Nota tecnica sull'imballaggio di OSPF, MTU e LSA

Sommario

Introduzione

Questo documento descrive l'interazione tra i pacchetti OSPF (Open Shortest Path First), MTU (Maximum Transition Unit), LSA (Link State Advertisements) e LS (Link State) e i pacchetti di aggiornamento nel contesto del bug Cisco con ID CSCse01519.

Dimensioni pacchetto OSPF

I collegamenti sui router hanno una MTU. I pacchetti in uscita, ad esempio i pacchetti OSPF, non possono essere più grandi dell'MTU dell'interfaccia.

Request for Comments (RFC) 2328 documenti versione 2 del protocollo OSPF. L'appendice A.1 della RFC 2328 descrive l'incapsulamento dei pacchetti OSPF nel modo seguente:

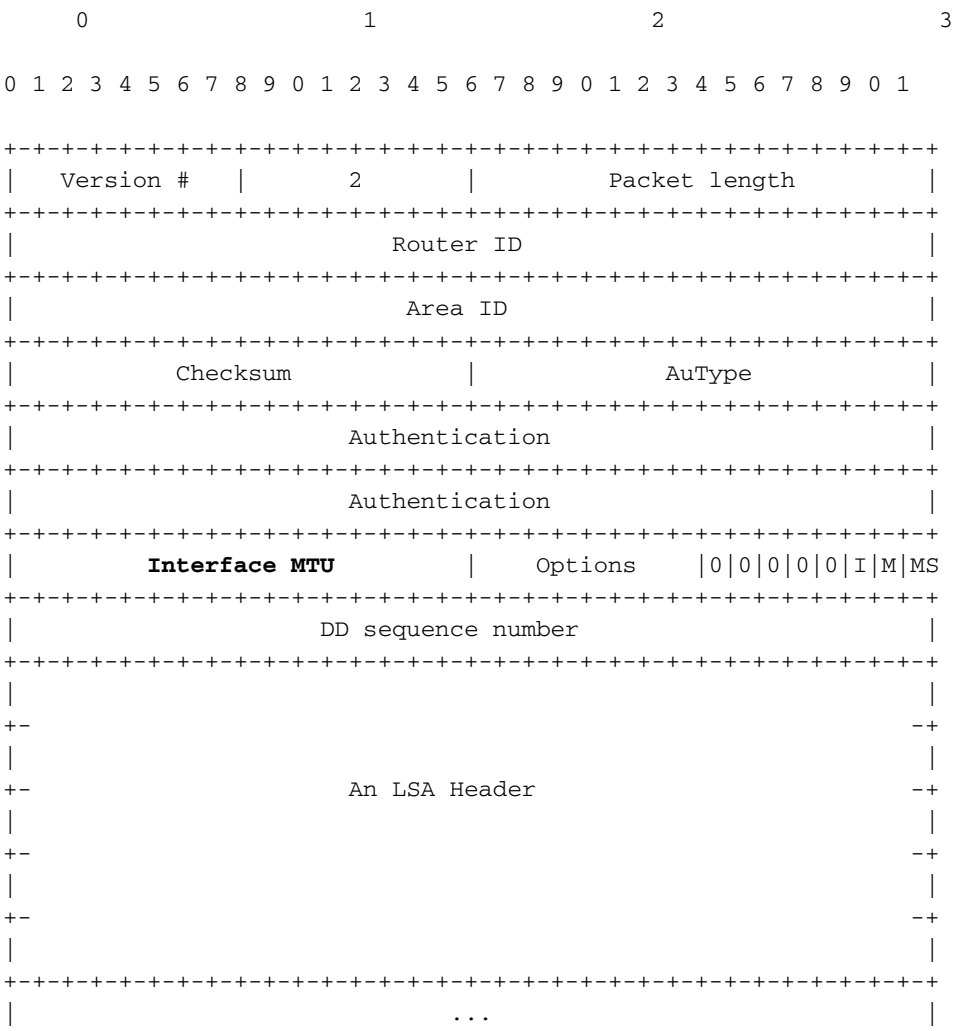
OSPF viene eseguito direttamente sul livello di rete del protocollo Internet. I pacchetti OSPF vengono quindi incapsulati solo dalle intestazioni IP e dal collegamento dati locale.

Il protocollo OSPF non definisce un modo per frammentare i pacchetti del protocollo e dipende dalla frammentazione IP quando trasmette pacchetti più grandi dell'MTU di rete. Se necessario, la lunghezza dei pacchetti OSPF può essere fino a 65.535 byte (inclusa l'intestazione IP). I tipi di pacchetti OSPF che potrebbero essere di grandi dimensioni (pacchetti Database Description, Link State Request, Link State Update e Link State Acknowledgment) possono in genere essere suddivisi in più pacchetti di protocollo separati, senza perdita di funzionalità. Si raccomanda quanto segue: Ove possibile, evitare la frammentazione IP.

Un pacchetto di aggiornamento LS può contenere una o più LSA. Molti LSA in un pacchetto LS Update sono noti come pacchetti LSA in un pacchetto LS Update.

MTU nel pacchetto DBD

Il pacchetto Database Description (DBD), specificato anche nella RFC 2328, descrive il contenuto del database dello stato del collegamento OSPF:



L'appendice A.3.3 della RFC 2328 descrive l'MTU dell'interfaccia come segue:

Dimensioni in byte del datagramma IP più grande che può essere inviato all'interfaccia associata, senza frammentazione.

I router collegati a un collegamento scambiano il valore MTU dell'interfaccia in pacchetti DBD quando viene inizializzata la adiacenza OSPF.

La sezione 10.6 della RFC 2328 afferma:

Se il campo MTU interfaccia nel pacchetto Database Description indica una dimensione del datagramma IP superiore a quella che il router può accettare sull'interfaccia ricevente senza frammentazione, il pacchetto Database Description viene rifiutato.

Quando si usa il comando **debug ip ospf adj**, è possibile verificare l'arrivo dei pacchetti DBD.

Nell'esempio, esiste una mancata corrispondenza nei valori MTU tra due router adiacenti OSPF. Questo router ha 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
```

L'altro router OSPF ha l'interfaccia 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
```

I pacchetti DBD vengono ritrasmessi continuamente fino a quando l'adiacenza OSPF non viene eliminata.

```
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
```

Comportamento OSPF e inserimento di LSA in un pacchetto di aggiornamento LS

Prima dell'ID bug Cisco CSCse01519

Prima dell'ID bug Cisco CSCse01519, l'OSPF nel software Cisco IOS® generava pacchetti OSPF non più grandi di 1500 byte, indipendentemente dall'MTU dell'interfaccia. Pertanto, se l'MTU dell'interfaccia era superiore a 1500 byte, il protocollo OSPF continua a contenere solo fino a 1500 byte in un pacchetto OSPF. Questa operazione non era efficiente in quanto OSPF poteva inviare pacchetti più grandi sul collegamento e ottenere un throughput maggiore.

> **Nota:** Si è verificata un'eccezione a questo scenario. Se un LSA contiene più di 1500 byte, l'OSPF lo ha creato, a prescindere dalle dimensioni, in quanto non può frammentare un LSA. Lo stack IP del router ha quindi frammentato il pacchetto per adattarlo alla MTU dell'interfaccia in uscita. Questo si verifica in genere quando un router OSPF ha molti collegamenti e la LSA del router diventa più grande della MTU del collegamento.

Analogamente, se l'MTU dell'interfaccia in uscita era inferiore a 1500 byte, il processo OSPF ha ancora creato o compresso pacchetti OSPF fino a 1500 byte e lo stack IP del router ha frammentato il pacchetto in pacchetti IP più piccoli per adattarlo alla MTU del collegamento in uscita. Ciò si è in genere verificato con un tunnel IPSec tra due router che eseguono OSPF. Il sovraccarico aggiunto dei byte di incapsulamento del tunnel ha portato a una MTU inferiore a 1500 byte. I pacchetti OSPF sono stati creati fino a 1500 byte e sono stati frammentati prima che il router li trasmettesse. Questa era un'ulteriore inefficienza.

Dopo l'ID bug Cisco CSCse01519

Dopo l'ID bug Cisco CSCse01519, l'OSPF in IOS può comprimere i pacchetti OSPF in modo che superino i 1500 byte. Questo si verifica se l'MTU dell'interfaccia in uscita è superiore a 1500 byte. Le trasmissioni sono più efficienti perché è possibile inserire una maggiore quantità di informazioni in un unico pacchetto di dimensioni maggiori. In altre parole, se un router OSPF deve trasmettere molti LSA esterni a un router adiacente OSPF, può comprimere più LSA esterni in un pacchetto di aggiornamento LS se il router esegue IOS con ID bug Cisco CSCse01519 implementato.

L'ID bug Cisco CSCse01519 consente anche alla piattaforma OSPF di generare pacchetti più piccoli di 1500 byte. In alcuni scenari, l'MTU tra due vicini OSPF è inferiore a 1500 byte. Nell'esempio precedente relativo a un tunnel IPSec, il protocollo OSPF trasmette pacchetti OSPF inferiori a 1500 byte ed evita la frammentazione IP; anche in questo caso, l'eccezione è rappresentata da un LSA più grande dell'MTU dell'interfaccia.

ID bug Cisco CSCse01519

Quando si aggiorna un router OSPF, è possibile che venga rilevato un problema di MTU OSPF causato dall'ID bug Cisco CSCse01519.

Panoramica

Molte reti dispongono di router adiacenti OSPF connessi tramite una rete commutata Layer 2 (L2) o una rete di trasporto, costituita da un servizio VPN L2 o da una rete SDH/SONET (Synchronous Digital Hierarchy/Synchronous Optical Network). Queste reti di trasporto possono avere impostazioni MTU diverse dai router che eseguono OSPF.

Anche se l'MTU deve essere impostata correttamente su tutti i router e riflettere la MTU effettiva, spesso gli errori passano inosservati.

Si tratta di una rete di esempio con due router che eseguono OSPF. Il router 1 (R1) e il router 2 (R2) sono collegati tramite uno switch L2.

Figure 1 : Example network

Nell'esempio, i router dispongono di interfacce Gigabit Ethernet con MTU impostata su 2000. L'MTU dello switch L2 è solo 1500 byte.

Se le dimensioni del traffico di dati non sono mai superiori a 1500 byte, è possibile usare IOS senza l'ID bug Cisco CSCse01519 perché i pacchetti OSPF non sono mai superiori a 1500 byte. Tuttavia, se ad esempio esiste un LSA di 1800 byte, il processo OSPF su R1 o R2 crea un pacchetto LS Update di dimensioni superiori a 1500 byte e lo trasmette, ma il pacchetto viene scartato dallo switch L2 tra i router.

Se il database OSPF su R2 ha un numero sufficiente di reti, le LSA originate localmente sono così grandi che un pacchetto di aggiornamento LS potrebbe essere più grande dell'MTU dell'interfaccia.

- Se queste reti sono state create dal comando cover network, le reti appaiono nel router LSA di R2. R2 crea un router LSA più grande di 2000 byte e lo trasmette, ma IP lo frammenta a 2000 byte, l'MTU dell'interfaccia. Tuttavia, lo switch L2 scarta questi pacchetti. A questo punto, OSPF ritrasmette il pacchetto all'infinito e lo stato di adiacenza OSPF non è mai pieno. Il problema viene quindi rilevato immediatamente, anche quando si esegue IOS senza l'ID bug Cisco CSCse01519.
- Se le reti sono state create con il comando **redistribute connected**, vengono visualizzate nelle LSA esterne. OSPF tenta di comprimere le LSA esterne in un pacchetto LS Update di dimensioni fino a 1500 byte. In questo caso, poiché l'MTU dell'interfaccia è di 2000 byte, la adiacenza OSPF raggiunge lo stato 'FULL'. Il problema dell'MTU sottostante inadeguata non viene rilevato immediatamente. Il problema verrà rilevato quando un router viene aggiornato a IOS con ID bug Cisco CSCse01519.

Scenario

Si supponga che entrambi i router eseguano una versione IOS senza l'ID bug Cisco CSCse01519.

Quando viene compilata la adiacenza OSFP, si noti che R1 non riceve mai un pacchetto OSPF più grande di 1500 byte, anche se l'MTU delle interfacce è 2000.

Abilitare il 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
...
```

In questo output di debug, 'l:1468' è la lunghezza del pacchetto OSPF, quindi è possibile notare che il pacchetto OSPF più grande era di 1468 byte. 't:4' indica che il pacchetto OSPF è di tipo 4, che è un pacchetto di aggiornamento dello stato del collegamento. La tabella riportata di seguito dalla sezione 4.3 della RFC 2328 definisce i diversi tipi di pacchetti OSPF:

Tipo	Nome pacchetto	Funzione Protocol
1	Salve	Rileva/gestisci vicini
2	Descrizione database	Riepiloga contenuto database
3	Richiesta stato collegamento	Download database
4	Aggiornamento stato collegamento	Aggiornamento database
5	ACK stato collegamento	Riconoscimento flooding

L'adiacenza OSPF raggiunge lo stato '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
```

Quindi, aggiornare IOS su R2 alla versione IOS con ID bug 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
```

L'adiacenza OSPF è bloccata nello stato 'LOADING' e non raggiunge lo stato 'FULL'. Le ritrasmissioni si verificano fino al raggiungimento del limite di 25 ritrasmissioni OSPF. OSPF tenta di stabilire nuovamente l'adiacenza, si verifica nuovamente lo stesso problema e il loop continua all'infinito.

Pertanto, l'aggiornamento su R2 individua un problema nascosto in precedenza: la MTU sottostante è più piccola di quella utilizzata dai router OSPF.

Quando lo switch cambia la MTU a 2000, un pacchetto OSPF più grande di 1500 byte ('l:1980') viene trasmesso senza problemi.

```
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
```

Per controllare i problemi MTU sottostanti, eseguire sempre il ping dell'indirizzo IP del router adiacente OSPF con una dimensione uguale all'MTU e al bit DF (non frammentare) impostati.

Per individuare il valore dell'MTU sottostante, eseguire un ping e ridimensionare il pacchetto. Contare il numero di punti esclamativi (!) nell'output per determinare la MTU corretta. In questo esempio, l'ultima risposta echo del comando **ping** ha una dimensione di 1500 byte.

```
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
```