

ARP数据包风暴是否会影响Nexus 7000平台上的BFD会话

目录

[简介](#)

[问：由于Cisco NX-OS可以将BFD操作分发到支持BFD的兼容模块，ARP数据包风暴是否会对Nexus 7000平台上的BFD会话产生任何影响？](#)

[实验设置详细信息](#)

[ARP风暴开始](#)

[ARP风暴开始影响控制平面](#)

[当ARP数据包风暴停止时会发生什么情况？](#)

[结论](#)

简介

本文档介绍ARP数据包风暴对Nexus 7000交换机上运行的控制平面协议（如BFD、OSPF等）的影响。

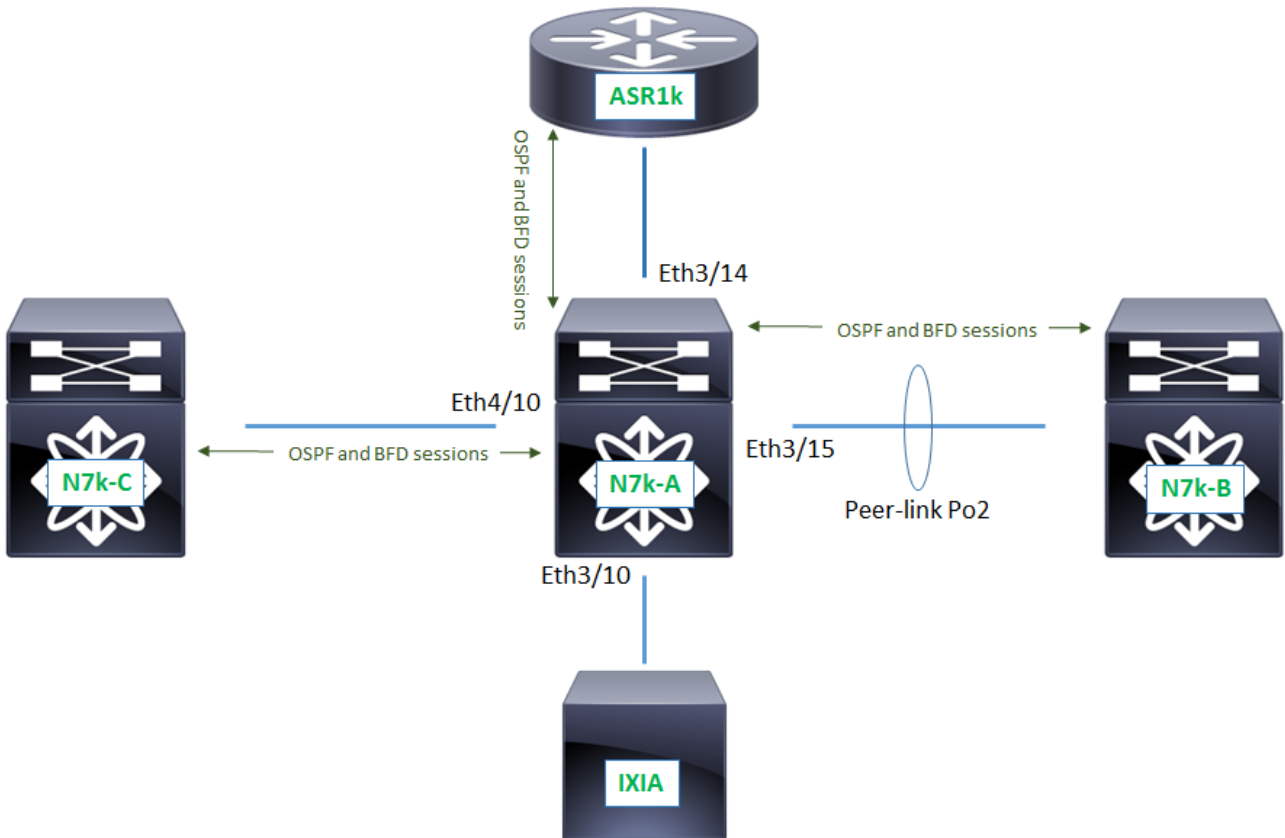
作者：Nishad Mohiuddin、Nikolay Kartashev，Cisco TAC工程师。

问：由于Cisco NX-OS可以将BFD操作分发到支持BFD的兼容模块，ARP数据包风暴是否会对Nexus 7000平台上的BFD会话产生任何影响？

答：通常，ARP数据包风暴对Nexus 7000交换机上运行的BFD会话的稳定性可能产生负面影响。确切的症状取决于ARP数据包风暴事件的寿命和大小。以下是Cisco TAC实验室网络的测试结果。

实验设置详细信息

以下实验设置旨在测试Nexus 7000交换机CPU上ARP流量的影响。



此处N7k-A用作测试设备(DUT)。DUT是具有以下硬件配置的Nexus 7009交换机

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连接了以下设备

- N7k-B是VPC对等体，连接到接口Ethernet 3/15
- ASR1k是第3层邻居，连接到接口Ethernet 3/14
- N7k-C是第3层邻居，连接到接口Ethernet 4/10
- IXIA流量生成器位于vlan 6中，连接到接口Ethernet 3/10，该接口配置为第2层接入端口

DUT有三个BFD会话，一个在插槽4中朝向N7k-C，另一个在插槽3中朝向N7k-B和ASR1k的线卡

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#

DUT还有三个OSPF会话，一个在插槽4中通向N7k-C，另一个在插槽3中通向N7k-B和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已向BFD注册

```

router ospf 1
bfd
router-id 10.80.0.1

```

此外，N7k-A上的ARP表包含所有三个BFD/OSPF邻居的条目

```

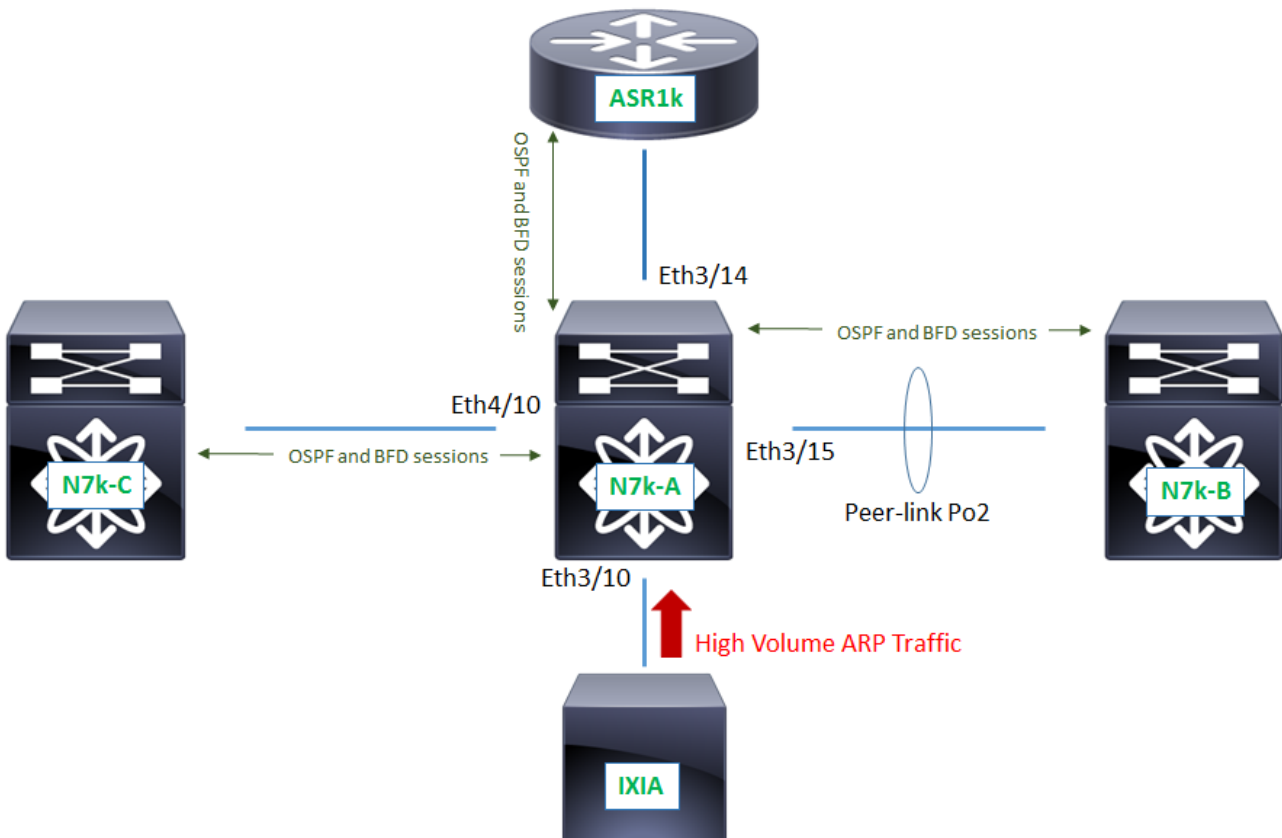
N7k-A# show ip arp

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#

```

ARP风暴开始

IXIA流量生成器用于模拟网络不稳定部分，这会导致大量ARP流量发送到DUT，如下图所示



以下输出显示接口Ethernet 3/10 (IXIA流量生成器连接) 上输入流量的增加。这些是VLAN 6中收到的广播ARP数据包

```
N7k-A# show interface Ethernet3/10 | grep "30 seconds input rate"
```

```
30 seconds input rate 3102999976 bits/sec, 6062053 packets/sec
```

```
N7k-A#
```

由于在此场景中，每个广播ARP数据包的副本都发送到N7k-A上的CPU，因此在CoPP中，模块3上违规的字节数会增加

```
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#
```

注意：请注意，插槽4中的模块上没有违反的字节，因为广播ARP风暴的源仅连接到模块3上的接口

在ARP风暴开始时，上述输出通常是指示网络问题的第一个（且仅有）符号。在大多数情况下，这些标志不会被网络运营商发现或忽略，并迅速进入导致重大连接问题的状态。

ARP风暴开始影响控制平面

默认情况下，Nexus 7000平台上的ARP超时值配置为25分钟或1500秒。Nexus交换机必须定期刷新本地ARP缓存条目，以便保持其下一跳第3层邻居的最新IP到MAC解析。

以下是ARP缓存条目过期后DUT上ARP缓存表的输出。

```
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#
```

请注意，连接到插槽3中线卡的设备的ARP缓存条目显示**INCOMPLETE** 状态，而连接到插槽4中线卡的交换机N7k-C的条目正按照预期成功刷新。

以下DUT日志消息指示对控制平面级别的影响

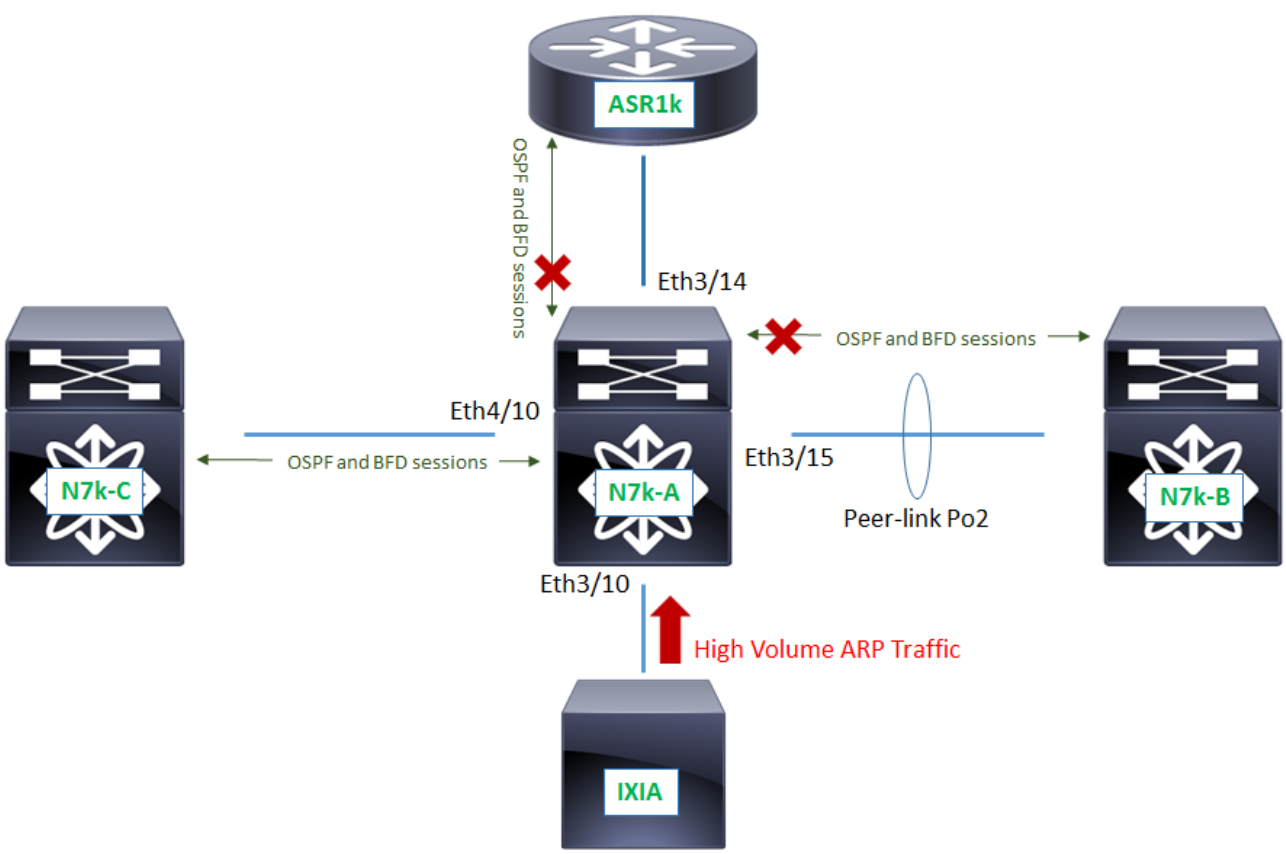
```
N7k-A# show logging log
```

```
...
```

```
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#
```

请注意，在此输出中，OSPF在DOWN状态和EXSTART状态之间切换，然后切换回INIT状态。这是因为OSPF在EXSTART状态期间使用单播交换前缀。由于ARP数据包风暴时插槽3中的模块上的ARP解析不完整，因此路由交换从未完成，导致OSPF邻接关系无法形成。

备注:ARP到IP到MAC的下一跳解析依赖于单播，BFD操作也是如此。假设我们可以断定BFD需要解析ARP才能正常运行。



以下输出确认了ARP数据包风暴对插槽3中模块上的BFD和OSPF会话的影响。与插槽4中模块上的此BFD和OSPF会话相反，ARP数据包风暴已建立并保持稳定。

```
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#

当ARP数据包风暴停止时会发生什么情况？

当ARP数据包风暴停止时，以下恢复会自动发生，网络开始收敛并处于ARP广播风暴之前的稳定状态。

1. 在N7k-A上解析ARP缓存条目
2. 插槽3中模块上的BFD会话重新建立
3. 插槽3中模块上的OSPF会话重新建立

结论

尽管Cisco NX-OS可以将BFD操作分发到支持BFD的兼容模块，但在超过刷新Nexus 7000平台上本地ARP缓存条目的剩余时间内，大量ARP流量进入交换机CPU将导致BFD会话和向BFD注册的任何客户端协议不稳定。

这可归因于BFD操作，它需要单播下一跳的ARP解析。如果下一跳的ARP缓存条目未及时刷新，BFD会话将失败。