Technischer Hinweis für OSPF-, MTU- und LSA-Pakete

Inhalt

Einführung

Dieses Dokument beschreibt die Interaktion von OSPF-Paketen (Open Shortest Path First), MTU (Maximum Transition Unit), LSAs (Link State Advertisements) und LS-Update-Paketen (Link State Advertisements) im Zusammenhang mit der Cisco Bug-ID CSCse01519.

OSPF-Paketgröße

Links auf Routern haben eine MTU. Ausgehende Pakete, z. B. OSPF-Pakete, dürfen nicht größer als die MTU der Schnittstelle sein.

Request for Comments (RFC) 2328-Dokumente, Version 2 des OSPF-Protokolls. In Anhang A.1 von RFC 2328 wird die Kapselung von OSPF-Paketen folgendermaßen beschrieben:

OSPF wird direkt über die Netzwerkebene des Internetprotokolls ausgeführt. OSPF-Pakete werden daher ausschließlich durch IP- und lokale Sicherungsschichtheader gekapselt.

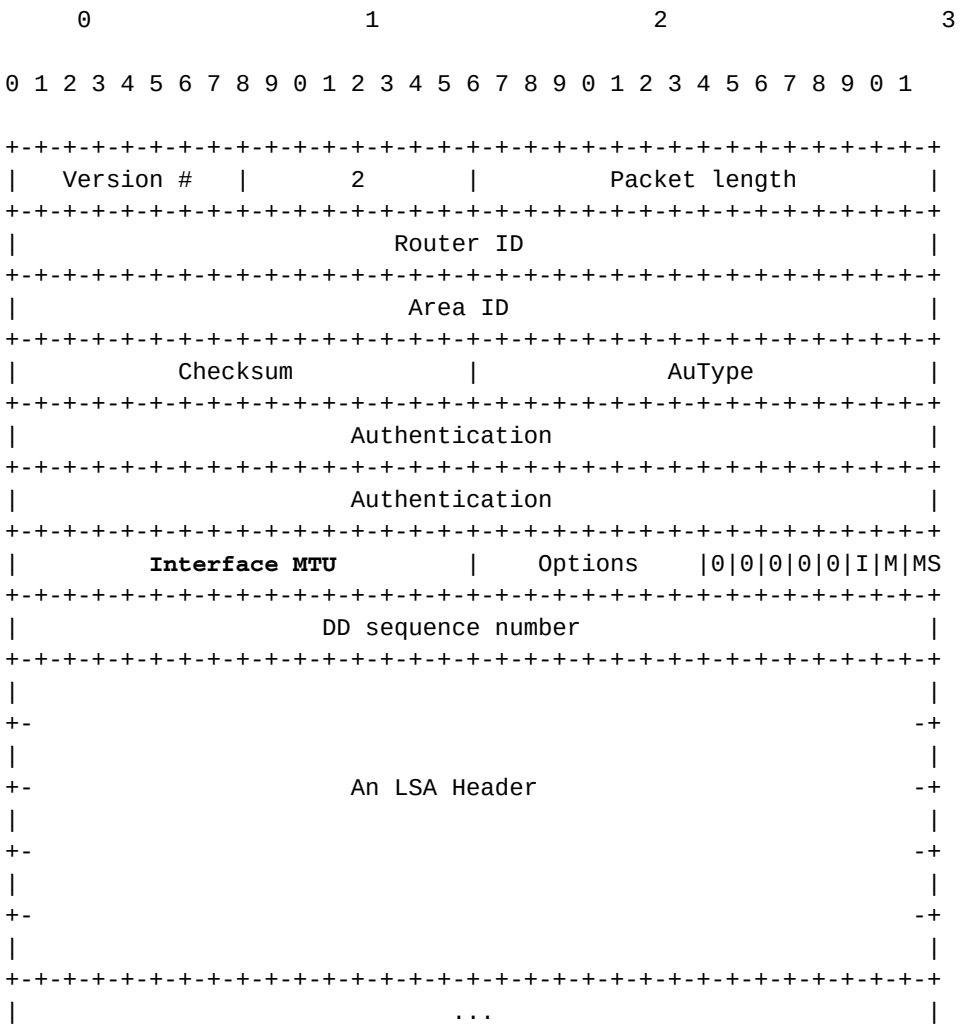
OSPF definiert keine Möglichkeit zur Fragmentierung seiner Protokollpakete und ist bei der Übertragung von Paketen, die größer als die Netzwerk-MTU sind, von der IP-Fragmentierung abhängig. Bei Bedarf kann die Länge der OSPF-Pakete bis zu 65.535 Byte (einschließlich des IP-Headers) betragen. Die OSPF-Pakettypen, die wahrscheinlich groß sind (Datenbankbeschreibungspakete, Verbindungsstatusanforderung, Link State Update und Link State Acknowledgment-Pakete), können in der Regel ohne Funktionsverlust in mehrere separate Protokollpakete aufgeteilt werden. Dies wird empfohlen. IP-Fragmentierung sollte möglichst vermieden werden.

Ein LS-Update-Paket kann einen oder mehrere LSAs enthalten. Viele LSAs in einem LS-Update-

Paket werden als Verpacken von LSAs in ein LS-Update-Paket bezeichnet.

MTU in DBD-Paket

Das ebenfalls in RFC 2328 angegebene Datenbankbeschreibungspaket (DBD) beschreibt den Inhalt der OSPF-Link-State-Datenbank:



Anlage A.3.3. von RFC 2328 beschreibt die MTU-Schnittstellengröße wie folgt:

Die Größe (in Byte) des größten IP-Datagramms, das ohne Fragmentierung über die zugeordnete Schnittstelle gesendet werden kann.

Router, die an eine Verbindung angeschlossen sind, tauschen ihre Schnittstellen-MTU-Werte in DBD-Paketen aus, wenn die OSPF-Adjacency initialisiert wird.

In Abschnitt 10.6 von RFC 2328 werden folgende Zustände angegeben:

Wenn das Feld Interface MTU (Schnittstelle-MTU) im Paket Database Description (Datenbankbeschreibung) eine IP-Datagrammgröße anzeigt, die größer ist, als der Router auf der empfangenden Schnittstelle ohne Fragmentierung akzeptieren kann, wird das Datenbankbeschreibungspaket abgelehnt.

Wenn der Befehl **debug ip ospf adj** verwendet wird, können Sie den Eingang dieser DBD-Pakete sehen.

In diesem Beispiel gibt es eine Diskrepanz bei den MTU-Werten zwischen zwei OSPF-Nachbarn. Dieser Router hat 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
```

Der andere OSPF-Router verfügt über die Schnittstelle 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
```

Die DBD-Pakete werden fortlaufend neu übertragen, bis die OSPF-Adjacency schließlich ausgeschaltet wird.

```
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
```

OSPF-Verhalten und Packen von LSAs in ein LS-Update-Paket

Vor Cisco Bug-ID CSCse01519

Vor der Cisco Bug-ID CSCse01519 erstellte OSPF in der Cisco IOS®-Software OSPF-Pakete mit maximal 1500 Byte, unabhängig von der MTU der Schnittstelle. Wenn die MTU-Größe der Schnittstelle also größer als 1.500 Byte ist, wird OSPF immer noch nur bis zu 1.500 Byte in einem OSPF-Paket gepackt. Dies war etwas ineffizient, da OSPF größere Pakete auf der Verbindung senden und einen höheren Durchsatz erzielen konnte.

> **Hinweis**: Es gab eine Ausnahme zu diesem Szenario. Wenn ein LSA mehr als 1500 Byte enthielt, erstellte OSPF dieses Paket unabhängig von der Größe, da OSPF kein einzelnes LSA fragmentieren kann. Der IP-Stack des Routers fragmentierte dann das Paket, um die MTU der ausgehenden Schnittstelle zu übernehmen. Dies ist in der Regel der Fall, wenn ein OSPF-Router viele Verbindungen hatte und das LSA des Routers größer als die Verbindungs-MTU wurde.

Wenn die MTU der ausgehenden Schnittstelle kleiner als 1.500 Byte war, wurde beim OSPF-Prozess weiterhin OSPF-Pakete von bis zu 1.500 Byte erstellt oder gepackt, und der IP-Stack des Routers fragmentierte das Paket in kleinere IP-Pakete, um die MTU der ausgehenden Verbindung zu übernehmen. Dies trat in der Regel bei einem IPSec-Tunnel zwischen zwei Routern auf, auf denen OSPF ausgeführt wurde. Der hinzugefügte Overhead der Kapselungsbyte des Tunnels führte zu einer MTU, die kleiner als 1500 Byte war. OSPF erstellte OSPF-Pakete mit bis zu 1.500 Byte, und die Pakete wurden anschließend fragmentiert, bevor der Router sie übertraf. Dies war eine zusätzliche Ineffizienz.

Nach Cisco Bug-ID CSCse01519

Nach der Cisco Bug-ID CSCse01519 kann OSPF in IOS OSPF-Pakete mit einer Größe von mehr als 1.500 Byte packen. Dies tritt auf, wenn die MTU der ausgehenden Schnittstelle größer als 1500 Byte ist. Übertragungen sind effizienter, da mehr Informationen in einem größeren Paket zusammengefasst werden können. Anders ausgedrückt: Wenn ein OSPF-Router viele externe LSAs an einen OSPF-Nachbarn senden muss, kann er weitere externe LSAs in einem LS-Update-Paket packen, wenn dieser Router IOS mit implementierter Cisco Bug-ID CSCse01519 ausführt.

Mit der Cisco Bug-ID CSCse01519 kann OSPF auch Pakete erstellen, die kleiner als 1.500 Byte sind. In einigen Szenarien ist die MTU zwischen zwei OSPF-Nachbarn kleiner als 1.500 Byte. Im vorherigen Beispiel sendet OSPF mit einem IPSec-Tunnel OSPF OSPF-Pakete, die kleiner als 1.500 Byte sind, und vermeidet die IP-Fragmentierung. Eine Ausnahme bildet wiederum ein LSA, das größer als die MTU der Schnittstelle ist.

Cisco Bug-ID CSCse01519

Beim Upgrade eines OSPF-Routers wird möglicherweise ein OSPF-MTU-Problem festgestellt, das durch die Cisco Bug ID CSCse01519 verursacht wird.

Übersicht

Viele Netzwerke verfügen über OSPF-Nachbarn, die über ein Layer-2-Switch-Netzwerk (L2) oder ein Transportnetzwerk verbunden sind, das aus einem L2-VPN-Service oder einem SDH/SONET-Netzwerk (Synchronous Digital Hierarchy/Synchronous Optical Network) besteht. Diese Transportnetzwerke können unterschiedliche MTU-Einstellungen haben als die Router, auf denen OSPF ausgeführt wird.

Obwohl die MTU-Einstellung auf allen Routern korrekt sein und die tatsächliche MTU widerspiegeln sollte, gibt es häufig Fehler, die nicht beachtet werden.

Dies ist ein Beispielnetzwerk mit zwei Routern, auf denen OSPF ausgeführt wird. Router 1 (R1) und Router 2 (R2) sind über einen L2-Switch verbunden.

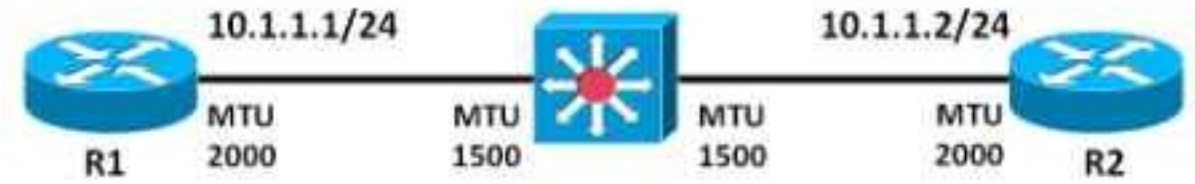


Figure 1 : Example network

In diesem Beispiel verfügen die Router über GigabitEthernet-Schnittstellen mit einer MTU von 2000. Die MTU des L2-Switches beträgt nur 1.500 Byte.

Wenn die Größe des Datenverkehrs nie größer als 1.500 Byte ist, können Sie IOS ohne Cisco Bug ID CSCse01519 verwenden, da die OSPF-Pakete nie größer als 1.500 Byte sind. Wenn es jedoch ein LSA mit 1800 Byte gibt, erstellt der OSPF-Prozess auf R1 oder R2 ein LS-Update-Paket mit mehr als 1500 Byte und überträgt es, aber das Paket wird vom L2-Switch zwischen den Routern verworfen.

Wenn die OSPF-Datenbank auf R2 über genügend Netzwerke verfügt, sind die lokal generierten LSAs so groß, dass ein LS-Update-Paket größer als die MTU der Schnittstelle sein kann.

- Wenn diese Netzwerke vom Befehl "cover network" (Netzwerk abdecken) stammen, werden die Netzwerke im Router-LSA von R2 angezeigt. R2 erstellt ein Router-LSA, das mehr als 2000 Byte umfasst, und überträgt es, aber das IP fragmentiert es auf 2000 Byte, die MTU der Schnittstelle. Der L2-Switch verwirft diese Pakete jedoch. OSPF überträgt dieses Paket dann endlos erneut, und der OSPF-Adjacency-Status ist niemals voll. Das Problem wird also sofort entdeckt, auch wenn Sie IOS ohne Cisco Bug ID CSCse01519 ausführen.
- Wenn diese Netzwerke durch den Befehl **redistribute connected** (Verbindung **neu verteilen**) generiert werden, werden die Netzwerke in externen LSAs angezeigt. OSPF versucht, externe LSAs in einem LS-Update-Paket mit einer Größe von bis zu 1.500 Byte zu packen. In diesem Fall erreicht die OSPF-Adjacency den Zustand "VULL", da die MTU-Schnittstellengröße 2000 Byte beträgt. Das Problem einer unzureichenden zugrunde liegenden MTU wird nicht sofort erkannt. Das Problem wird erkannt, wenn ein Router mit der Cisco Bug-ID CSCse01519 auf IOS aktualisiert wird.

Szenario

Nehmen Sie an, dass auf beiden Routern eine IOS-Version ohne Cisco Bug ID CSCse01519 ausgeführt wird.

Beachten Sie bei der Erstellung der OSFP-Adjacency, dass R1 niemals ein OSPF-Paket mit einer Größe von mehr als 1500 Byte empfängt, obwohl die MTU der Schnittstellen 2000 beträgt.

Aktivieren Sie den Befehl **debug ip ospf packages**.

```
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 dieser Debug-Ausgabe ist 'l:1468' die Länge des OSPF-Pakets. Sie sehen also, dass das größte OSPF-Paket 1468 Byte betrug. 't:4' gibt an, dass das OSPF-Paket vom Typ 4 ist, d. h. ein Link State Update-Paket. In dieser Tabelle in Abschnitt 4.3 von RFC 2328 werden die verschiedenen OSPF-Pakettypen definiert:

Typ	Paketname	Protokollfunktion
1	Hallo	Erkennung/Pflege von Nachbarn
2	Datenbankbeschreibung	Zusammenfassung des Datenbankinhalts

1	Link-State-Anforderung	Download der Datenbank
4	Link-State-Update	Datenbankaktualisierung
5	Link-State-Bestätigung	Flooding-Bestätigung

Die OSPF-Adjacency erreicht den Status 'VOLLSTÄNDIG'.

```
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
```

Als Nächstes aktualisieren Sie IOS auf R2 mit der Cisco Bug-ID CSCse01519 auf eine IOS-Version.

```
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
```

Die OSPF-Adjacency ist im 'LOADING'-Status fixiert und erreicht nicht den 'VOLLSTÄNDIGEN' Status. Neuübertragungen erfolgen, bis OSPF den Grenzwert von 25 Neuübertragungen erreicht. OSPF versucht erneut, die Adjacency einzurichten, das gleiche Problem tritt erneut auf, und die Schleife wird endlos fortgesetzt.

Das Upgrade auf R2 deckt also ein bereits ausgeblendetes Problem auf: Die zugrunde liegende MTU ist kleiner als die MTU, die von den OSPF-Routern verwendet wird.

Wenn der Switch die MTU auf 2000 ändert, wird problemlos ein OSPF-Paket mit einer Größe von mehr als 1500 Byte ('l:1980') übertragen.

```
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
```

Um die zugrunde liegenden MTU-Probleme zu überprüfen, pingen Sie immer die IP-Adresse des OSPF-Nachbarn mit einer Größe, die der MTU- und der DF-Bitmenge (nicht fragmentieren) entspricht.

Um den Wert der zugrunde liegenden MTU zu ermitteln, führen Sie einen Ping aus, und kehren Sie die Größe zurück. Zählen Sie die Anzahl der Ausrufezeichen (!) in der Ausgabe, um die korrekte MTU zu bestimmen. In diesem Beispiel hat die letzte Echo-Antwort des Befehls **ping** eine Größe von 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
```