

Note technique d'emballage OSPF, MTU et LSA

Contenu

[Introduction](#)

[Taille de paquet OSPF](#)

[MTU dans le paquet DBD](#)

[Comportement OSPF et emballage des LSA dans un paquet de mise à jour LS](#)

[Avant l'ID de bogue Cisco CSCse01519](#)

[Après l'ID de bogue Cisco CSCse01519](#)

[ID de bogue Cisco CSCse01519](#)

[Aperçu](#)

[Scénario](#)

Introduction

Ce document décrit l'interaction des paquets OSPF (Open Shortest Path First), MTU (Maximum Transition Unit), LSA (Link State Advertisements) et LS (Link State Update) dans le contexte de l'ID de bogue Cisco [CSCse01519](#).

Taille de paquet OSPF

Les liaisons sur les routeurs ont un MTU. Les paquets sortants, tels que les paquets OSPF, ne peuvent pas être plus grands que le MTU de l'interface.

[Request for Comments \(RFC\) 2328](#) documente la version 2 du protocole OSPF. L'annexe A.1 de la RFC 2328 décrit l'encapsulation des paquets OSPF de cette manière :

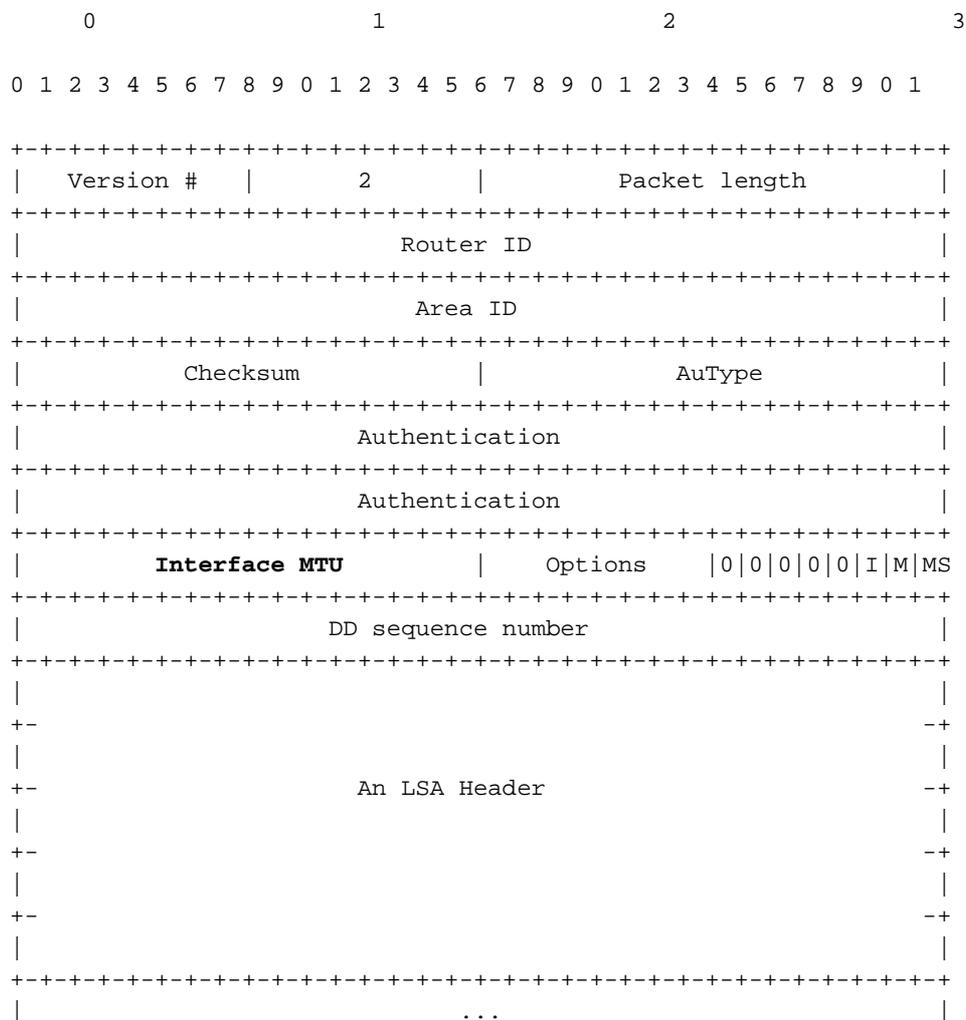
Le protocole OSPF s'exécute directement sur la couche réseau du protocole Internet. Les paquets OSPF sont donc encapsulés uniquement par des en-têtes IP et de liaison de données locale.

Le protocole OSPF ne définit pas de manière à fragmenter ses paquets de protocole et dépend de la fragmentation IP lors de la transmission de paquets plus importants que le MTU du réseau. Si nécessaire, la longueur des paquets OSPF peut atteindre 65 535 octets (y compris l'en-tête IP). Les types de paquets OSPF qui sont susceptibles d'être volumineux (paquets de description de base de données, requête d'état de liaison, mise à jour d'état de liaison et paquets d'accusé de réception d'état de liaison) peuvent généralement être divisés en plusieurs paquets de protocole distincts, sans perte de fonctionnalité. Il est recommandé de le faire ; La fragmentation IP doit être évitée autant que possible.

Il peut y avoir une ou plusieurs LSA dans un paquet de mise à jour LS. De nombreuses LSA d'un paquet de mise à jour LS sont connues sous le nom de lot de LSA dans un paquet de mise à jour LS.

MTU dans le paquet DBD

Le paquet DBD (Database Description), également spécifié dans la RFC 2328, décrit le contenu de la base de données d'état des liaisons OSPF :



Appendice A.3.3 de RFC 2328 décrit le MTU de l'interface comme suit :

Taille en octets du datagramme IP le plus grand qui peut être envoyé à l'interface associée, sans fragmentation.

Les routeurs connectés à une liaison échangent leur valeur MTU d'interface dans des paquets DBD lorsque la contiguïté OSPF est initialisée.

La section 10.6 de la RFC 2328 stipule :

Si le champ MTU de l'interface dans le paquet Description de la base de données indique une taille de datagramme IP supérieure à ce que le routeur peut accepter sur l'interface de réception sans fragmentation, le paquet Description de la base de données est rejeté.

Lorsque la commande **debug ip ospf adj** est utilisée, vous pouvez voir l'arrivée de ces paquets DBD.

Dans cet exemple, les valeurs MTU ne correspondent pas entre deux voisins OSPF. Ce routeur a 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'autre routeur OSPF possède l'interface 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
```

Les paquets DBD sont retransmis en continu jusqu'à ce que la contiguïté OSPF soit finalement désactivée.

```
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
```

Comportement OSPF et emballage des LSA dans un paquet de mise à jour LS

Avant l'ID de bogue Cisco CSCse01519

Avant l'ID de bogue Cisco [CSCse01519](#), OSPF dans le logiciel Cisco IOS® a construit des paquets OSPF ne dépassant pas 1 500 octets, quelle que soit la MTU de l'interface. Ainsi, si la MTU de l'interface était supérieure à 1 500 octets, OSPF ne contenait que jusqu'à 1 500 octets dans un paquet OSPF. Cela était quelque peu inefficace car le protocole OSPF pouvait envoyer des paquets plus importants sur la liaison et obtenir un débit plus important.

Note: Il y avait une exception à ce scénario. Si une LSA contenait plus de 1 500 octets, OSPF a généré ce paquet, quelle que soit sa taille, car OSPF ne peut pas fragmenter une LSA. La pile IP du routeur a ensuite fragmenté le paquet afin de tenir compte de la MTU de l'interface sortante. Cela se produit généralement lorsqu'un routeur OSPF possède de nombreuses liaisons et que la LSA du routeur est devenue plus grande que la MTU de liaison.

De même, si la MTU de l'interface sortante était inférieure à 1 500 octets, le processus OSPF a toujours généré ou empaqueté des paquets OSPF jusqu'à 1 500 octets, et la pile IP du routeur a fragmenté le paquet en paquets IP plus petits afin de tenir compte de la MTU de la liaison sortante. Cela se produit généralement avec un tunnel IPSec entre deux routeurs qui exécutaient OSPF. La surcharge ajoutée des octets d'encapsulation du tunnel a conduit à un MTU inférieur à 1 500 octets. OSPF a généré des paquets OSPF pouvant atteindre 1 500 octets et les paquets ont été fragmentés avant que le routeur ne les transmette. C'était une inefficacité supplémentaire.

Après l'ID de bogue Cisco CSCse01519

Après l'ID de bogue Cisco [CSCse01519](#), OSPF dans IOS peut emballer des paquets OSPF de plus de 1 500 octets. Cela se produit si la MTU de l'interface sortante est supérieure à 1 500 octets. Les transmissions sont plus efficaces car plus d'informations peuvent être regroupées dans un paquet plus grand. En d'autres termes, si un routeur OSPF doit transmettre de nombreuses LSA externes à un voisin OSPF, il peut regrouper davantage de LSA externes dans un paquet de mise à jour LS si ce routeur exécute IOS avec l'ID de bogue Cisco CSCse01519 implémenté.

L'ID de bogue Cisco CSCse01519 permet également au protocole OSPF de générer des paquets de moins de 1 500 octets. Dans certains scénarios, la MTU entre deux voisins OSPF est inférieure à 1 500 octets. Dans l'exemple précédent avec un tunnel IPSec, OSPF transmet des paquets OSPF de moins de 1 500 octets et évite la fragmentation IP ; encore une fois, l'exception est le cas d'une LSA qui est plus grande que la MTU de l'interface.

ID de bogue Cisco CSCse01519

Lorsque vous mettez à niveau un routeur OSPF, vous pouvez découvrir un problème de MTU OSPF causé par l'ID de bogue Cisco [CSCse01519](#).

Aperçu

De nombreux réseaux ont des voisins OSPF qui sont connectés via un réseau commuté de couche 2 (L2), ou réseau de transport, composé d'un service VPN de couche 2 ou d'un réseau SDH/SONET (Synchronous Digital Hierarchy/Synchronous Optical Network). Ces réseaux de transport peuvent avoir des paramètres MTU différents des routeurs qui exécutent OSPF.

Bien que le paramètre MTU doive être correct sur tous les routeurs et refléter le vrai MTU, il y a souvent des erreurs qui passent inaperçues.

Voici un exemple de réseau avec deux routeurs qui exécutent le protocole OSPF. Les routeurs 1 (R1) et 2 (R2) sont connectés via un commutateur L2.

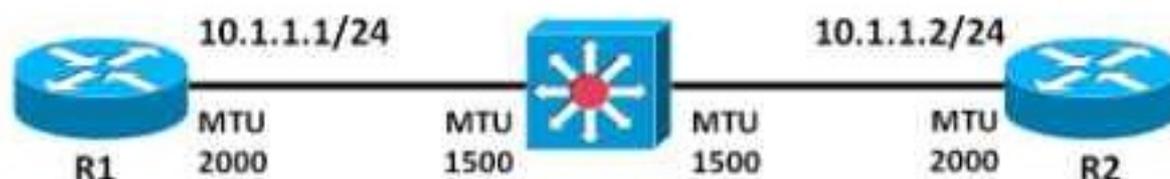


Figure 1 : Example network

Dans cet exemple, les routeurs ont des interfaces GigabitEthernet avec un MTU défini sur 2000. Le MTU du commutateur L2 est de seulement 1 500 octets.

Si la taille du trafic de données n'est jamais supérieure à 1 500 octets, vous pouvez utiliser IOS sans l'ID de bogue Cisco [CSCse01519](#) parce que les paquets OSPF ne sont jamais supérieurs à 1 500 octets. Cependant, s'il existe une LSA de 1 800 octets, par exemple, le processus OSPF sur R1 ou R2 crée un paquet LS Update de plus de 1 500 octets et le transmet, mais le paquet est abandonné par le commutateur L2 entre les routeurs.

Si la base de données OSPF sur R2 dispose de suffisamment de réseaux, les LSA d'origine locale sont si grandes qu'un paquet de mise à jour LS peut être plus grand que le MTU de l'interface.

- Si ces réseaux proviennent de la commande `cover network`, les réseaux apparaissent dans la LSA du routeur de R2. R2 crée une LSA de routeur de plus de 2 000 octets et la transmet, mais IP la fragmente à 2 000 octets, le MTU de l'interface. Le commutateur L2 abandonne cependant ces paquets. OSPF retransmet ensuite ce paquet sans fin et l'état de contiguïté OSPF n'est jamais plein. Ainsi, le problème est immédiatement découvert, même lorsque vous exécutez IOS sans l'ID de bogue Cisco CSCse01519.
- Si ces réseaux proviennent de la commande **redistribute connected**, les réseaux apparaissent dans des LSA externes. Le protocole OSPF tente de regrouper les LSA externes dans un paquet de mise à jour LS de taille maximale de 1 500 octets. Dans ce cas, comme la MTU de l'interface est de 2 000 octets, la contiguïté OSPF atteint l'état 'FULL'. La question d'une MTU sous-jacente inadéquate n'est pas immédiatement découverte. Le problème sera détecté lorsqu'un routeur est mis à niveau vers IOS avec l'ID de bogue Cisco CSCse01519.

Scénario

Supposons que les deux routeurs exécutent une version IOS sans l'ID de bogue Cisco [CSCse01519](#).

Lorsque la contiguïté OSPF se crée, notez que R1 ne reçoit jamais un paquet OSPF de plus de 1 500 octets, bien que la MTU des interfaces soit 2 000.

Activez la commande `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
...
```

Dans cette sortie de débogage, 'l:1468' est la longueur du paquet OSPF, vous pouvez donc voir que le plus grand paquet OSPF était de 1468 octets. 't:4' indique que le paquet OSPF est de type 4, qui est un paquet de mise à jour d'état de liens. Ce tableau de la section 4.3 de la RFC 2328 définit les différents types de paquets OSPF :

Type	Nom du paquet	Fonction de protocole
1	hello	Détecter/gérer les voisins
2	Description de la base de données	Résumer le contenu de la base de données
3	Demande d'état de liaison	Téléchargement de base de données
4	Mise à jour de l'état des liaisons	Mise à jour de base de données
5	Accusé de réception d'état de liaison	Accusé de réception d'inondation

La contiguïté OSPF atteint l'état 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

Ensuite, mettez à niveau IOS sur R2 vers une version IOS avec l'ID de bogue 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 contiguïté OSPF est bloquée dans l'état 'LOADING' et n'atteint pas l'état 'FULL'. Les retransmissions se produisent jusqu'à ce que le protocole OSPF atteigne sa limite de 25 retransmissions. Le protocole OSPF tente de rétablir la contiguïté, le même problème se reproduit et la boucle se poursuit sans fin.

Ainsi, la mise à niveau sur R2 détecte un problème précédemment masqué : la MTU sous-jacente est plus petite que celle utilisée par les routeurs OSPF.

Lorsque le commutateur change la MTU en 2000, un paquet OSPF de plus de 1 500 octets ('1:1980') est transmis sans problème.

```
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
```

Afin de vérifier les problèmes de MTU sous-jacents, envoyez toujours une requête ping à l'adresse IP du voisin OSPF avec une taille égale au MTU et au bit DF (ne pas fragmenter) défini.

Afin de découvrir la valeur du MTU sous-jacent, exécutez une requête ping et balayez la taille. Comptez le nombre de points d'exclamation (!) dans le résultat afin de déterminer la MTU correcte. Dans cet exemple, la dernière réponse d'écho de la commande ping a une taille de 1 500 octets.

```
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
```