Nexus 7000风暴控制：选择适当的抑制值

目录

简介

当数据包泛洪LAN时，会发生流量风暴，造成过多流量并降低网络性能。您可以使用流量风暴控制功能防止物理接口上的广播、组播或单播流量风暴对第2层端口造成中断。

流量风暴控制（也称为流量抑制）允许您在10毫秒的间隔内监控传入广播、组播和单播流量的级别。在此间隔内，流量级别（即端口总可用带宽的百分比）与您配置的流量风暴控制级别进行比较。当入口流量达到端口上配置的流量风暴控制级别时，流量风暴控制会丢弃流量，直到间隔结束。

流量风暴控制阈值数和时间间隔允许流量风暴控制算法以不同的粒度级别工作。阈值越高，允许更多数据包通过。

默认情况下，当流量超过配置级别时，Cisco Nexus操作系统(NX-OS)软件不会采取纠正措施。但是，您可以配置嵌入式事件管理(EEM)操作，以在特定时间段内如果流量不在子端（低于阈值）时错误禁用接口

流量风暴控制的准则和限制

配置流量风暴控制级别时，请注意以下准则和限制：

- 您可以在端口通道接口上配置流量风暴控制。
- 请勿在属于端口通道接口的接口上配置流量风暴控制。在配置为端口通道成员的接口上配置流量风暴控制会使端口进入挂起状态。
- 将级别指定为总接口带宽的百分比： 级别可以是0到100。级别的可选分数可以是0到99。100%表示没有流量风暴控制。0%抑制所有流量。

由于硬件限制以及计算不同大小数据包的方法，级别百分比是近似值。根据构成传入流量的帧的大小，实际实施级别可能与配置级别相差几个百分点。

流量风暴控制的默认设置

参数	默认
流量风暴控制	禁用
阈值百分比	100

配置流量风暴控制

您可以设置受控流量可以使用的总可用带宽的百分比。

1. configure terminal
2. 接口 {以太网 插槽/端口 | 端口通道 号码}
3. 风暴控制 {广播 | 组播 | 单播} 级别 百分比[。分数]

注意：流量风暴控制使用10毫秒的间隔，可影响流量风暴控制的行为。

检验流量风暴控制配置

要显示流量风暴控制配置信息，请执行以下任务之一：

命令	目的
show interface [以太网 插槽/端口 | 端口通道 号码] 计数器风暴控制	显示接口的流量风暴控制配置。
show running-config interface	显示流量风暴控制配置。

监控流量风暴控制计数器

您可以监控Cisco NX-OS设备维护的流量风暴控制活动的计数器。

```
switch# show interface counters storm-control
```

Nexus 7000风暴控制：选择适当的抑制值

为帮助客户选择适当的阈值，本部分提供有关使用阈值背后的逻辑的深入见解。

注意：此处提供的信息不提供任何最佳实践编号，但客户在浏览信息后可以做出合理的决定。

使用的组件

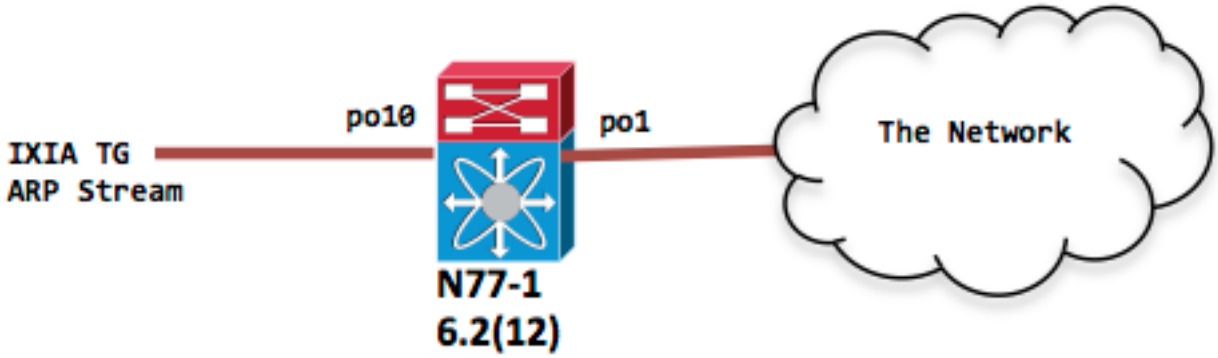
本文档中的信息基于以下软件和硬件版本：

- Nexus 7700，带6.2.12及更高版本。
- F3系列线卡。

本文档中的信息都是基于特定实验室环境中的设备编写的。本文档中使用的所有设备最初均采用原始（默认）配置。如果您使用的是真实网络，请确保您已经了解所有命令的潜在影响。

实验室测试

风暴控制是一种流量抑制机制，应用于特定端口上的入口流量。




```
N77-1(config-if)# sh port-c sum
1     Po1(SU)     Eth      LACP      Eth2/4(P)
10    Po10(SU)    Eth      LACP      Eth1/1(P)

interface port-channel1
switchport

interface port-channel10
switchport
```

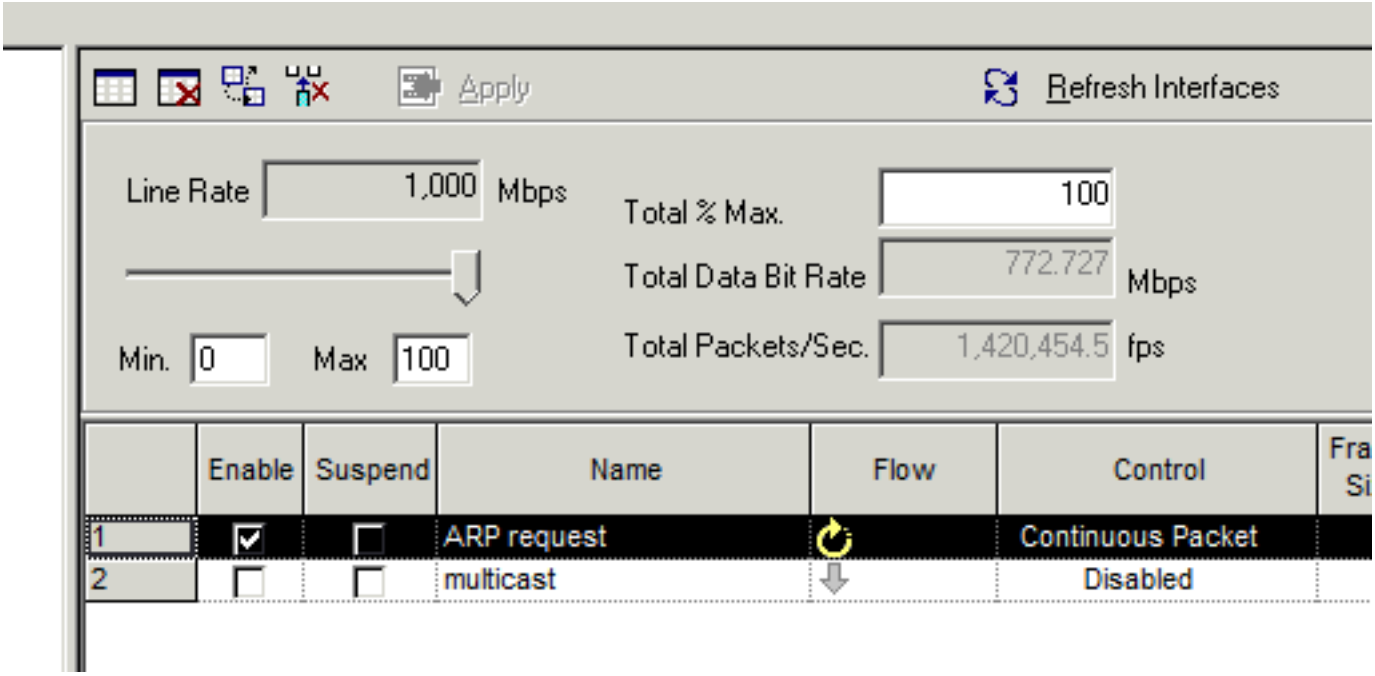
场景1:抑制率为0.01%

入口流量速率设置为1Gbps的ARP请求流量

config

interface port-channel10
风暴控制广播级别0.01

IXIA快照供参考




```
N77-1(config-if)# sh int po10 | in rate | in "30 sec"
   30 seconds input rate 954649416 bits/sec, 1420607 packets/sec
   30 seconds output rate 1856 bits/sec, 0 packets/sec
input rate 954.82 Mbps, 1.42 Mpps; output rate 1.97 Kbps, 0 pps

N77-1(config-if)# sh int po1 | in rate | in "30 sec"
   30 seconds input rate 8656 bits/sec, 8 packets/sec
   30 seconds output rate 853632 bits/sec, 1225 packets/sec >>>> Output rate is ~ 1200 pps
input rate 8.74 Kbps, 8 pps; output rate 875.32 Kbps, 1.22 Kpps

N77-1# sh int po10 counters storm-control
--------------------------------------------------------------------------------
Port        UcastSupp %     McastSupp %     BcastSupp %     TotalSuppDiscards
--------------------------------------------------------------------------------
Po10             100.00          100.00            0.01           67993069388
```

显示风暴控制丢弃以供参考。

场景2:抑制率为0.1%

入口流量速率设置为1Gbps的ARP请求流量

config

interface port-channel10
风暴控制广播级别0.10

仅显示出口接口，因为入口接口po10的传入流量速率相同，为1gbps

```
N77-1(config-if)# sh int po1 | in rate | in "30 sec"
  30 seconds input rate 8840 bits/sec, 8 packets/sec
  30 seconds output rate 8253392 bits/sec, 12271 packets/sec >>>> Output rate is ~ 12k pps
```

场景3:抑制率为1%

入口流量速率设置为1Gbps的ARP请求流量

config

interface port-channel10

风暴控制广播级别1

仅显示出口接口，因为入口接口po10的传入流量速率相同，为1gbps

```
N77-1(config-if)# sh int po1 | in rate
  30 seconds input rate 8784 bits/sec, 7 packets/sec
  30 seconds output rate 86601056 bits/sec, 129293 packets/sec >>>> Output rate is ~ 120k pps
  input rate 8.78 Kbps, 7 pps; output rate 86.60 Mbps, 129.29 Kpps
```

场景4:抑制率为10%

入口流量速率设置为1Gbps的ARP请求流量

config

interface port-channel10

风暴控制广播级别10.00

```
 N77-1(config-if)# sh int po1 | in rate
  30 seconds input rate 8496 bits/sec, 7 packets/sec
  30 seconds output rate 839570968 bits/sec, 1249761 packets/sec >>>> Output rate is ~ 1.2mil
pps
  input rate 8.50 Kbps, 7 pps; output rate 839.57 Mbps, 1.25 Mpps
```

摘要:

上述所有场景都处理可能由环路或网卡故障导致的持续流量。风暴控制在流量注入网络之前限制流量速率的场景中是有效的。不同的抑制级别可指示您将注入网络的流量。

当风暴控制到位时，如果将阈值保持在主动级别，是否会导致正常ARP丢弃？

需要考虑以下几点

1. 首先，如果ARP是首次丢弃的，则始终由应用层发起重试，因此在后续重试期间ARP被解析的机率更高，并将导致IP到MAC的成功解析。
2. 风暴控制是入口监察器，应尽可能靠近边缘应用。因此，您可能需要处理一台物理主机或VM群集。如果一台主机，则ARP的数量在正常工作情况下并不是问题。如果这是VM集群，则您可能有一定数量的主机，但同样不会显示边缘端口后面的整个第2层域。
3. 如果在核心端口上应用风暴控制配置，则在广播流量到达核心层之前，应了解广播流量如何聚合。

回到我们的测试 — 此处是针对突发ARP流量的一些测试 —

测试 1：5000个数据包突发量@ 5000pps单突发量

抑制级别0.01%

config

interface port-channel10

风暴控制广播级别0.01

```
N77-1# sh int po10
port-channel10 is up
admin state is up
RX
  12985158 unicast packets 27 multicast packets 5000 broadcast packets
   12990674 input packets 1091154042 bytes
  0 jumbo packets 2560 storm suppression packets

N77-1#Sh int po1
port-channel1 is up
admin state is up
TX
  0 unicast packets 507 multicast packets 2440 broadcast packets

N77-1(config-if)# sh int po10 counters storm-control
--------------------------------------------------------------------------------
Port       UcastSupp %     McastSupp %     BcastSupp %     TotalSuppDiscards
--------------------------------------------------------------------------------
Po10             100.00         100.00            0.01                  2560
```

上面显示2560个丢弃的ARP数据包。当然，如果一个接口后面有5000台主机，那么其中一半在第一次迭代期间通过，而后一半在下一次迭代期间通过。如果您的应用程序只发送一个ARP请求来获取IP到MAC的解析，则如果没有响应，则可能需要修改应用程序以重新传输ARP请求。 在这种情况下，请咨询应用供应商以获取更改此行为的帮助。

测试 2：5000个数据包突发量@ 50000pps单突发量

抑制级别0.01%

config

interface port-channel10

风暴控制广播级别0.01

```
N77-1(config-if)# sh int po10
port-channel10 is up
admin state is up
RX
  0 unicast packets 19 multicast packets 5000 broadcast packets
```

```
  5019 input packets 435550 bytes
  0 jumbo packets 3771 storm suppression packets

N77-1(config-if)# sh int po1
port-channel1 is up
admin state is up
TX
  0 unicast packets 712 multicast packets 1229 broadcast packets

N77-1(config-if)# sh int po10 counters storm-control
--------------------------------------------------------------------------------
Port       UcastSupp %     McastSupp %     BcastSupp %     TotalSuppDiscards
--------------------------------------------------------------------------------
Po10             100.00         100.00           0.01                 3771
```

在上述输出中，由于数据包突发速率较高，丢包数更多。

类似结果也显示，5000个数据包突发速率在100kpps时提高到1 gbps数据包速率

以下选项可用于检测风暴状况。

在数据层面发出警报：

- 配置风暴控制会生成警报的系统日志消息，您可以将EEM绑定以生成简单网络管理协议(SNMP)陷阱或关闭端口作为预防措施。

在控制平面发出警报：

- 配置“logging drop threshold”选项：

在Nexus 7k上，有一个默认策略映射 — 控制平面：

此策略映射正在调节哪些流量正在传递到CPU。在此策略映射中，您可以看到一个类，该类用于调节ARP流向CPU的量。

在此类下配置“logging drop threshold”将报告系统日志中的任何违规，您可以进一步使用EEM生成SNMP陷阱。

- 控制平面策略(CoPP)MIB轮询

从NX-OS 6.2(2)开始，CoPP支持思科基于类的QoS MIB(cbQoSMIB)，并且其所有元素都可以使用SNMP进行监控

结论

风暴控制是防止物理接口上的广播、组播或单播流量风暴对第2层端口造成中断的有用功能。此功能在影响控制平面和CoPP之前控制数据平面上的风暴。