

Solucionar problemas de polarização no balanceamento de carga do canal de porta

Contents

[Introduction](#)

[Background](#)

[Prerequisites](#)

[Topologia](#)

[Configuração](#)

[Fluxo de tráfico](#)

[Troubleshooting](#)

[Solução](#)

Introduction

Este documento explica os cenários sob os quais pode ocorrer polarização no balanceamento de carga do canal de porta e fornece sugestões sobre como evitá-los.

Background

Polarização é um problema em que o algoritmo de hash seleciona certos caminhos na rede e deixa os caminhos redundantes não utilizados

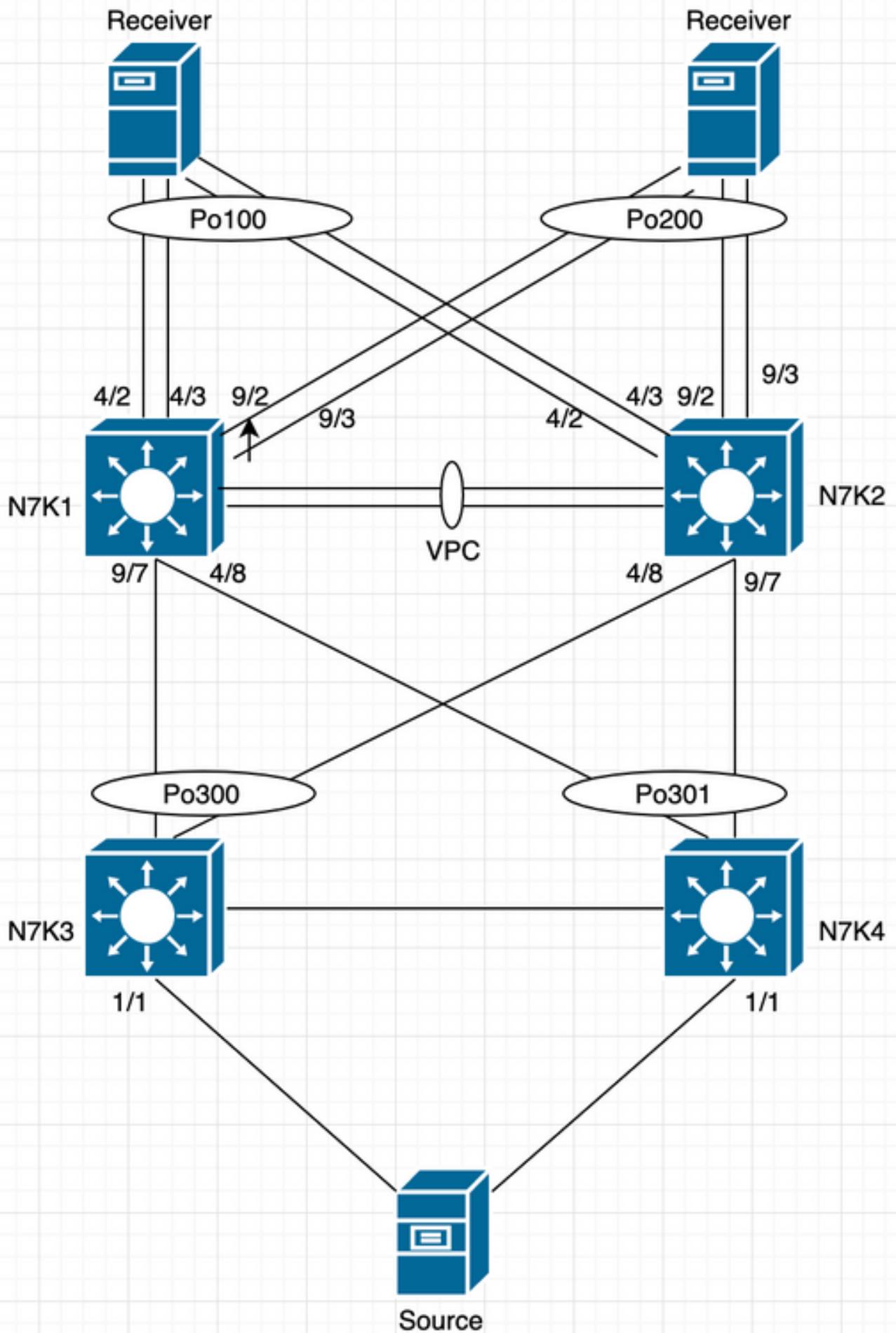
Prerequisites

Recomenda-se ter conhecimento sobre os tópicos a seguir.

[Protocolo de controle de agregação de link](#)

Plataformas Cisco Nexus

Topologia



Configuração

N7K1 e N7K2 conectados em VPC e Po100, Po200, Po300 e Po301 estão no canal de porta VPC.

O N7K1 e o N7K2 atuam como um switch L2 puro, sem que ocorra roteamento nesses switches.

Todos os switches estão executando o mesmo algoritmo de balanceamento de carga do canal de porta

O problema de polarização é observado no tráfego que sai de N7K1 e N7K2, independentemente de o tráfego da origem para o destino estar na mesma vlan (sem roteamento) ou se eles estavam em uma vlan diferente com o roteamento acontecendo em N7K3 ou N7k4.

Fluxo de tráfico

A origem envia vários fluxos para o destino (com vários endereços IP de origem e destino, e as informações da porta L4 também variam de pacote para pacote). Uma boa combinação de tráfego é usada para garantir que, em uma situação ideal, o tráfego seja distribuído uniformemente entre as interfaces membro do canal de porta.

O tráfego da terra de origem em N7k3/N7k4 e depois vai através de N7K1/N7K2 para o destino.

Um link entre os links de membros de Po100 e Po200 em cada um dos N7K1 e N7K2 está enviando quase 99% do tráfego e o outro link permanece ocioso. (ou seja, em cada switch N7K1 e N7K2, um link entre 4/2 e 4/3 transporta 99% de tráfego unicast e o outro link transporta menos de 1%, da mesma forma que um link entre 9/2 e 9/3 transporta 99% de tráfego e o outro link transporta menos de 1%. a saída na seção Troubleshooting mostra o tráfego nas interfaces dos membros po100 e po200 em N7K1, uma saída semelhante pode ser vista em N7K2).

Independentemente do tipo de algoritmo de balanceamento de carga do canal de porta usado, o problema pode ser visto desde que o mesmo algoritmo de balanceamento de carga do canal de porta seja usado no par N7K1/N7K2 e no par N7K3/N7K4. O comando para verificar o algoritmo de balanceamento de carga do canal de porta é fornecido abaixo.

```
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
```

Troubleshooting

Se o balanceamento de carga irregular for visto em um canal de porta, ele pode ser devido à polarização.

Quando o tráfego chega aos switches N7K3 e N7K4, eles são encaminhados aos switches N7K1/N7K2 através do Po301 do N7K4 e Po300 do N7K3. Aqui, o algoritmo de balanceamento de carga entra e alguns fluxos são encaminhados para N7K1 e outros fluxos encaminhados para N7K2.

Inicialmente, todo o tráfego chega aos switches N7K3/N7K4 em eth1/1 e, com base nas informações de porta src-dest ip e l4, certos fluxos são colocados no link indo para N7K1 e outros fluxos são colocados no link indo para N7K2. O hash é feito com base no valor rbh calculado pelo switch. Para simplificar, vamos supor que, com base no algoritmo de balanceamento de carga usado, o switch segregue o tráfego de entrada em dois fluxos (fluxo X e fluxo Y). O fluxo X foi enviado de um link de membro do canal de porta e o fluxo Y foi enviado para fora do outro link de membro do canal de porta.

Agora, quando o tráfego pousa no par N7K1/N7K2, pode haver duas possibilidades. (Considerando que X e Y são intercambiáveis)

Caso 1:

N7K3 enviou fluxo X para N7K1 e fluxo Y para N7K2

e

O fluxo enviado de N7K4 para N7K1 e o fluxo X para N7K2

Caso 2:

N7K3 enviou fluxo X para N7K1 e fluxo Y para N7K2

e

N7K4 enviou fluxo X para N7K1 e fluxo Y para N7K2

No Caso 1, N7K1 e N7K2 recebem ambos os tipos de fluxos (fluxo X e fluxo Y) e mesmo depois de usar o mesmo algoritmo de balanceamento de carga do canal de porta usado pelo N7K3/N7K4, nenhuma polarização seria vista como a saída de fluxos do Po100 e Po200 em links diferentes e, portanto, vemos uma melhor distribuição de tráfego entre os membros do canal de porta interfaces.

No Caso 2, o N7K1 recebe somente o fluxo X e N7K2 recebe somente o fluxo Y e isso pode criar polarização se o algoritmo de balanceamento de carga do canal de porta usado nesses switches

for o mesmo usado no par N7K3/N7K4. Como N7K1 e N7K2 estão usando o mesmo algoritmo de balanceamento de carga de canal de porta, N7K1 envia o fluxo X em apenas um link membro do Po100/Po200 e o outro link membro não encaminha nenhum tráfego. Da mesma forma, o N7K2 envia o fluxo Y em apenas um link membro do Po100/Po200 e o outro link membro não encaminhará nenhum tráfego.

Como o tráfego que os switches N7K1 e N7K2 estão recebendo já está classificado para começar, somente um link membro do canal de porta será usado para enviar todo o tráfego de entrada do switch N7K1/N7K2 e nada será enviado do outro link membro. Caso a taxa de tráfego de entrada exceda a largura de banda do link de canal de porta único, o tráfego adicional pode ser descartado, já que o outro link de membro de canal de porta não encaminharia esse tráfego.

Um problema semelhante pode ser observado quando mais de dois links são usados no canal de porta. Por exemplo, se quatro enlaces forem usados em um canal de porta, dependendo do hash que está ocorrendo, não haverá polarização ou veremos polarização parcial onde apenas dois dos quatro enlaces de membro do canal de porta seriam usados para encaminhar todo o tráfego de entrada e os outros dois enlaces não encaminharão nada

A polarização é causada por causa do projeto e, portanto, é importante analisar o projeto para garantir que não ocorra polarização. A saída que indica polarização ocorrendo Po100 e Po200 em N7k1 é fornecida abaixo (a saída semelhante também pode ser vista em N7K2).

```
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](#) foi arquivado para documentação externa.

Solução

Seguindo algumas das soluções alternativas que podem ser usadas para garantir que a polarização não ocorra.

1. Design adequado: Como a principal causa da polarização é o design inapropriado, é melhor ter certeza de que alteramos o projeto de rede para garantir que não haja espaço para polarização na topologia

Se nenhuma alteração no projeto for possível, podemos fazer o seguinte.

2. Use diferentes algoritmos de balanceamento de carga do canal de porta em cada nível de switches (um algoritmo no par N7K1/N7k2 e um algoritmo diferente no par N7K3/N7k4). Quando o algoritmo de balanceamento de carga é alterado, os switches N7k1/N7k2 agora hash o tráfego de entrada com base em outras informações além das usadas pelos switches N7k3/N7k4 e, portanto, o tráfego de saída usa todos os links de membro do canal de porta. (A decisão sobre qual algoritmo escolher depende do tipo de tráfego recebido pelo switch)

3. Se o cliente quiser usar o mesmo algoritmo de balanceamento de carga, use valores de rotação diferentes em cada nível de switches. O comando Rotate introduz aleatoriedade no algoritmo de hash, ao compensar a entrada de hash por bytes configurados pelo usuário e ajuda a evitar a polarização. (Use um valor de rotação para o par N7k1/N7k2 e um valor de rotação diferente para o par N7k3/N7k4)