排除端口通道负载均衡中的极化故障

目录

[简介](#)
[背景](#)
[先决条件](#)
[拓扑](#)
[配置](#)
[流量传输](#)
[故障排除](#)
[解决方法](#)

简介

本文档说明了端口通道负载均衡中可能出现极化的场景，并提供了如何防止极化的建议。

背景

极化 散列算法选择网络中的某些路径，而使冗余路径未使用

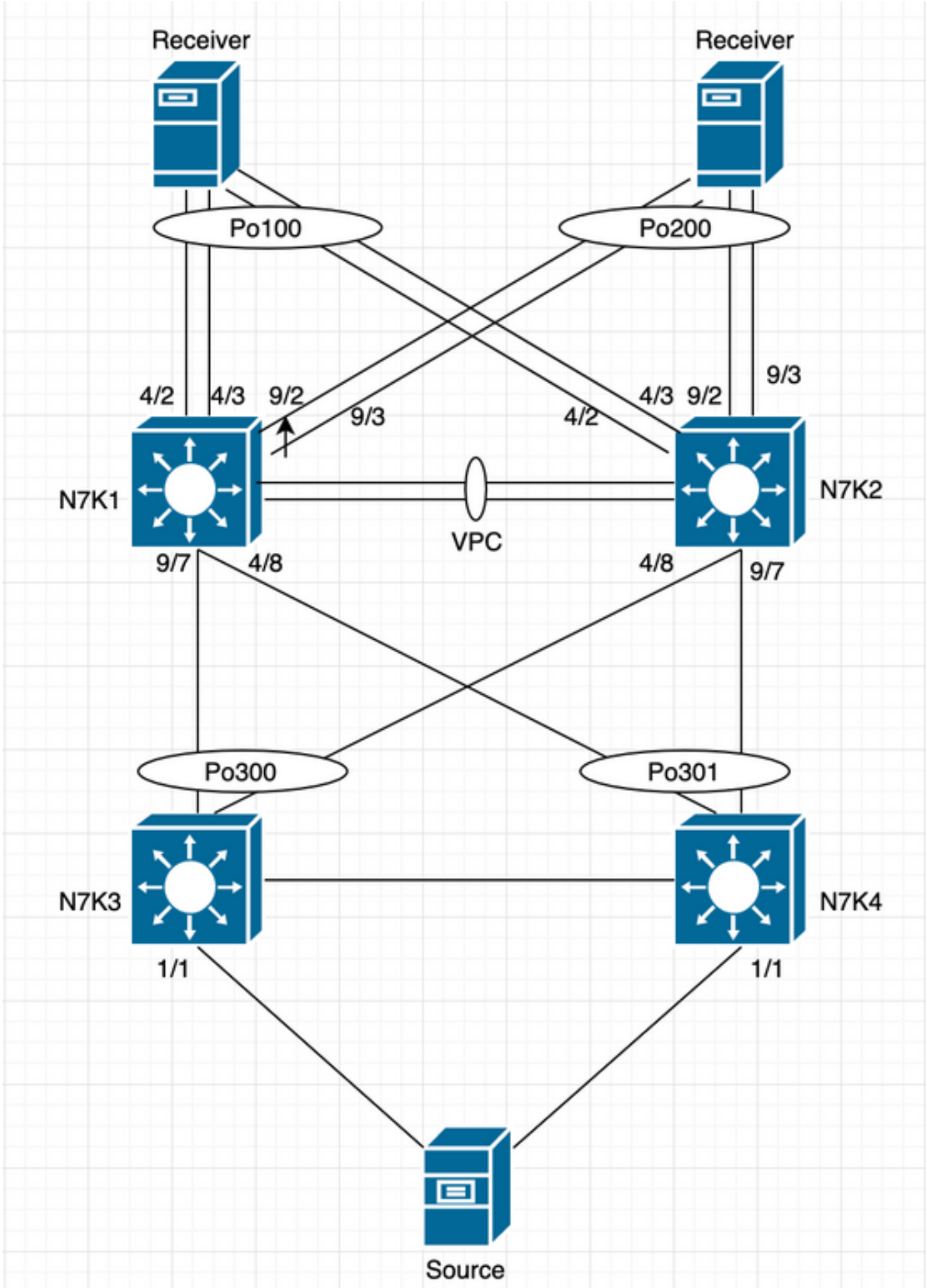
先决条件

建议您了解以下主题。

[链路聚合控制协议](#)

Cisco Nexus平台

拓扑


Receiver
Receiver
Po100
Po200
4/2
4/3
9/2
9/3
4/2
4/3
9/2
9/3
N7K1
N7K2
VPC
9/7
4/8
4/8
9/7
Po300
Po301
N7K3
N7K4
1/1
1/1
Source

配置

N7K1和N7K2连接在VPC中，Po100、Po200、Po300和Po301连接在VPC端口通道中。

N7K1和N7K2充当纯L2交换机，这些交换机上没有路由。

所有交换机都运行相同的端口通道负载均衡算法

无论从源到目的地的流量是在同一VLAN中（无路由），还是在N7K3或N7k4上发生路由时处于不同VLAN中，N7K1和N7K2的流量都会出现极化问题。

流量传输

源设备向目的设备发送多个流（具有多个源和目的IP地址，并且第4层端口信息也因数据包而异）。使用良好的流量组合来确保在理想情况下，流量会均匀地分布在端口通道成员接口之间。

从N7k3/N7k4上的源地址发来的流量，然后通过N7K1/N7K2到达目的地。

N7K1和N7K2上Po100和Po200成员链路中的一条链路发送了近99%的流量，而另一条链路保持空闲状态。（即，在每台交换机N7K1和N7K2上，4/2和4/3中的一条链路传输99%的单播流量，另一条链路传输1%以下，9/2和9/3中的一条链路传输99%的流量，另一条链路传输1%以下。故障排除部分的输出显示po100和po200成员接口上的流量N7K1，在N7K2上可以看到类似输出）。

无论所使用的端口通道负载均衡算法类型如何，只要在N7K1/N7K2对和N7K3/N7K4对上使用相同的端口通道负载均衡算法，问题就可以看到。下面提供了检查端口通道负载均衡算法的命令。

```
N7K1# show port-channel load-balance
Warning: Per Packet Load balance configuration has higher precedence
System config:
  Non-IP: src-dst mac
  IP: src-dst ip-l4port-vlan rotate 0
Port Channel Load-Balancing Configuration for all modules:
Module 1:
  Non-IP: src-dst mac
  IP: src-dst ip rotate 0
Module 2:
  Non-IP: src-dst mac
  IP: src-dst ip rotate 0
Module 3:
  Non-IP: src-dst mac
  IP: src-dst ip rotate 0
Module 4:
  Non-IP: src-dst mac
  IP: src-dst ip-l4port-vlan rotate 0
Module 7:
  Non-IP: src-dst mac
  IP: src-dst ip-l4port-vlan rotate 0
Module 8:
  Non-IP: src-dst mac
  IP: src-dst ip-l4port-vlan rotate 0
Module 9:
  Non-IP: src-dst mac
  IP: src-dst ip-l4port-vlan rotate 0
```

故障排除

如果在端口通道上看到负载均衡不均，则可能是由于极化。

当流量到达N7K3和N7K4交换机时，它们会通过N7K4的Po301和N7K3的Po300转发到N7K1/N7K2交换机。此时，负载均衡算法开始运行，某些流会转发到N7K1，其他流会转发到N7K3k2。

最初，所有流量进入eth1/1上的交换机N7K3/N7K4，并且根据src-dest ip和l4端口信息，在通往N7K1的链路上对某些流进行散列处理，在通往N7K2的链路上对其他流进行散列处理。散列处理基于交换机计算的rbh值。为简单起见，我们假设交换机根据使用的负载均衡算法将传入流量分为两个流（流X和流Y）。 从一个端口通道成员链路发送的流X和从另一个端口通道成员链路发送的流Y。

现在，当流量降落到N7K1/N7K2线对上时，可能有两种情况。（考虑X和Y可互换）

案例1:

N7K3已发送流X至N7K1，流Y至N7K2

和

N7K4已发送流Y到N7K1，流X到N7K2

案例2:

N7K3已发送流X至N7K1，流Y至N7K2

和

N7K4已发送流X至N7K1，流Y至N7K2

在案例1中，N7K1和N7K2接收两种类型的流（流X和流Y），即使使用与N7K3/N7K4相同的端口通道负载均衡算法，也不会看到任何极化，因为流从Po100和Po200流出在不同的链路上，因此，我们看到在端口通道成员接口之间更好地分配流量。

在案例2中，N7K1仅接收流X，N7K2仅接收流Y，如果这些交换机上使用的端口通道负载均衡算法与N7K3/N7K4对中使用的算法相同，这可能会产生极化。由于N7K1和N7K2使用相同的端口通道负载均衡算法，N7K1仅在Po100/Po200的一个成员链路上发送流X，而另一个成员链路不会转发任何流量。同样，N7K2仅在Po100/Po200的一个成员链路上发送流Y，而另一个成员链路不会转发任何流量。

由于交换机N7K1和N7K2接收的流量已分类为开始，因此仅使用一个端口通道成员链路将所有传入流量从交换机N7K1/N7K2发送出去，而不会从另一成员链路发送任何流量。如果传入流量速率超过单个端口通道链路的带宽，其他流量可能会被丢弃，因为其他端口通道成员链路不会转发此流量。

当端口通道中使用两个以上的链路时，也会出现类似问题。例如，如果在端口通道中使用四条链路，则根据散列的发生，不会发生极化，或者我们会看到部分极化，其中只有四条端口通道成员链路中的两条用于转发所有传入流量，而另外两条链路不会转发任何内容

偏振是由设计引起的，因此分析设计必须确保不发生偏振。下面给出了指示N7k1上发生极化的Po100和Po200的输出（N7K2上也可看到类似的输出）。

```
N7K1# show port-channel summary | i 200
200   Po200(SU)   Eth      LACP      Eth9/2(P)    Eth9/3(P)
```

```
N7K1# show port-channel traffic interface port-channel 200
NOTE: Clear the port-channel member counters to get accurate statistics
ChanId      Port Rx-Ucst Tx-Ucst Rx-Mcst Tx-Mcst Rx-Bcst Tx-Bcst
------ --------- ------- ------- ------- ------- ------- -------
   200    Eth9/2    0.0%  99.99%  44.44%   4.00%    0.0% 100.00%
   200    Eth9/3    0.0%   0.00%  55.55%  96.00%    0.0%    0.0%

N7K1# show port-channel summary | i 100
100   Po100(SU)   Eth      LACP      Eth4/2(P)    Eth4/3(P)

N7K1# show port-channel traffic interface port-channel 100
NOTE: Clear the port-channel member counters to get accurate statistics
ChanId      Port Rx-Ucst Tx-Ucst Rx-Mcst Tx-Mcst Rx-Bcst Tx-Bcst
------ --------- ------- ------- ------- ------- ------- -------
   100    Eth4/2    0.0%  99.99%  40.55%   7.00%    0.0% 100.00%
   100    Eth4/3    0.0%   0.00%  54.44%  93.00%    0.0%    0.0%
```

CSCvq26885 已归档以获取外部文档。

解决方法

遵循一些可用于确保不发生极化的解决方法。

1.正确设计：由于极化的主要原因是设计不当，因此最好确保我们更改网络设计以确保拓扑中没有极化空间

如果设计不可能更改，我们可以执行以下操作。

2.在每级交换机上使用不同的端口通道负载均衡算法（N7K1/N7k2对上使用一种算法，N7K3/N7k4对上使用不同算法）。 当负载均衡算法发生更改时，N7k1/N7k2交换机现在根据N7k3/N7k4交换机使用的其它信息以外的信息对传入流量进行散列处理，因此传出流量使用所有端口通道成员链路。（选择哪种算法取决于交换机接收的流量类型）

3.如果客户希望使用相同的负载均衡算法，请在每级交换机上使用不同的旋转值。Rotate命令通过偏移用户配置的字节来引入散列算法的随机性，并有助于避免极化。（对N7k1/N7k2线对使用一个旋转值，对N7k3/N7k4线对使用不同的旋转值）