

포트 채널 로드 밸런싱에서 양극화 문제 해결

목차

[소개](#)

[배경](#)

[사전 요구 사항](#)

[토폴로지](#)

[구성](#)

[트래픽 흐름](#)

[문제 해결](#)

[해결 방법](#)

소개

이 문서에서는 포트 채널 로드 밸런싱의 양극화가 발생할 수 있는 시나리오에 대해 설명하고 이를 방지하는 방법에 대한 제안을 제공합니다.

배경

양극화 해시 알고리즘이 네트워크의 특정 경로를 선택하고 이중 경로를 사용하지 않는 상태로 두는 문제

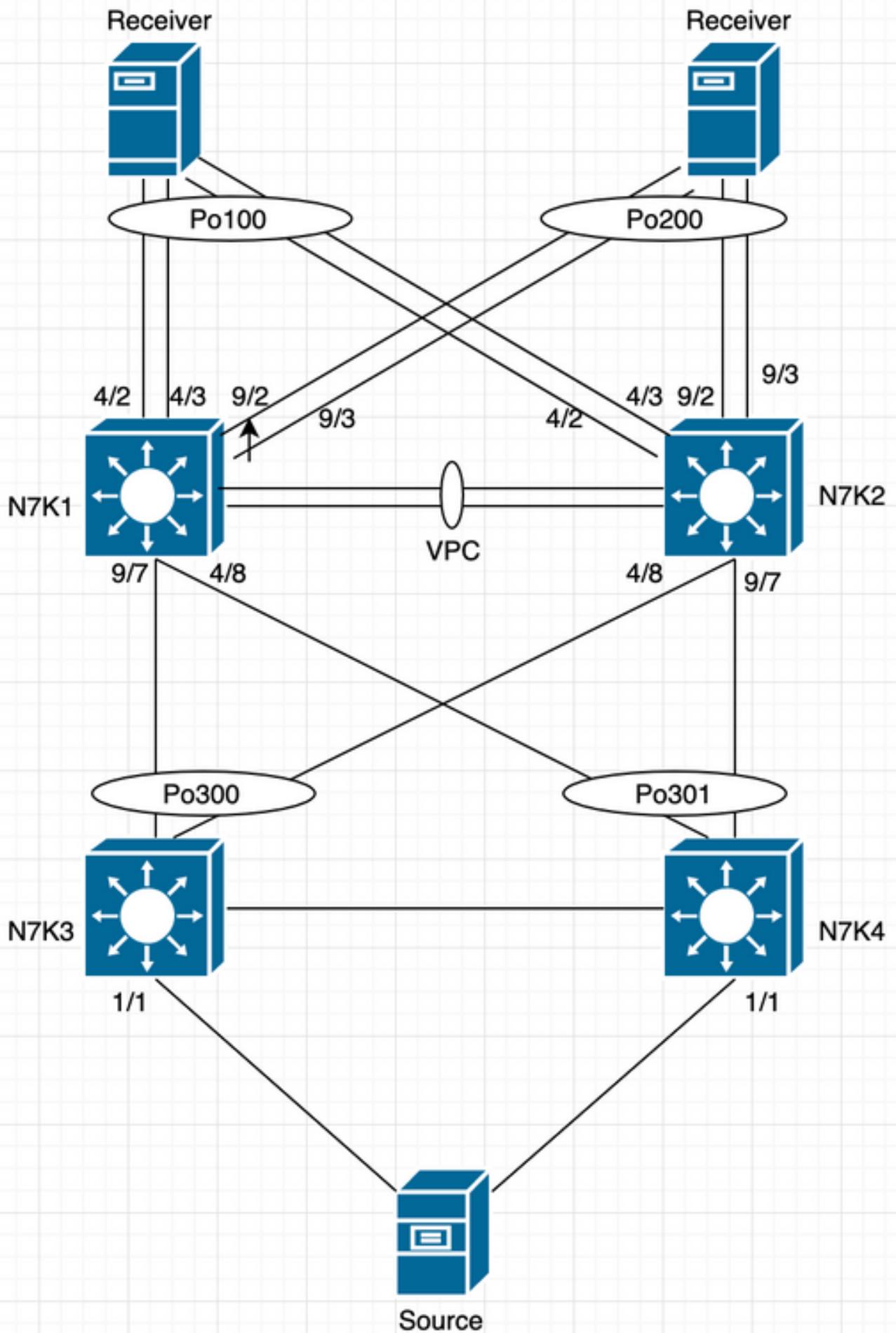
사전 요구 사항

다음 주제에 대한 지식을 습득하는 것이 좋습니다.

[링크 집계 제어 프로토콜](#)

Cisco Nexus 플랫폼

토폴로지



구성

VPC 및 Po100, Po200, Po300 및 Po301에 연결된 N7K1 및 N7K2는 VPC 포트 채널에 있습니다.

N7K1 및 N7K2는 이 스위치에서 라우팅 없이 순수 L2 스위치로 작동합니다.

모든 스위치가 동일한 포트 채널 로드 밸런싱 알고리즘을 실행 중입니다.

소스에서 대상으로의 트래픽이 동일한 VLAN에 있는지(라우팅 없음) 또는 N7K3 또는 N7K4에서 발생하는 라우팅과 함께 다른 VLAN에 있었는지에 관계없이 N7K1 및 N7K2에서 나가는 트래픽에서 양극화 문제가 발생합니다.

트래픽 흐름

소스는 여러 스트림을 대상으로 보냅니다(여러 소스 및 목적지 IP 주소가 있고 L4 포트 정보는 패킷마다 다름). 이상적인 상황에서 트래픽이 포트 채널 멤버 인터페이스 간에 균일하게 분산되도록 트래픽을 적절히 혼합하여 사용합니다.

N7k3/N7k4의 소스 랜드 트래픽을 보낸 다음 N7K1/N7K2를 통해 목적지로 이동합니다.

각 N7K1 및 N7K2에 있는 Po100 및 Po200의 멤버 링크 중 하나의 링크는 트래픽의 거의 99%를 전송하고 다른 링크는 유휴 상태로 남아 있습니다.(예: 각 스위치 N7K1 및 N7K2에서 4/2 및 4/3 중 하나의 링크는 99%의 유니캐스트 트래픽을 전송하고, 다른 링크는 1% 미만을 전달하며, 9/2 및 9/3 중 하나의 링크는 99%의 트래픽을 전달하며, 다른 링크는 1% 미만을 전달합니다. 문제 해결 섹션의 출력은 N7K1의 po100 및 po200 멤버 인터페이스에 대한 트래픽을 보여주며 N7K2에서는 유사한 출력을 볼 수 있습니다.)

사용된 포트 채널 로드 밸런싱 알고리즘의 유형에 관계없이 N7K1/N7K2 쌍 및 N7K3/N7K4 쌍에서 동일한 포트 채널 로드 밸런싱 알고리즘을 사용하는 한 이 문제가 발생할 수 있습니다. 포트 채널 로드 밸런싱 알고리즘을 확인하는 명령은 아래에 나와 있습니다.

```
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

문제 해결

포트 채널에서 고르지 않은 로드 밸런싱이 확인되면 양극화 때문일 수 있습니다.

트래픽이 N7K3 및 N7K4 스위치에 도달하면 N7K4의 Po301 및 N7K3의 Po300을 통해 N7K1/N7K2 스위치로 전달됩니다. 여기서 로드 밸런싱 알고리즘이 작동하며 일부 플로는 N7K1로 전달되고 다른 플로는 N7K2로 전달됩니다.

처음에는 모든 트래픽이 eth1/1의 스위치 N7K3/N7K4로 들어오고 src-dest ip 및 l4 포트 정보를 기반으로 특정 플로우가 N7K1로 향하는 링크와 N7K2로 향하는 링크에 해시된 다른 플로우에 해시됩니다. 해시는 스위치에 의해 계산된 rbh 값을 기반으로 수행됩니다. 간소화를 위해 사용된 로드 밸런싱 알고리즘을 기반으로, 스위치는 수신 트래픽을 두 개의 플로우(플로우 X 및 플로우 Y)로 분리한다고 가정해 보겠습니다. 하나의 포트 채널 멤버 링크에서 Flow X가 전송되고 다른 포트 채널 멤버 링크에서 Flow Y가 전송되었습니다.

이제 트래픽이 N7K1/N7K2 쌍에 도달하면 두 가지 가능성이 있습니다.(X와 Y의 상호 교환 가능 고려)

사례1:

N7K3은 N7K1에 플로우 X를 전송하고 Y는 N7K2에 플로우 Y를 전송했습니다.

및

N7K4는 N7K1에 플로우 Y를 보내고 N7K2에 플로우 X를 보냈습니다.

사례2:

N7K3은 N7K1에 플로우 X를 전송하고 Y는 N7K2에 플로우 Y를 전송했습니다.

및

N7K4는 N7K1에 플로우 X를 보내고 Y는 N7K2에 흐름 Y를 보냈습니다.

케이스 1, N7K1 및 N7K2는 플로우 유형(플로우 X 및 플로우 Y)을 모두 수신하며, N7K3/N7K4에서 사용하는 것과 동일한 포트 채널 로드 밸런싱 알고리즘을 사용한 후에도 다른 링크에서 Po100 및 Po200에서 나가는 플로우로 인해 균열이 나타나지 않으므로 포트 채널 멤버 인터페이스 간에 트래픽 분배가 더 나은 것으로 보입니다.

사례 2의 경우 N7K1은 플로우 X만 수신하고 N7K2는 플로우 Y만 수신하며, 이 스위치에서 사용되는 포트 채널 로드 밸런싱 알고리즘이 N7K3/N7K4 쌍에서 사용된 것과 동일한 경우 양극화가 발생할 수 있습니다. N7K1 및 N7K2가 동일한 포트 채널 로드 밸런싱 알고리즘을 사용하므로 N7K1은 Po100/Po200의 멤버 링크 한 개에서만 흐름 X를 전송하고 다른 멤버 링크는 트래픽을 전달하지 않습니다. 마찬가지로, N7K2는 Po100/Po200의 멤버 링크 하나에서만 흐름 Y를 전송하고 다른 멤버 링크는 트래픽을 전달하지 않습니다.

N7K1 및 N7K2를 전환하는 트래픽이 이미 다음으로 분류되었으므로 하나의 포트 채널 멤버 링크만 스위치 N7K1/N7K2에서 모든 수신 트래픽을 보내는 데 사용되며 다른 멤버 링크에서는 아무것도 전송되지 않습니다. 수신 트래픽 속도가 단일 포트 채널 링크의 대역폭을 초과하는 경우 다른 포트

채널 멤버 링크가 이 트래픽을 전달하지 않으므로 추가 트래픽을 삭제할 수 있습니다.

포트 채널에서 2개 이상의 링크가 사용되는 경우에도 비슷한 문제가 발생할 수 있습니다. 예를 들어 포트 채널에 4개의 링크가 사용되는 경우 해싱에 따라 양극화가 발생하지 않거나 4개의 포트 채널 멤버 링크 중 2개만 사용하여 모든 수신 트래픽을 전달하고 나머지 2개의 링크는 아무것도 전달하지 않는 부분적인 양극화를 볼 수 있습니다

양극화는 디자인 때문에 발생하므로, 양극화가 발생하지 않도록 디자인을 분석하는 것이 중요합니다. N7k1에서 Po100 및 Po200이 발생하는 양극화를 나타내는 출력이 아래에 표시됩니다(N7K2에서도 유사한 출력이 표시될 수 있음).

```
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](#) 외부 문서에 대해 제출되었습니다.

해결 방법

양극화가 발생하지 않도록 하기 위해 사용할 수 있는 몇 가지 해결 방법을 따릅니다.

1. 설계양극화의 주된 원인은 부적절한 설계이므로 토폴로지에 양극화가 발생하지 않도록 네트워크 설계를 변경하는 것이 좋습니다.

설계를 변경할 수 없으면 다음을 수행할 수 있습니다.

2. 각 스위치 레벨에서 서로 다른 포트 채널 로드 밸런싱 알고리즘을 사용합니다(N7K1/N7k2 쌍의 알고리즘 하나와 N7K3//N7k4 쌍의 다른 알고리즘). 로드 밸런싱 알고리즘이 변경되면 N7k1/N7k2 스위치는 이제 N7k3/N7k4 스위치에서 사용하는 정보 이외의 다른 정보를 기반으로 수신 트래픽을 해시하므로 발신 트래픽은 모든 포트 채널 멤버 링크를 사용합니다.(선택할 알고리즘에 대한 결정은 스위치에서 수신한 트래픽 유형에 따라 달라집니다.)

3. 고객이 동일한 로드 밸런싱 알고리즘을 사용하려는 경우 스위치 레벨마다 서로 다른 회전 값을 사용합니다.회전 명령은 사용자가 구성한 바이트로 해시 입력을 오프셋하여 해싱 알고리즘의 임의 설정을 도입하고 분화를 방지하는 데 도움이 됩니다.(N7k1/N7k2 쌍에 대해 하나의 회전 값을 사용하고 N7k3/N7k4 쌍에 대해 다른 회전 값을 사용합니다.)