Catalyst 9000 交换机输出丢包故障排除

目录

简介

先决条件

- 要求
- 使用的组件

背景信息

什么是输出丢弃

- 拥塞类型
- 低吞吐量拥塞
- 验证缓冲区拥塞

修改缓冲区以解决输出丢弃

- SoftMax乘数
- 每队列缓冲区修改
- 管理拥塞的备用方法

使用Wireshark分析输出丢包

- 查看I/O速率
- 查看I/O速率（以毫秒为单位）

简介

本文档介绍如何对 Catalyst 9000 系列平台上的输出丢包进行故障排除。

先决条件

要求

要对Catalyst 9000系列平台的服务质量(QoS)进行故障排除，您必须了解：

- 标准QoS概念
- 模块化QoS命令行界面(CLI)

使用的组件

本文档中的信息基于此硬件和软件版本，但该方法和大多数命令可应用到其他代码上的其他Catalyst 9000系列交换机：

- Cisco Catalyst 9300
- 思科IOS® XE 16.12.3

注意：有关在其他思科平台上启用这些功能所使用的命令，请参阅相应的配置指南。

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。如果您的网络处于活动状态，请确保您了解所有命令的潜在影响。

背景信息

有关Catalyst 9000系列平台上的QoS的详细说明（包括默认QoS配置、队列结构和缓冲区说明），请参阅Catalyst 9000 QoS和排队白皮书查看推荐的版本指南，以确保您使用的是针对您的平台的最新推荐软件。这些建议可确保您的软件得到支持，并帮助避免旧代码中的已知错误。推荐的适用于Catalyst的版本

什么是输出丢弃

有关缓冲区分配的知识可帮助您了解缓冲区拥塞如何导致输出丢弃。当目标接口的数据包数量超过其输出速率时，就会发生拥塞。这些数据包必须存储在缓冲区中，直到可以传输。假设这些交换机每个ASIC最多有36 MB缓冲区，然后在ASIC上的所有端口之间共享。虽然出口接口能够以线速清空该缓冲区，但任何导致以更高速率缓冲数据包的场景都可能导致拥塞。即使流量突发仅持续几分之一秒，也可能会发生拥塞，并且可能导致流量出现延迟；如果缓冲区被完全填满，则会导致输出丢弃。

注:show interface中显示的输出丢弃计数器默认以字节显示。在版本16.9.3及更高版本中，这些计数器默认为数据包。

拥塞类型

如图1所示，有两种拥塞类型。

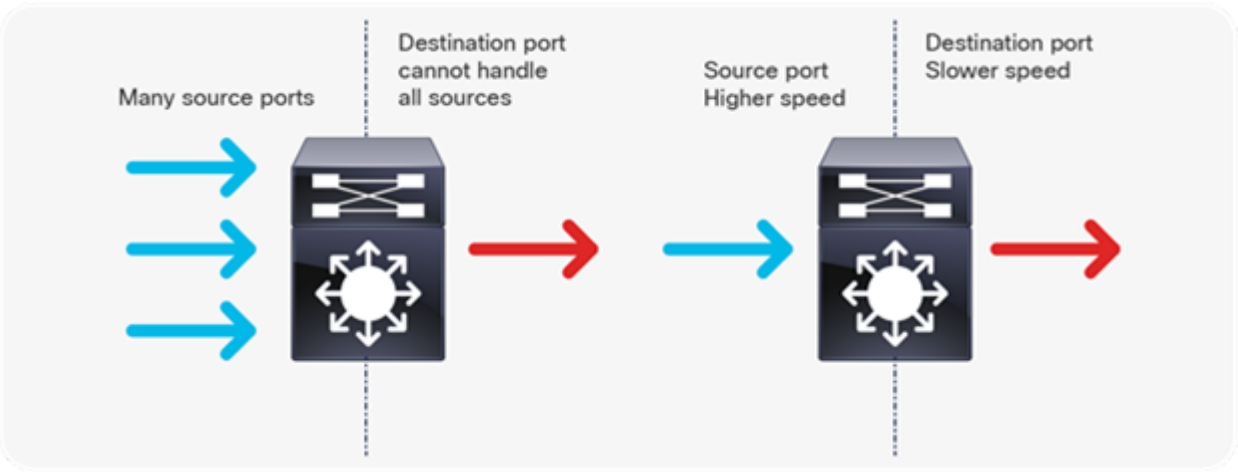


图1.拥塞类型

图1中显示的两种拥塞类型是：

- 多对一：当多个源端口同时向单个目标发送流量时，目标端口可能会因从多个源接收的流量量而拥塞。
- 速度不匹配：当速度较高的端口传输到速度较低的端口（例如10 Gbps到1 Gbps）时，数据包必须花费时间从出口端口排出，这可能导致延迟和/或丢包。

低吞吐量拥塞

流量突发可能会导致输出丢弃，即使接口输出速率明显低于最大接口容量。默认情况下，show interface命令的输出速率在五分钟内取平均值，这不足以捕获任何短期突发。最好将其平均30秒以上，即使在这种情况下，爆发几毫秒的流量也可能导致输出丢弃，而不会导致30秒平均速率增加。本文档可用于对Catalyst 9000系列交换机上看到的任何其他类型的拥塞进行故障排除。

验证缓冲区拥塞

有两个命令可用于验证缓冲区拥塞。第一个命令是show platform hardware fed switch active qos queue config interface <interface>。此命令允许您查看端口上的当前缓冲区分配，如图2所示。

<#root>

9300#

show platform hardware fed switch active qos queue config interface gigabitEthernet 1/0/48

```
Asic:0 Core:0 DATA Port:47 GPN:48 LinkSpeed:0x1
AFD:Disabled FlatAFD:Disabled QoSMap:0 HW Queues: 376 - 383
  DrainFast:Disabled PortSoftStart:2 - 1800
   DTS
```

Hardmax

Softmax

	PortSMin		GlblSMin		PortStEnd				
0	1	6	200	7	800	19	475	0	0	3	2400
1	1	5	0	8	1200	19	712	8	300	3	2400
2	1	5	0	6	0	0	0	0	0	3	2400
3	1	5	0	6	0	0	0	0	0	3	2400
4	1	5	0	6	0	0	0	0	0	3	2400
5	1	5	0	6	0	0	0	0	0	3	2400
6	1	5	0	6	0	0	0	0	0	3	2400
7	1	5	0	6	0	0	0	0	0	3	2400

图 2. 队列缓冲区分配

您特别想查看Hardmax和Softmax列，其中显示队列可用的缓冲区数量。有关这些缓冲区的内容以及默认分配方式的信息，请参阅Catalyst 9000 QoS和排队白皮书。

第二个命令是show platform hardware fed switch active qos queue stats interface <interface>。此命令允许您查看接口上的每队列统计信息，包括有多少字节已排入缓冲区，以及有多少字节由于缺少可用缓冲区而被丢弃。

<#root>

9300#

show platform hardware fed switch active qos queue stats interface Gig 1/0/1

```
DATA Port:0 Enqueue Counters
-------------------------------------------------------------------------------------------
Q Buffers          Enqueue-TH0          Enqueue-TH1          Enqueue-TH2             Qpolicer
  (Count)              (Bytes)              (Bytes)              (Bytes)               (Bytes)
- ------- -------------------- -------------------- -------------------- --------------------
0       0                    0                    0
```

384251797

```
                    0
1       0                   0                   0

488393930284

                    0
2       0                   0                   0                   0                   0
3       0                   0                   0                   0                   0
4       0                   0                   0                   0                   0
5       0                   0                   0                   0                   0
6       0                   0                   0                   0                   0
7       0                   0                   0                   0                   0

DATA Port:0 Drop Counters
-------------------------------------------------------------------------------------------------------
Q             Drop-TH0             Drop-TH1             Drop-TH2             SBufDrop             QebD
               (Bytes)              (Bytes)              (Bytes)              (Bytes)              (Byt
- -------------------- -------------------- -------------------- -------------------- ----------------
0                     0                    0                    0                    0
1                     0                    0

192308101

                    0                    0                    0
2                     0                    0                    0                    0
3                     0                    0                    0                    0
4                     0                    0                    0                    0
5                     0                    0                    0                    0
6                     0                    0                    0                    0
7                     0                    0                    0                    0
```

图3.包含丢弃的队列缓冲区统计信息

如图3所示，Queue 0和Queue 1的字节都已排入队列，但正是队列1在Drop-TH2列中遇到丢包。此信息表明队列0流量未受此拥塞影响，并且拥塞的原因特别是队列1流量。

修改缓冲区以解决输出丢弃

SoftMax乘数

要增加每个队列可以从共享池请求的缓冲区数量，请使用配置qos queue-softmax-multiplier <100 - 1200>增加SoftMax阈值。最大值为1200，并且以12的倍数增加单个端口队列吸收微突发的能力。此命令会增加端口队列阈值，以便端口队列可以占用共享池中的更多缓冲单元。如图4所示，配置和增加的缓冲区分配。

```
9300(config)#

qos queue-softmax-multiplier 1200



9300#

show platform hardware fed switch active qos queue config interface


gigabitEthernet


1/0/48


Asic:0 Core:0 DATA Port:47 GPN:48 LinkSpeed:0x1
AFD:Disabled FlatAFD:Disabled QoSMap:0 HW Queues: 376 - 383
  DrainFast:Disabled PortSoftStart:3 - 14400
   DTS  Hardmax  Softmax   PortSMin  GlblSMin  PortStEnd
  ----- --------  --------  --------  --------  ---------
 0   1  6   200   9  9600   2   600   0     0   1 15000
 1   1  5     0  10 14400   2   900   1   450   1 15000
 2   1  5     0   6     0   0     0   0     0   1 15000
 3   1  5     0   6     0   0     0   0     0   1 15000
 4   1  5     0   6     0   0     0   0     0   1 15000
 5   1  5     0   6     0   0     0   0     0   1 15000
 6   1  5     0   6     0   0     0   0     0   1 15000
 7   1  5     0   6     0   0     0   0     0   1 15000
```

图4.使用SoftMax乘数1200的队列配置

这是一种常用配置，可快速解决输出丢包问题。在映像4中，此配置适用于所有接口上的所有非优先级队列。缓冲区分配本身假设不会在所有端口上同时发生微爆发。如果微突发在随机时刻发生，则共享缓冲器可以专门提供额外的缓冲单元来吸收它们。

每队列缓冲区修改

在无法使用SoftMax乘数的情况下，或者尝试微调缓冲区以适应流量量变曲线的情况下，可以使用每队列缓冲区修改。要修改队列缓冲区分配（即基于每个接口的交换机），必须使用策略映射。 在大多数情况下，可以修改接口的当前策略映射并根据类更改缓冲区。

在本示例中，接口GigabitEthernet1/0/48出现输出丢弃。如图5所示，应用于此接口的出口策略映射。

```
policy-map MYPOL
 class Voice
  priority level 1 percent 20
 class Video
  priority level 2 percent 10
 class Control
  bandwidth percent 10
 class Data
  bandwidth percent 5
 class class-default
```

图5.策略映射示例

此策略映射具有5个类映射，导致接口上共有5个出口队列。每个类都有根据它们的优先级分配给它的缓冲区的默认数量。

图6显示当前缓冲区分配。

<#root>

9300#

show platform hardware fed switch active qos queue config interface gigabitEthernet 1/0/48

```
Asic:0 Core:0 DATA Port:47 GPN:48 LinkSpeed:0x1
AFD:Disabled FlatAFD:Disabled QoSMap:0 HW Queues: 376 - 383
  DrainFast:Disabled PortSoftStart:3 - 600
   DTS  Hardmax  Softmax   PortSMin  GlblSMin  PortStEnd
  ----- --------  --------  --------  --------  ---------
 0   1  7   100    9   100   0     0   0     0   3   800
 1   1  7   100   10   400  19   237   0     0   3   800
 2   1  5     0   10   400  19   237   8   100   3   800
 3   1  5     0   10   400  19   237   8   100   3   800
 4   1  5     0   10   400  19   237   8   100   3   800
 5   1  5     0    6     0   0     0   0     0   3   800
 6   1  5     0    6     0   0     0   0     0   3   800
 7   1  5     0    6     0   0     0   0     0   3   800
```

图6.使用示例策略的队列缓冲区配置

由于此接口遇到输出丢弃，请查看接口的队列统计信息，了解拥塞所在。

<#root>

9300#

show platform hardware fed switch active qos queue stats interface gigabitEthernet 1/0/48

```
DATA Port:0 Enqueue Counters
-------------------------------------------------------------------------------------------
Q Buffers          Enqueue-TH0          Enqueue-TH1          Enqueue-TH2             Qpolicer
  (Count)              (Bytes)              (Bytes)              (Bytes)              (Bytes)
- ------- -------------------- -------------------- -------------------- --------------------
0       0                    0                    0               489094                    0
1       0                    0                    0              4846845                    0
2       0                    0                    0             89498498                    0
3       0                    0                    0

21297827045


                   0
4       0                    0                    0             74983184                    0
5       0                    0                    0                    0                    0
6       0                    0                    0                    0                    0
7       0                    0                    0                    0                    0

DATA Port:0 Drop Counters
-------------------------------------------------------------------------------------------
Q             Drop-TH0             Drop-TH1             Drop-TH2             SBufDrop             QebD
                (Bytes)              (Bytes)              (Bytes)              (Bytes)             (Byt
- -------------------- -------------------- -------------------- -------------------- ----------------
0                    0                    0                    0                    0
1                    0                    0                    0                    0
2                    0                    0                    0                    0
3                    0                    0

3854484


                   0                    0                    0
4                    0                    0                    0                    0
5                    0                    0                    0                    0
6                    0                    0                    0                    0
7                    0                    0                    0                    0
```

图7.带丢弃的队列缓冲区统计信息及示例策略

图7显示，队列3加入队列的流量比任何其他队列都多，也是唯一一个遇到输出丢弃的队列。由于队列编号从0开始，队列3将映射到第四个类映射Data类。

要缓解此队列上的丢弃，请向队列3分配更多缓冲区。要更改此缓冲区分配，请在策略映射中使用queue-buffers ratio <0-100>配置。如果在策略中的每个类上配置，则最大值必须为100。如果仅使用此命令配置单个类，则系统尝试从其他队列中平均减去缓冲区。

在映像8中，Data类已配置了队列缓冲区比率40。

<#root>

```
policy-map MYPOL
```

```
class Voice
  priority level 1 percent 20
 class Video
  priority level 2 percent 10
 class Control
  bandwidth percent 10
 class Data
  bandwidth percent 5


queue-buffers ratio 40
```

图8.带有已修改的队列缓冲区的策略映射示例

在映像9中，您可以看到Data类现在具有40%的接口缓冲区，总共具有800个缓冲区。

<#root>

```
9300#

show platform hardware fed switch active qos queue config interface gigabitEthernet 1/0/48


Asic:0 Core:0 DATA Port:47 GPN:48 LinkSpeed:0x1
AFD:Disabled FlatAFD:Disabled QoSMap:0 HW Queues: 376 - 383
  DrainFast:Disabled PortSoftStart:3 - 1200
   DTS  Hardmax  Softmax   PortSMin  GlblSMin  PortStEnd
  ----- --------  --------  --------  --------  ---------
 0   1  7    75   9    75   0     0   0     0   3  1600
 1   1  7    75  10   300  19   178   0     0   3  1600
 2   1  5     0  10   300  19   178   8    75   3  1600
 3   1  5     0   7

800

  19   475   8   200   3  1600
 4   1  5     0  10   300  19   178   8    75   3  1600
 5   1  5     0   6     0   0     0   0     0   3  1600
 6   1  5     0   6     0   0     0   0     0   3  1600
 7   1  5     0   6     0   0     0   0     0   3  1600
```

图9.使用更新的示例策略的队列缓冲区配置

这也会导致其他队列的Softmax缓冲区较少。必须以较小的增量更改这些缓冲区，以确保更改不会导致其他队列上的输出丢弃。

进行更改后，检查队列统计信息，并查看此队列或任何其他队列上的丢弃是否仍会增加。如果丢弃继续，请进一步修改队列缓冲区配置，直到解决输出丢弃。

管理拥塞的备用方法

QoS主要是确定流量优先顺序的方法，它不是每个输出丢弃场景的解决方案。在某些情况下，对队列缓冲区的修改不足以解决所有输出丢弃。在这些情况下，您可以通过几种其他方法管理拥塞：

- 降低超订用率。

这包括增加出口带宽的方法，例如端口通道或等价多路径(ECMP)，但也可能需要更多涉及的配置，例如流量工程。

- 使用排队调度程序确定流量的优先级。

虽然队列调度程序不会停止拥塞，但它可以保护重要流量不受拥塞的影响

- 使用加权随机早期丢弃(WRED)或加权尾部丢弃(WTD)等拥塞管理算法提前丢弃某些流量。

- 管制入口上的流量以减少出口上的流量。

使用Wireshark分析输出丢包

Wireshark是识别导致缓冲区拥塞和丢弃的突发流量的有用工具。 如果在接口遇到丢包时在出口方向上跨越，Wireshark可以绘制输出速率图，以查看流量触发丢包的时间和内容。这在识别低吞吐量场景中的输出丢弃时尤其有用。

查看I/O速率

使用Wireshark打开SPAN捕获后，请依次选择Statistics和I/O Graph，如图10所示。

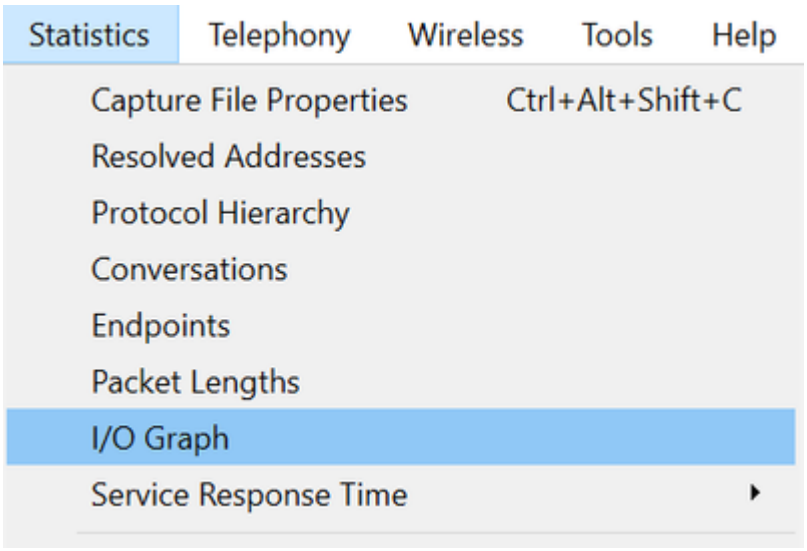


图10.选择I/O图

选中后，Wireshark会生成流量图（以比特/秒为单位）。图11显示接口遇到输出丢弃时的示例图形。

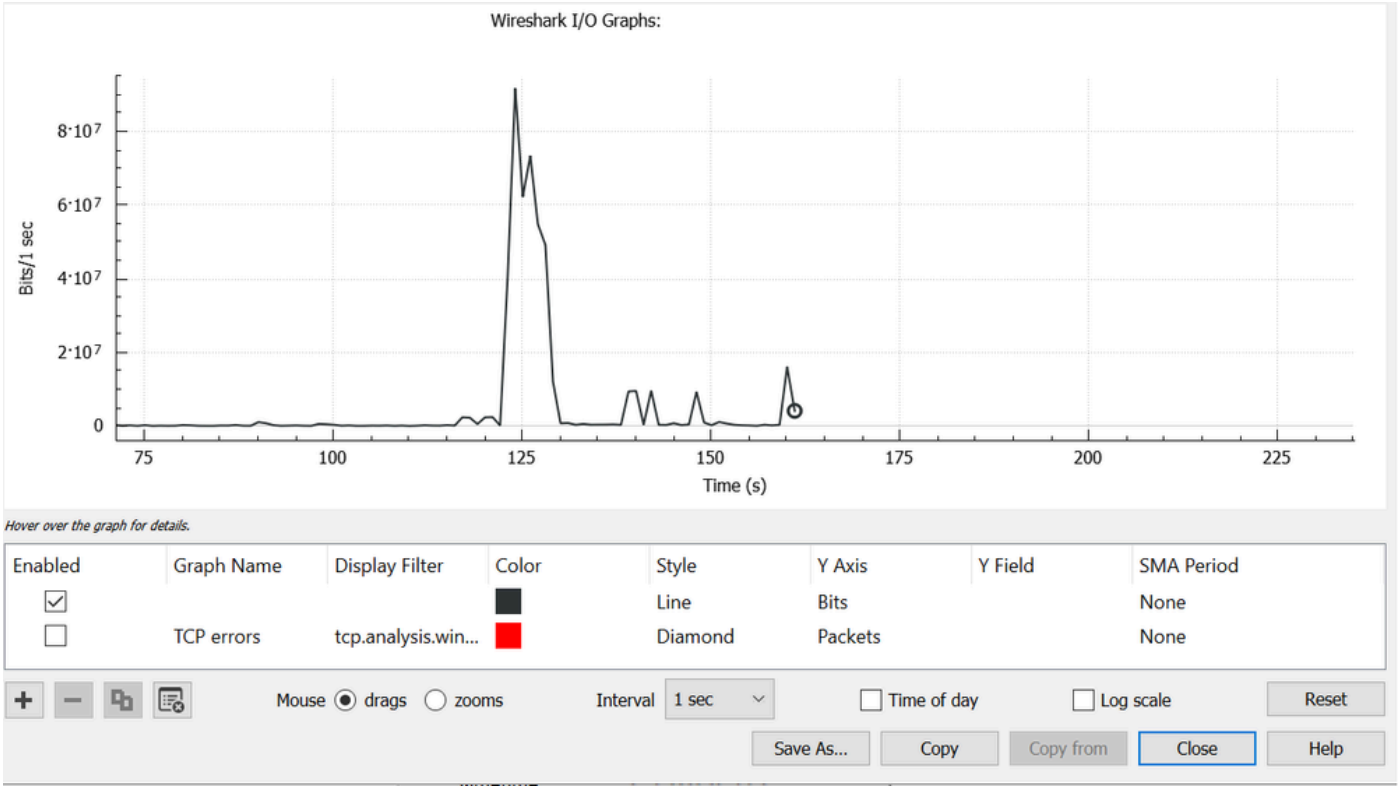


图11. I/O图形比特/毫秒

图11表示接口的最大吞吐量仅略高于80 Mbps。默认图形视图不够精确，无法识别导致丢包的小型流量突发。它是每秒的平均流量速率。要了解此速率如何导致缓冲区拥塞，请考虑以毫秒为单位的吞吐量。

千兆接口每秒可转发1,000,000,000位。转换到毫秒后，这等于每毫秒1,000,000（或10^6）位。

当接口速率超过接口的转发速度时，交换机必须缓冲这些数据包，从而导致拥塞和输出丢弃。

查看I/O速率（以毫秒为单位）

Wireshark允许用户将I/O速率图形化为每毫秒的位数。为此，请将Interval（时间间隔）从1秒减小到1毫秒，然后单击Reset以正确查看图形。此步骤如图12所示。

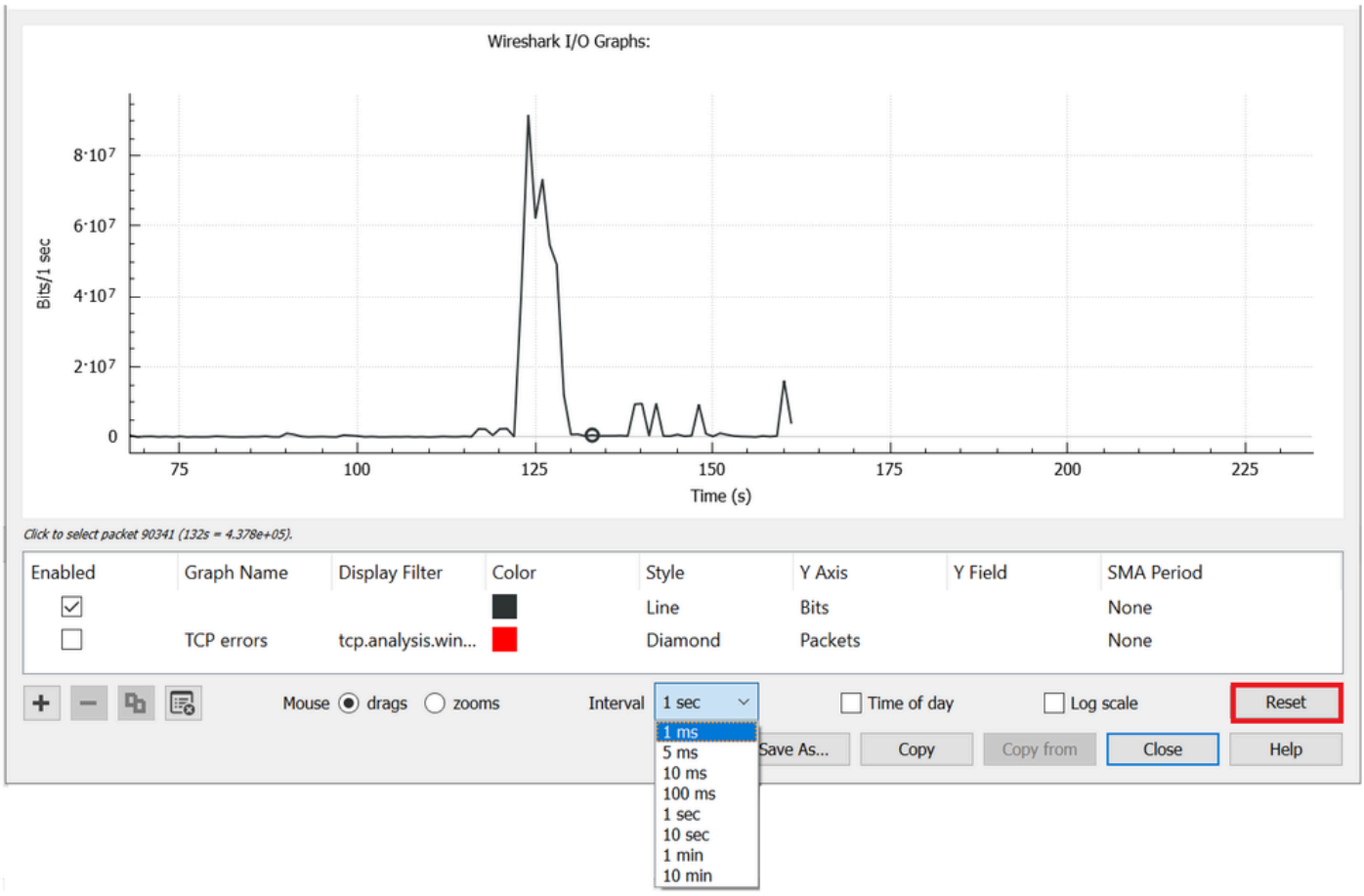


图12.将时间间隔缩短至1ms并重置图形

更新后的图表更准确地显示了接口的实际I/O速率。当速率达到或超过10^6位/毫秒时，交换机将遇到拥塞或输出丢弃。 图13显示出现输出丢弃的接口的更新I/O图。

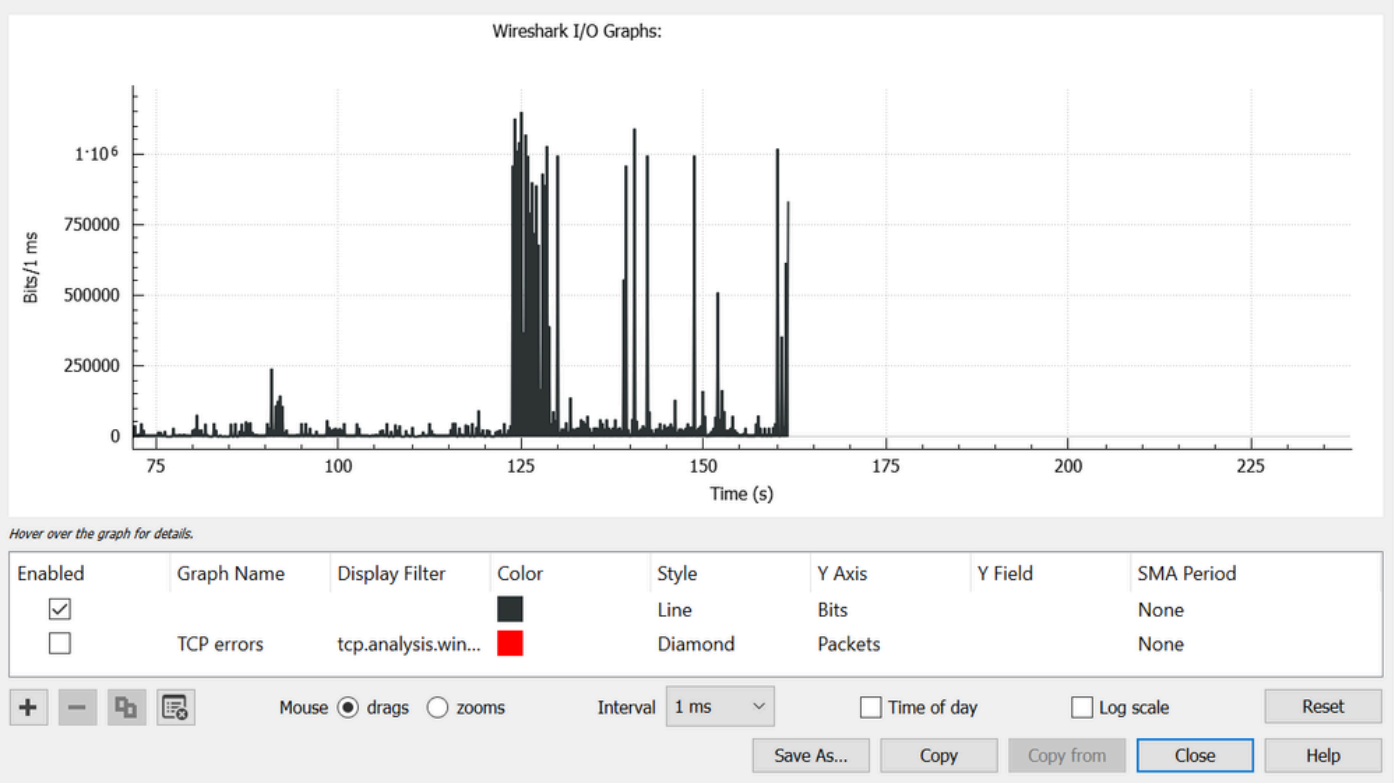


图13. I/O图形比特/毫秒

图13显示有多个流量峰值达到或超过10^6阈值。流量会受到缓冲，如果流量超过出口缓冲区大小，则会被丢弃。

注意：如果SPAN目标通过1 Gbps接口连接，则Wireshark中的I/O速率不能超过每毫秒10^6位，无论源接口速率如何。SPAN目标接口会缓冲或丢弃这些数据包。经常会看到I/O图形在最大吞吐量处停滞不前或呈现似乎会升高的平均业务速率。

关于此翻译
思科采用人工翻译与机器翻译相结合的方式将此文档翻译成不同语言，希望全球的用户都能通过各自的语言得到支持性的内容。

请注意：即使是最好的机器翻译，其准确度也不及专业翻译人员的水平。

Cisco Systems, Inc. 对于翻译的准确性不承担任何责任，并建议您总是参考英文原始文档（已提供链接）。