera, a menudo hay errores que pasan desapercibidos.

Ésta es una red de ejemplo con dos routers que ejecutan OSPF. El router 1 (R1) y el router 2 (R2) están conectados a través de un switch L2.

Figure 1 : Example network

En este ejemplo, los routers tienen interfaces GigabitEthernet con una MTU configurada en 2000. La MTU del switch L2 es sólo 1500 bytes.

Si el tamaño del tráfico de datos nunca es mayor que 1500 bytes, puede utilizar IOS sin el ID de bug de Cisco CSCse01519 porque los paquetes OSPF nunca son mayores que 1500 bytes. Sin embargo, si hay un LSA que es de 1800 bytes, por ejemplo, el proceso OSPF en R1 o R2 construye un paquete LS Update mayor que 1500 bytes y lo transmite, pero el paquete es descartado por el switch L2 entre los routers.

Si la base de datos OSPF en R2 tiene suficientes redes, los LSAs originados localmente son tan grandes que un paquete LS Update puede ser mayor que la MTU de interfaz.

- Si estas redes se originan por el comando de red de cobertura, las redes aparecen en el LSA del router de R2. R2 construye un LSA de router que es mayor a 2000 bytes y lo transmite, pero IP lo fragmenta a 2000 bytes, la MTU de interfaz. Sin embargo, el switch L2 descarta estos paquetes. Luego OSPF retransmite este paquete sin fin y el estado de adyacencia OSPF nunca está completo. Por lo tanto, el problema se descubre inmediatamente, incluso cuando está ejecutando IOS sin el ID de bug Cisco CSCse01519.
- Si estas redes se originan por el comando **redistribute connected**, las redes aparecen en LSA externos. OSPF intenta empaquetar los LSA externos en un paquete LS Update que tiene un tamaño de hasta 1500 bytes. En este caso, debido a que la MTU de interfaz es de 2000 bytes, la adyacencia OSPF alcanza el estado 'FULL'. El problema de una MTU subyacente inadecuada no se descubre inmediatamente. El problema se detectará cuando un router se actualice a IOS con el ID de bug de Cisco CSCse01519.

Situación

Suponga que ambos routers ejecutan una versión de IOS sin el ID de bug de Cisco CSCse01519.

Cuando se genera la adyacencia OSFP, observe que R1 nunca recibe un paquete OSPF mayor a 1500 bytes, aunque la MTU de las interfaces es 2000.

Habilite el comando **debug ip ospf packets**.

```
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
...
```

En esta salida de depuración, 'l:1468' es la longitud del paquete OSPF, por lo que puede ver que el paquete OSPF más grande era de 1468 bytes. 't:4' indica que el paquete OSPF es el tipo 4, que es un paquete de actualización de estado de link. Esta tabla de la sección 4.3 de RFC 2328 define los diferentes tipos de paquetes OSPF:

Tipo	Nombre del paquete	Función de protocolo
1	Hello	Descubra/mantenga vecinos
2	Descripción de la base de datos	Resumir el contenido de la base de datos
3	Petición de estado de link	Descarga de base de datos
4	Actualización del estado de los links	Actualización de la base de datos
5	Ack De Estado De Link	Reconocimiento de inundación

La adyacencia OSPF alcanza el estado 'FULL'.

```
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
```

A continuación, actualice el IOS en R2 a una versión del IOS con el ID de bug de Cisco CSCse01519.

```
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
```

La adyacencia OSPF está atascada en el estado 'CARGANDO' y no alcanza el estado 'FULL'. Las retransmisiones se producen hasta que OSPF alcanza su límite de 25 retransmisiones. OSPF intenta establecer la adyacencia de nuevo, el mismo problema vuelve a ocurrir y el loop continúa interminablemente.

Por lo tanto, la actualización en R2 descubre un problema previamente oculto: la MTU subyacente

es más pequeña que la utilizada por los routers OSPF.

Cuando el switch cambia la MTU a 2000, un paquete OSPF mayor a 1500 bytes ('l:1980') se transmite sin ningún problema.

```
R1#
OSPF: rcv. v:2 t:3 l:1980 rid:10.100.1.2
     aid:0.0.0.0 chk:AC5B aut:0 auk: from GigabitEthernet0/1
```

Para verificar los problemas de MTU subyacentes, haga ping siempre a la dirección IP de vecino OSPF con un tamaño igual a la MTU y el conjunto de bits DF (no fragmentar).

Para detectar el valor de la MTU subyacente, realice un ping y barre el tamaño. Cuente el número de signos de exclamación (!) en la salida para determinar la MTU correcta. En este ejemplo, la última respuesta de eco del comando **ping** tiene un tamaño de 1500 bytes.

```
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
```