¿Puede una tormenta de paquetes ARP afectar a las sesiones de BFD en la plataforma Nexus 7000?

Contenido

Introducción

Este documento describe el impacto de la tormenta de paquetes ARP en los protocolos del plano de control como BFD, OSPF y otros, que se ejecutan en los switches Nexus 7000.

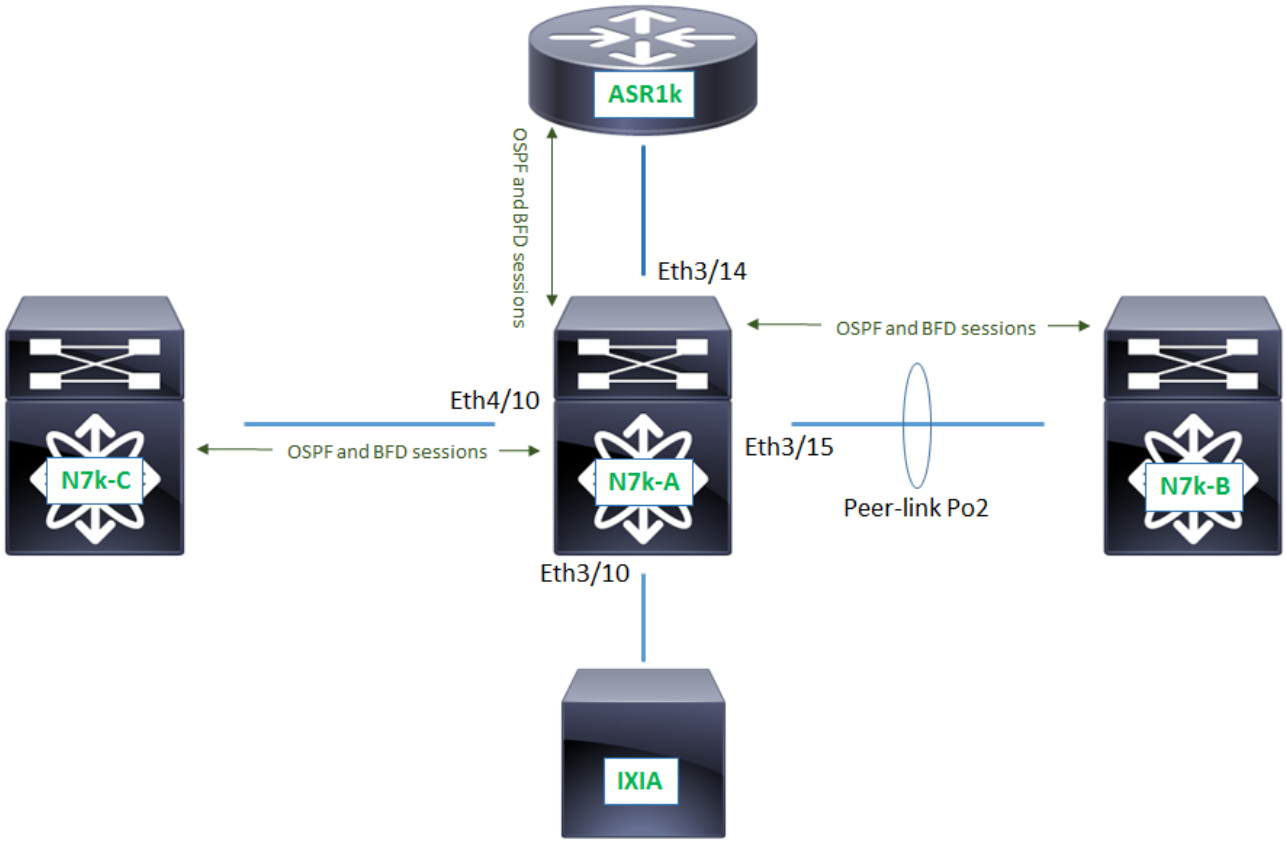
Colaborado por Nishad Mohiuddin, Nikolay Kartashev, Ingenieros del TAC de Cisco.

P. Dado que Cisco NX-OS puede distribuir la operación BFD a módulos compatibles que admiten BFD, ¿una tormenta de paquetes ARP tendría algún impacto en las sesiones BFD en la plataforma Nexus 7000?

A. En general, una tormenta de paquetes ARP puede tener un impacto negativo en la estabilidad de las sesiones BFD que se ejecutan en el switch Nexus 7000. Los síntomas exactos dependen de la longitudad y magnitud del evento ARP Packet Storm. A continuación se muestran los resultados de las pruebas de la red de laboratorio de Cisco TAC.

Detalles de la configuración del laboratorio

La siguiente configuración de laboratorio está diseñada para probar el impacto de las cantidades de tráfico ARP que afectan a la CPU del switch Nexus 7000.



Aquí se utiliza N7k-A como dispositivo sometido a prueba (DUT). DUT es un switch Nexus 7009 con la siguiente configuración de hardware

```
N7k-A# show module
Mod Ports Module-Type Model Status
--- ----- ---------------------------------- ------------------ ----------
1 0 Supervisor module-1X N7K-SUP1 active *
2 0 Supervisor module-1X N7K-SUP1 ha-standby
3 32 10 Gbps Ethernet Module N7K-M132XP-12 ok
4 32 10 Gbps Ethernet Module N7K-M132XP-12 ok
N7k-A#
```

N7k-A tiene conectados los siguientes dispositivos

- N7k-B es un peer VPC, conectado a la interfaz Ethernet 3/15
- ASR1k es un vecino de Capa 3, conectado a la interfaz Ethernet 3/14
- N7k-C es un vecino de Capa 3, conectado a la interfaz Ethernet 4/10
- El generador de tráfico IXIA se encuentra en vlan 6, conectado a la interfaz Ethernet 3/10, que se configura como puerto de acceso de capa 2

DUT tiene tres sesiones BFD, una en la tarjeta de línea en la ranura 4 hacia N7k-C y dos en la tarjeta de línea en la ranura 3 hacia N7k-B y ASR1k

```
N7k-A# show bfd neighbors

OurAddr NeighAddr LD/RD RH/RS Holdown(mult) State Int
10.80.6.173 10.80.6.174 1090519061/4105 Up 4951(3) Up Eth3/14

10.80.1.162 10.80.1.161 1090519054/1090519044 Up 4203(3) Up Eth4/10

10.80.1.61 10.80.1.62 1090519060/1090519059 Up 5921(3) Up Vlan6
```

```
N7k-A#
```

El DUT también tiene tres sesiones OSPF, una en la tarjeta de línea en la ranura 4 hacia N7k-C y dos en la tarjeta de línea en la ranura 3, hacia N7k-B y ASR1k.

```
N7k-A# show ip ospf neighbors
OSPF Process ID 1
Total number of neighbors: 3
 Neighbor ID Pri State Up Time Address Interface
 10.80.0.2 1 FULL/ - 00:13:26 10.80.1.62 Vlan6
10.80.4.25 1 FULL/DR 00:12:40 10.80.6.174 Eth3/14
10.80.0.3 1 FULL/DR 20:15:07 10.80.1.161 Eth4/10
N7k-A#
```

OSPF se registra con BFD

```
router ospf 1
 bfd
 router-id 10.80.0.1
```

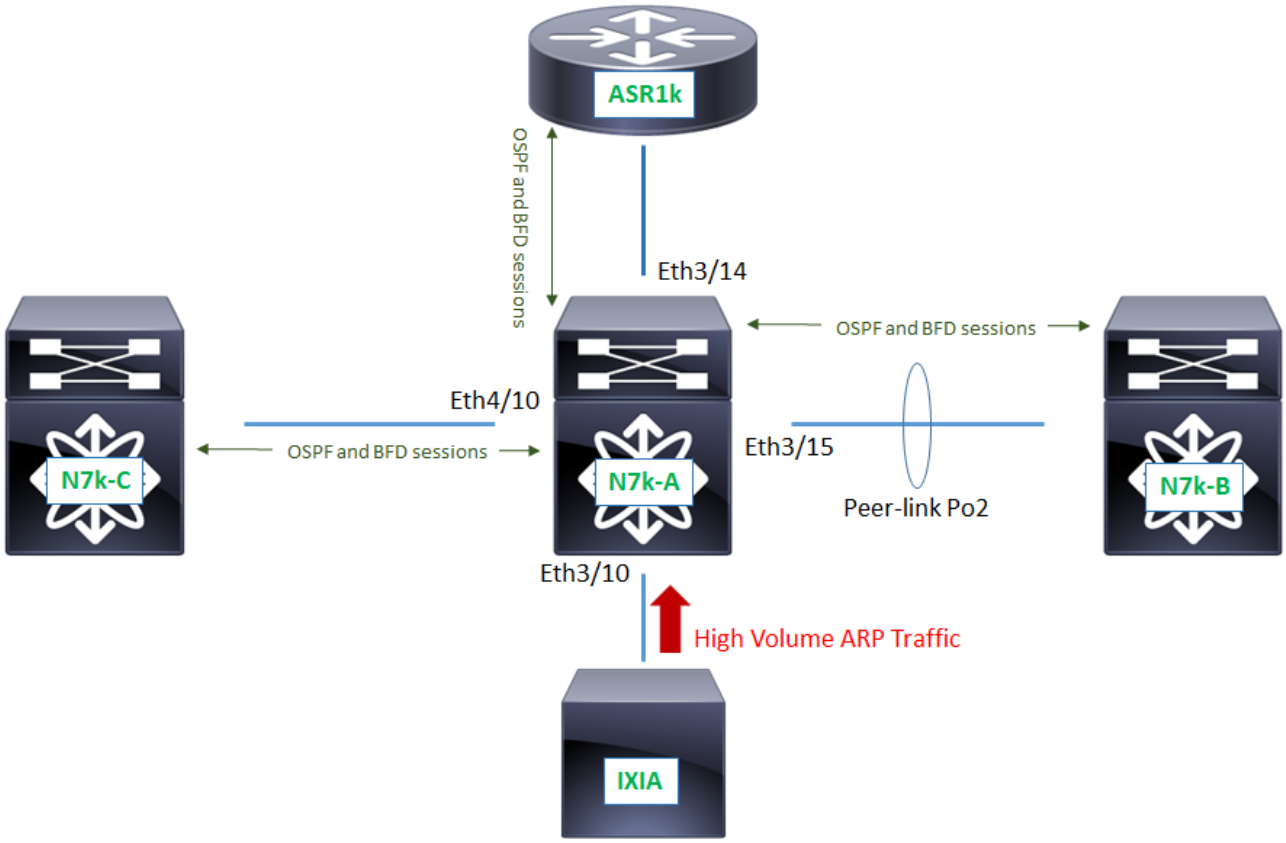
Además, la tabla ARP en N7k-A tiene entradas para los tres vecinos BFD/OSPF

```
N7k-A# show ip arp

Address Age MAC Address Interface
10.80.1.62 00:13:30 4055.390f.48c1 Vlan6
10.80.6.174 00:12:46 88f0.774b.0700 Ethernet3/14
10.80.1.161 00:15:13 6c9c.ed44.6841 Ethernet4/10
N7k-A#
```

Comienza la tormenta ARP

IXIA Traffic Generator se utiliza para simular la parte inestable de la red, lo que da lugar a un gran volumen de tráfico ARP enviado a DUT, como se puede ver en el diagrama siguiente



El siguiente resultado muestra un aumento del tráfico de entrada en la interfaz Ethernet 3/10, donde se conecta el generador de tráfico IXIA. Estos son paquetes ARP de broadcast recibidos en vlan 6

```
N7k-A# show interface Ethernet3/10 | grep "30 seconds input rate"
30 seconds input rate 3102999976 bits/sec, 6062053 packets/sec
N7k-A#
```

Dado que se envía una copia de cada paquete ARP de broadcast a la CPU en N7k-A en este escenario, se observa un aumento de bytes violados en el módulo 3 en CoPP

```
N7k-A# show policy-map interface control-plane class copp-system-p-class-normal
Control Plane

 service-policy input: copp-system-p-policy-strict

 class-map copp-system-p-class-normal (match-any)
 match access-group name copp-system-p-acl-mac-dot1x
 match protocol arp
 set cos 1
 police cir 680 kbps , bc 250 ms
module 3 :
 conformed 2295040 bytes; action: transmit
violated 20569190016 bytes; action: drop

module 4 :
 conformed 128 bytes; action: transmit
violated 0 bytes; action: drop

N7k-A#
```

Nota: Tenga en cuenta que *no hay bytes violados en el módulo en la ranura 4*, ya que el origen de la tormenta ARP de broadcast está conectado a la interfaz solamente en el

módulo 3

En el punto en que comienza la tormenta ARP, los resultados anteriores son generalmente los primeros (y solamente) signos que indican un problema en la red. En la mayoría de los casos, estas señales pasan desapercibidas o son ignoradas por los operadores de red y avanzan rápidamente a una situación que conduce a problemas de conectividad importantes.

La tormenta ARP comienza a afectar al plano de control

De forma predeterminada, el valor de tiempo de espera ARP en la plataforma Nexus 7000 se configura durante 25 minutos o 1500 segundos. El switch Nexus debe actualizar periódicamente las entradas de caché ARP locales para mantener la resolución IP a MAC actualizada de sus vecinos de Capa 3 de salto siguiente.

El siguiente es el resultado de la tabla de caché ARP en DUT después de que las entradas de la memoria caché ARP hayan caducado.

```
N7k-A# show ip arp

Address Age MAC Address Interface
10.80.1.62 00:00:06 INCOMPLETE Vlan6
10.80.6.174 00:00:10 INCOMPLETE Ethernet3/14
10.80.1.161 00:12:59 6c9c.ed44.6841 Ethernet4/10
N7k-A#
```

Observe que las entradas de la memoria caché ARP para los dispositivos conectados a la tarjeta de línea en la ranura 3 muestran el estado **INCOMPLETE**, mientras que la entrada para el switch N7k-C, que está conectado a la tarjeta de línea en la ranura 4 se actualiza correctamente como se esperaba.

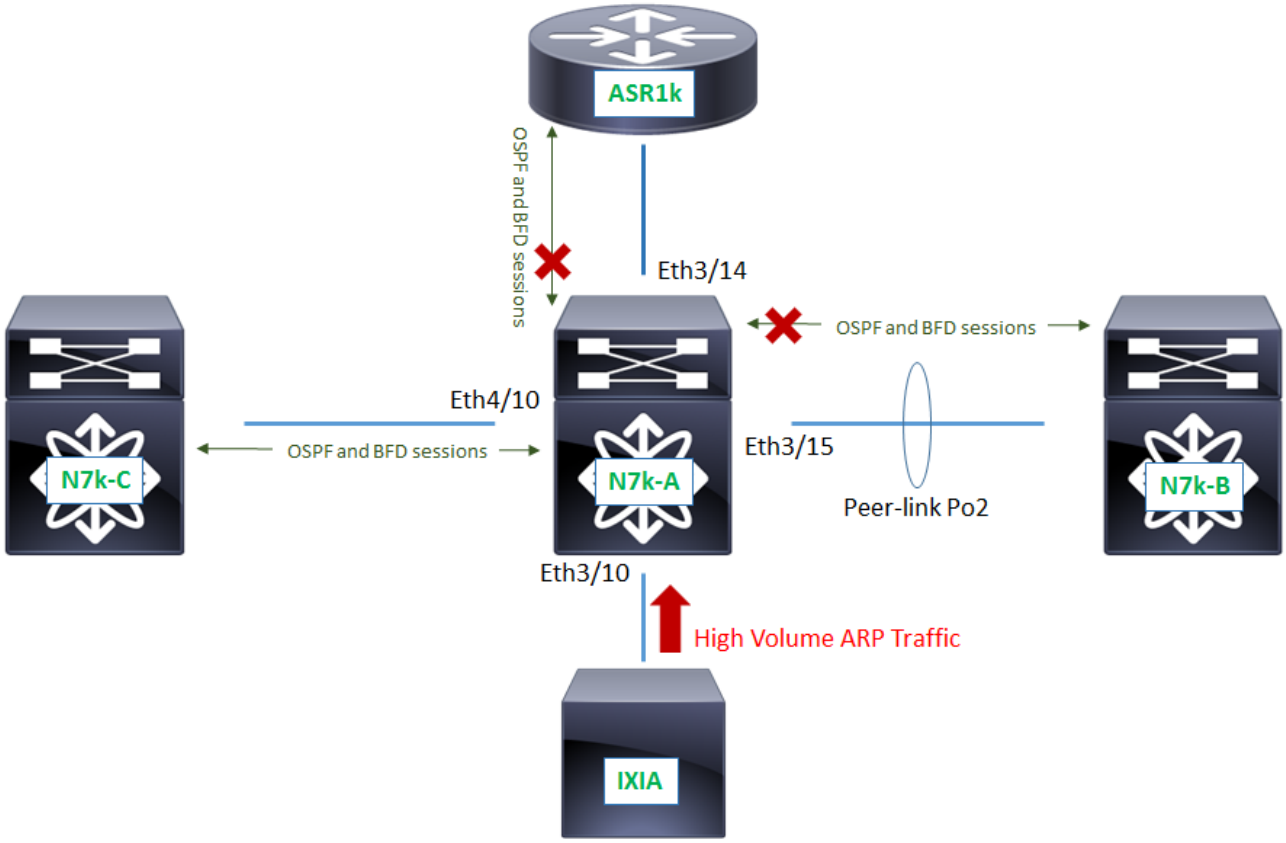
Los siguientes mensajes de registro DUT indican el impacto en el nivel del plano de control

```
N7k-A# show logging log
...
2016 Nov 16 22:12:55 N7k-A %BFD-5-SESSION_STATE_DOWN: BFD session 1090519060 to neighbor
10.80.1.62 on interface Vlan6 has gone down. Reason: 0x3.
2016 Nov 16 22:12:55 N7k-A %OSPF-5-ADJCHANGE: ospf-1 [10600] Nbr 10.80.1.62 on Vlan6 went DOWN
2016 Nov 16 22:12:55 N7k-A %BFD-5-SESSION_REMOVED: BFD session to neighbor 10.80.1.62 on
interface Vlan6 has been removed
2016 Nov 16 22:12:56 N7k-A %OSPF-5-ADJCHANGE: ospf-1 [10600] Nbr 10.80.1.62 on Vlan6 went
EXSTART
2016 Nov 16 22:13:40 N7k-A %OSPF-5-ADJCHANGE: ospf-1 [10600] Nbr 10.80.6.174 on Ethernet3/14
went DOWN
2016 Nov 16 22:13:40 N7k-A %BFD-5-SESSION_STATE_DOWN: BFD session 1090519061 to neighbor
10.80.6.174 on interface Eth3/14 has gone down. Reason: 0x3.
2016 Nov 16 22:13:40 N7k-A %OSPF-5-ADJCHANGE: ospf-1 [10600] Nbr 10.80.6.174 on Ethernet3/14
went EXSTART
2016 Nov 16 22:13:46 N7k-A %BFD-5-SESSION_REMOVED: BFD session to neighbor 10.80.6.174 on
interface Eth3/14 has been removed
2016 Nov 16 22:15:45 N7k-A %OSPF-5-ADJCHANGE: ospf-1 [10600] Nbr 10.80.6.174 on Ethernet3/14
went INIT
...
N7k-A#
```

Observe en este resultado que OSPF cambia entre el estado INACTIVO a EXSTART y luego vuelve al estado INIT. Esto ocurre porque OSPF utiliza unicast para intercambiar prefijos durante

el estado EXSTART. Debido a que la resolución ARP está incompleta en el módulo en la ranura 3 en el momento de la tormenta de paquetes ARP, el intercambio de ruta nunca se completa, lo que resulta en que la adyacencia OSPF no se forma.

> **Nota**:La resolución ARP a IP a MAC del siguiente salto depende de unicast como lo hace el funcionamiento BFD. Dado que podemos concluir que BFD requiere que ARP se resuelva para un funcionamiento adecuado.



Los siguientes resultados confirman el impacto de una tormenta de paquetes ARP en las sesiones BFD y OSPF en el módulo en la ranura 3. A diferencia de esta sesión BFD y OSPF en el módulo en la ranura 4 se establecen y se mantienen estables.

```
N7k-A# show bfd neighbors

OurAddr NeighAddr LD/RD RH/RS Holdown(mult) State Int
10.80.1.162 10.80.1.161 1090519054/1090519044 Up 5764(3) Up Eth4/10

N7k-A#


N7k-A# show ip ospf neighbors
OSPF Process ID 1
Total number of neighbors: 3
 Neighbor ID Pri State Up Time Address Interface
 10.80.0.2 1 EXSTART/ - 00:02:54 10.80.1.62 Vlan6
10.80.4.25 1 INIT/DR 00:00:05 10.80.6.174 Eth3/14
10.80.0.3 1 FULL/DR 20:29:28 10.80.1.161 Eth4/10
N7k-A#
```

¿Qué sucede cuando se detiene una tormenta de paquetes ARP?

Cuando se detiene una tormenta de paquetes ARP, la siguiente recuperación se produce automáticamente y la red comienza a converger y disfruta del estado estable que tenía antes de la tormenta de difusión ARP.

1. Las entradas de caché ARP se resuelven en N7k-A
2. Sesiones BFD en el módulo en el slot 3 restablecer
3. Sesiones OSPF en el módulo en el slot 3 restablecer

Conclusión

Aunque Cisco NX-OS puede distribuir la operación BFD a módulos compatibles que admiten BFD, los grandes volúmenes de tráfico ARP que afectan a la CPU del switch durante un período mayor que el tiempo restante para actualizar las entradas de caché ARP locales en la plataforma Nexus 7000 causarán inestabilidad en las sesiones BFD y en los protocolos de cliente registrados con BFD.

Esto se puede atribuir a la operación BFD que requiere la resolución ARP del salto siguiente que es unicast. Si la entrada de caché ARP para el salto siguiente no se actualiza a tiempo, las sesiones BFD fallarán.