OSPF, MTU 및 LSA 포장 기술 노트

목차

소개

이 문서에서는 Cisco 버그 ID CSCse01519 컨텍스트에서 OSPF(Open Shortest Path First) 패킷, MTU(Maximum Transition Unit), LSA(Link State Advertisements) 및 LS(Link State) 업데이트 패킷의 상호 작용에 대해 [설명합니다](#).

OSPF 패킷 크기

라우터의 링크에는 MTU가 있습니다.OSPF 패킷과 같은 발신 패킷은 인터페이스 MTU보다 클 수 없습니다.

[RFC(Request for Comments) 2328](#) 문서 버전 2(OSPF 프로토콜)RFC 2328의 부록 A.1에서는 다음과 같은 방법으로 OSPF 패킷의 캡슐화에 대해 설명합니다.

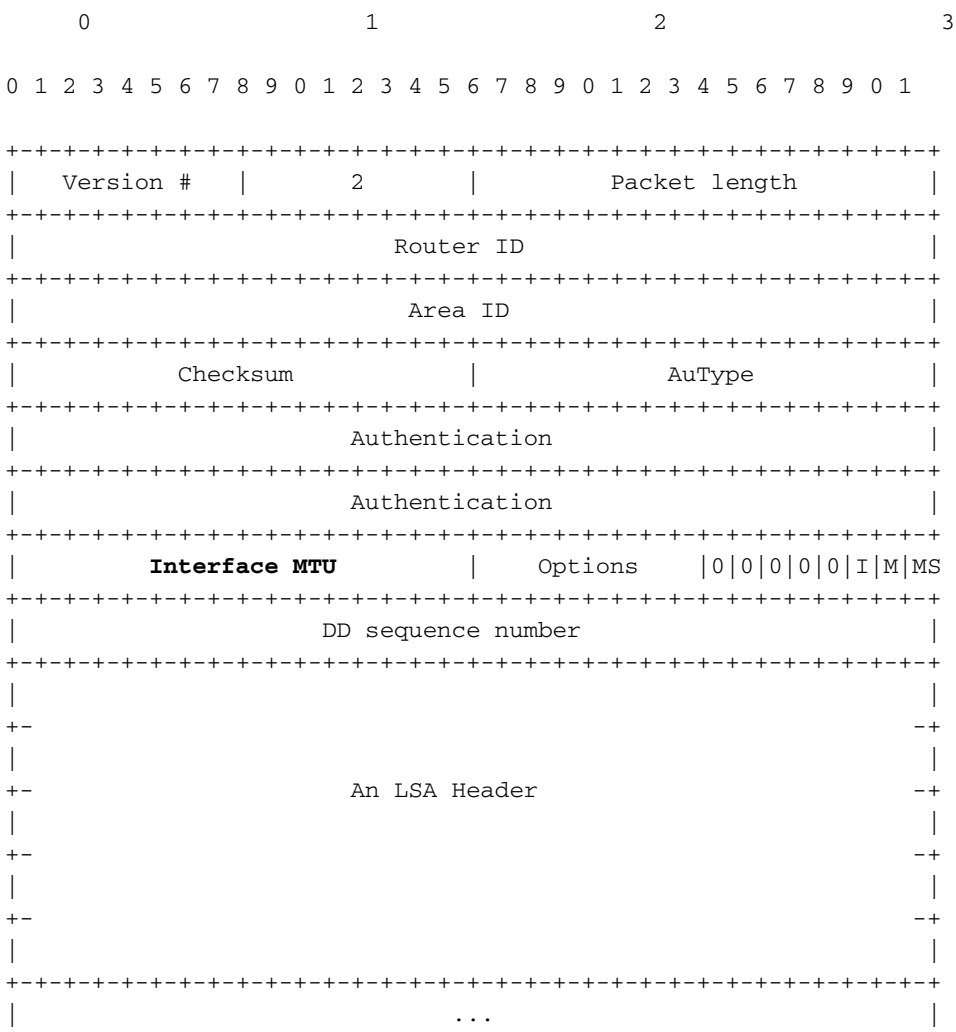
OSPF는 인터넷 프로토콜의 네트워크 레이어를 통해 직접 실행됩니다.따라서 OSPF 패킷은 IP 및 로컬 데이터 링크 헤더로만 캡슐화됩니다.

OSPF는 프로토콜 패킷을 단편화하는 방법을 정의하지 않으며, 네트워크 MTU보다 큰 패킷을 전송할 때 IP 단편화에 의존합니다.필요한 경우 OSPF 패킷의 길이는 최대 65,535바이트(IP 헤더 포함)가 될 수 있습니다. 대규모 OSPF 패킷 유형(Database Description Packets, Link State Request, Link State Update 및 Link State Acknowledgment Packets)은 일반적으로 여러 개의 개별 프로토콜 패킷으로 분할될 수 있으며, 기능 손실 없이 분할할 수 있습니다.권장됩니다.IP 프래그먼트화는 가능하면 피해야 합니다.

LS 업데이트 패킷에 하나 이상의 LSA가 있을 수 있습니다.하나의 LS 업데이트 패킷에 있는 많은 LSA를 LS 업데이트 패킷으로 LSA를 패키지하는 것으로 알려져 있습니다.

DBD 패킷의 MTU

RFC 2328에도 지정된 DBD(Database Description) 패킷은 OSPF 링크 상태 데이터베이스의 내용을 설명합니다.



RFC 2328의 부록 A.3.3에서는 인터페이스 MTU를 다음과 같이 설명합니다.

프래그먼트화 없이 연결된 인터페이스에서 전송할 수 있는 가장 큰 IP 데이터그램의 크기(바이트)입니다.

링크에 연결된 라우터는 OSPF 인접성이 초기화될 때 DBD 패킷에서 인터페이스 MTU 값을 교환합니다.

RFC 2328의 섹션 10.6은 다음과 같습니다.

Database Description 패킷의 Interface MTU 필드가 프래그먼트화 없이 수신 인터페이스에서 라우터가 수용할 수 있는 크기보다 큰 IP 데이터그램 크기를 나타내는 경우 Database Description 패킷이 거부됩니다.

debug **ip ospf adj** 명령을 사용하면 이러한 DBD 패킷의 도착을 확인할 수 있습니다.

이 예에서는 두 OSPF 네이버 간에 MTU 값이 일치하지 않습니다.이 라우터에는 MTU 1600이 있습니다.

```
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
```

다른 OSPF 라우터에는 MTU 2000 인터페이스가 있습니다.

```
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
```

DBD 패킷은 OSPF 인접성이 결국 해제될 때까지 지속적으로 재전송됩니다.

```
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 동작 및 LSA를 LS 업데이트 패킷으로 포장

Cisco 버그 ID CSCse01519 이전

Cisco 버그 ID CSCse01519에 앞서 Cisco IOS® 소프트웨어의 OSPF는 인터페이스 MTU와 상관없이 1500바이트보다 크지 않은 OSPF 패킷을 구축했습니다.따라서 인터페이스 MTU가 1500바이트보다 큰 경우 OSPF는 여전히 OSPF 패킷에 최대 1500바이트까지만 압축됩니다.OSPF가 링크에서 더 큰 패킷을 전송하고 더 많은 처리량을 얻을 수 있기 때문에 이 방법은 다소 비효율적이었습니다.

> **참고:**이 시나리오에는 한 가지 예외가 있습니다.하나의 LSA가 1500바이트를 초과하는 경우 OSPF는 LSA를 프래그먼트화할 수 없으므로 크기와 상관없이 해당 패킷을 구축했습니다.그런 다음 라우터의 IP 스택은 나가는 인터페이스의 MTU에 맞게 패킷을 조각화했습니다.이 문제는 일반적으로 OSPF 라우터에 많은 링크가 있고 라우터 LSA가 링크 MTU보다 커졌을 때 발생합니다.

마찬가지로, 발신 인터페이스의 MTU가 1500바이트보다 작으면 OSPF 프로세스는 1500바이트까지 OSPF 패킷을 계속 빌드하거나 압축하며, 라우터의 IP 스택은 나가는 링크의 MTU에 맞도록 패킷을 더 작은 IP 패킷으로 프래그먼트화했습니다.일반적으로 OSPF를 실행 중인 두 라우터 간의 IPSec 터널에서 이 오류가 발생했습니다.터널의 캡슐화 바이트의 추가된 오버헤드는 1500바이트보다 작은 MTU로 이어졌습니다.OSPF는 최대 1500바이트의 OSPF 패킷을 구축했으며, 그런 다음 라우터가 전송하기 전에 패킷이 프래그먼트화되었습니다.이는 추가적인 비효율성이었습니다.

Cisco 버그 ID CSCse01519 이후

Cisco 버그 ID CSCse01519 이후, IOS의 OSPF는 1500바이트보다 큰 OSPF 패킷을 압축할 수 있습니다.이는 발신 인터페이스의 MTU가 1500바이트보다 큰 경우 발생합니다.더 많은 정보를 더 큰 하나의 패킷으로 압축할 수 있으므로 전송 효율성이 향상됩니다.즉, 하나의 OSPF 라우터가 여러 외부 LSA를 OSPF 네이버로 전송해야 하는 경우, 해당 라우터가 Cisco 버그 ID CSCse01519가 구현된 IOS를 실행하는 경우 더 많은 외부 LSA를 하나의 LS 업데이트 패킷으로 압축할 수 있습니다.

Cisco 버그 ID CSCse01519를 사용하면 OSPF에서 1500바이트보다 작은 패킷을 작성할 수 있습니다.일부 시나리오에서는 두 OSPF 네이버 간의 MTU가 1500바이트보다 작습니다.IPSec 터널의 이전 예에서 OSPF는 1500바이트보다 작고 IP 단편화를 방지하는 OSPF 패킷을 전송합니다.다시, 인터페이스 MTU보다 큰 LSA의 경우는 예외입니다.

Cisco 버그 ID CSCse01519

OSPF 라우터를 업그레이드할 때 Cisco 버그 ID CSCse01519로 인해 발생한 OSPF MTU 문제를 발견할 수 있습니다.

개요

많은 네트워크에는 L2 VPN 서비스 또는 SDH/SONET(Synchronous Digital Hierarchy/Synchronous Optical Network) 네트워크로 구성된 레이어 2(L2) 스위치 네트워크 또는 전송 네트워크를 통해 연결된 OSPF 인접 디바이스가 있습니다.이러한 전송 네트워크는 OSPF를 실행하는 라우터와 다른 MTU 설정을 가질 수 있습니다.

MTU 설정은 모든 라우터에서 올바르며 실제 MTU를 반영해야 하지만 자주 간과되는 오류가 있습니다.

이것은 OSPF를 실행 중인 두 라우터가 있는 네트워크의 예입니다.라우터 1(R1) 및 라우터 2(R2)는 L2 스위치를 통해 연결됩니다.

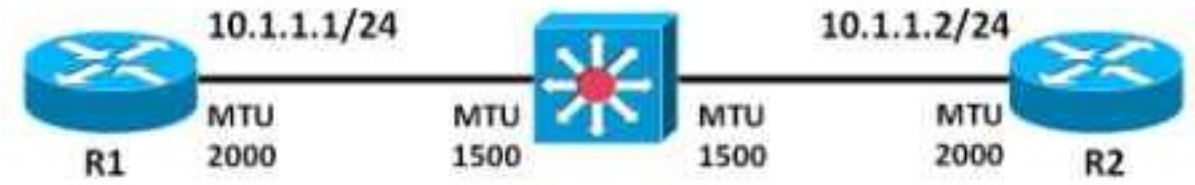


Figure 1 : Example network

이 예에서는 라우터에 MTU가 2000으로 설정된 GigabitEthernet 인터페이스가 있습니다.L2 스위치의 MTU는 1500바이트입니다.

데이터 트래픽의 크기가 1500바이트보다 크지 않을 경우 OSPF 패킷이 1500바이트보다 크지 않으므로 Cisco 버그 ID CSCse01519 없이 IOS를 사용할 수 있습니다.그러나 예를 들어 1800바이트인 LSA가 있는 경우 R1 또는 R2의 OSPF 프로세스는 1500바이트보다 큰 LS 업데이트 패킷을 작성하여 전송하지만 라우터 간 L2 스위치에서 패킷을 삭제합니다.

R2의 OSPF 데이터베이스에 충분한 네트워크가 있는 경우 로컬에서 시작된 LSA가 너무 커서 LS 업데이트 패킷이 인터페이스 MTU보다 클 수 있습니다.

- 이러한 네트워크가 covering network 명령으로 시작되는 경우 R2의 라우터 LSA에 네트워크가 나타납니다. R2는 2000바이트보다 큰 라우터 LSA를 구축하여 전송하지만 IP는 2000바이트인 인터페이스 MTU로 프래그먼트합니다.그러나 L2 스위치는 이러한 패킷을 삭제합니다.그런 다음 OSPF는 이 패킷을 끊임없이 재전송하며, OSPF 인접성 상태가 절대 꽉 차지 않습니다.따라서 Cisco 버그 ID CSCse01519 없이 IOS를 실행하는 경우에도 문제가 즉시 발견됩니다.
- 이러한 네트워크가 **redistribute connected** 명령에 의해 시작된 경우 네트워크가 외부 LSA에 나타납니다.OSPF는 외부 LSA를 최대 1500바이트 크기의 단일 LS 업데이트 패킷으로 압축하려고 시도합니다.이 경우 인터페이스 MTU는 2000바이트이므로 OSPF 인접성은 'FULL' 상태에 도달합니다.부적절한 기본 MTU의 문제는 즉시 발견되지 않습니다.이 문제는 Cisco 버그 ID CSCse01519가 있는 IOS로 한 라우터를 업그레이드하면 검색됩니다.

시나리오

두 라우터가 모두 Cisco 버그 ID CSCse01519 없이 IOS 버전을 실행한다고 가정합니다.

OSFP 인접성이 구축되면 인터페이스의 MTU는 2000이지만 R1은 1500바이트보다 큰 OSPF 패킷을 수신하지 않습니다.

debug **ip ospf packets** 명령을 활성화합니다.

```
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
...
```

이 디버그 출력에서 'l:1468'은 OSPF 패킷의 길이이므로 가장 큰 OSPF 패킷이 1468바이트임을 확인할 수 있습니다.'t:4'는 OSPF 패킷이 링크 상태 업데이트 패킷인 유형 4임을 나타냅니다.RFC 2328의 섹션 4.3의 이 표에서는 다음과 같은 다양한 OSPF 패킷 유형을 정의합니다.

유형	패킷 이름	프로토콜 함수
1	여보세요	네이버 검색/유지 관리
2	데이터베이스 설명	데이터베이스 내용 요약
3	링크 상태 요청	데이터베이스 다운로드
4	링크 상태 업데이트	데이터베이스 업데이트
5	링크 상태 ACK	플러딩 승인

OSPF 인접성이 'FULL' 상태에 도달합니다.

```
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
```

다음으로, R2의 IOS를 Cisco 버그 ID CSCse01519의 IOS 버전으로 업그레이드합니다.

```
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 인접성이 'LOADING' 상태로 유지되며 'FULL' 상태에 도달하지 않습니다.재전송은 OSPF가 재전송 한도인 25개에 도달할 때까지 발생합니다.OSPF는 인접성을 다시 설정하려고 시도하며, 동일한 문제가 다시 발생하고 루프가 끊임없이 계속됩니다.

따라서 R2의 업그레이드는 이전에 숨겨진 문제를 해결합니다.기본 MTU는 OSPF 라우터에서 사용하는 것보다 작습니다.

스위치가 MTU를 2000으로 변경하면 1500바이트보다 큰 OSPF 패킷('l:1980')이 문제 없이 전송됩니다.

```
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
```

기본 MTU 문제를 확인하려면 항상 MTU 및 DF(조각화 안 함) 비트 세트와 같은 크기의 OSPF 인접 디바이스 IP 주소를 ping합니다.

기본 MTU의 값을 검색하려면 ping을 수행하고 크기를 청소합니다.올바른 MTU를 결정하기 위해 출력에서 느낌표(!)의 수를 계산합니다.이 예에서는 ping 명령의 마지막 에코 회신의 **크기**가 1500바이

트입니다.

```
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
```