解释ACI中的丢包故障

目录

简介

背景信息

托管对象

硬件丢弃计数器类型

- 转发
- 错误
- 缓冲区

在CLI中查看丢弃统计信息

- 托管对象
- 硬件计数器
- 枝叶
- 主干

故障

- F112425 — 入口丢弃数据包速率(l2IngrPktsAg15min:dropRate)
- F100264 — 入口缓冲区丢弃数据包速率(eqptIngrDropPkts5min:bufferRate)
- F100696 — 入口转发丢弃数据包(eqptIngrDropPkts5min:forwardingRate)

统计信息阈值

- eqptIngrDropPkts中的转发丢弃数据包速率
- l2IngrPktsAg中的入口丢弃数据包速率

简介

本文档介绍每种故障类型，以及发生此故障时的操作步骤。

背景信息

在思科以应用为中心的基础设施(ACI)交换矩阵的正常运行期间，管理员可以看到特定类型的数据包丢弃的故障。

托管对象

在思科ACI中，所有故障均在托管对象(MO)下发生。 例如，故障F11245 — 入口丢弃数据包速率(l2IngrPktsAg15min:dropRate)与MO l2IngrPktsAg15min中的参数dropRate有关。

本节介绍与丢弃数据包故障相关的托管对象(MO)的一些示例。

	示例	描述	示例参数	MO示例 引起哪些故障
l2IngrPkts	l2IngrPkts5min l2IngrPkts15min l2IngrPkts1h 等等。	这表示每个VLAN在每个时间段内的入口数据包统计信息。	dropRate floodRate multicastRate unicastRate	vlanCktEp(VLAN)
l2IngrPktsAg	l2IngrPktsAg15min l2IngrPktsAg1h l2IngrPktsAg1d 等等。	这表示每个EPG、BD、VRF等的入口数据包统计信息。例如，EPG统计信息表示属于EPG的VLAN统计信息的聚合。	dropRate floodRate multicastRate unicastRate	fvAEPg(EPG) fvAp（应用配置文件） fvBD(BD) l3extOut(L3OUT)
eqptIngrDropPkts	eqptIngrDropPkts15min eqptIngrDropPkts1h eqptIngrDropPkts1d 等等。	这表示每个接口在每个时间段内的入口丢弃数据包统计信息。	*1转发速率 *1错误率 *1 bufferRate	l1PhysIf（物理端口） pcAggrIf（端口通道）

*1:eqptIngrDropPkts中的这些计数器可能会由于若干Nexus 9000平台中的ASIC限制而错误地引发，因为SUP_REDIRECT数据包被记录为转发丢弃。另请参阅Cisco Bug ID CSCvo68407 和思科漏洞ID CSCvn72699 了解更多详情和固定版本。

硬件丢弃计数器类型

在ACI模式下运行的Nexus 9000交换机上，ASIC上有三个主要硬件计数器用于入口接口丢弃原因。

l2IngrPkts和l2IngrPktsAg中的dropRate包括这些计数器。eqptIngrDropPkts的表中的三个参数(forwardingRate、errorRate、bufferRate)代表每个三个接口计数器。

转发

转发丢弃是在ASIC的查找块(LU)上丢弃的数据包。在LU块中，根据数据包报头信息做出数据包转发决策。如果决定丢弃数据包，则计算转发丢弃。出现这种情况的原因有很多。主要包括：

SECURITY_GROUP_DENY

由于缺少允许通信的合同而丢弃。

当数据包进入交换矩阵时，交换机将查看源EPG和目标EPG，以查看是否存在允许此通信的合同。如果源和目标位于不同的EPG中，并且没有允许此数据包类型的合同，则交换机将丢弃该数据包，并将其标记为SECURITY_GROUP_DENY。这会增加Forward Drop计数器。

VLAN_XLATE_MISS

由于VLAN不当而丢弃。

当数据包进入交换矩阵时，交换机将查看数据包，以确定端口上的配置是否允许该数据包。 例如，帧以802.1Q标记10进入交换矩阵。如果交换机端口上有VLAN 10，它会检查内容并根据目标MAC做出转发决策。但是，如果VLAN 10不在端口上，它会丢弃该端口并将其标记为VLAN_XLATE_MISS。这会增加Forward Drop计数器。

使用XLATE或Translate的原因是，在ACI中，枝叶交换机采用具有802.1Q封装的帧，并将其转换为一个新的VLAN，该新的VLAN用于VXLAN和交换矩阵内的其他规范化。如果帧来自未部署的VLAN，则转换失败。

ACL_DROP

因为sup-tcam而下降。

aci交换机中的sup-tcam包含要在正常L2/L3转发决策之上应用的特殊规则。sup-tcam中的规则是内置的，用户不可配置。Sup-tcam规则的主要目的是处理某些异常或某些控制平面流量，而不是由用户检查或监控。当数据包达到sup-tcam规则且规则为丢弃数据包时，丢弃的数据包将计为ACL_DROP，并递增Forward Drop计数器。发生这种情况时，通常意味着数据包将根据基本的ACI转发主体进行转发。

即使丢弃名称为ACL_DROP，此ACL与可在独立NX-OS设备或任何其他路由/交换设备

上配置的正常访问控制列表不同。

SUP_REDIRECT

这不是一滴水。

即使数据包已正确处理并转发到CPU，超级重定向的数据包（例如，CDP/LLDP/UDLD/BFD等）也可以计为转发丢弃。

这是由于基于云扩展ASIC(EX/FX/FX2/FX3/GX/GX2)的ACI N9k交换机的ASIC实施限制。

错误

当交换机在一个前面板接口上收到无效帧时，该帧将作为错误丢弃。 示例包括带有FCS或CRC错误的帧。在查看上行链路/下行链路枝叶端口或主干端口时，最好使用show interface检查FCS/CRC错误。但是，在正常操作下，预期会在枝叶的上行链路/下行链路端口或主干端口上看到错误数据包增加，因为此计数器还包括系统修剪的帧，且预期不会从接口发送出去的帧。

示例：路由数据包的TTL故障，相同的接口广播/泛洪帧。

缓冲区

当交换机收到帧时，没有可用于入口或出口的缓冲区信用数，该帧可能会随缓冲区一起丢弃。 这通常提示网络中的某处存在拥塞。显示故障的链路可能已满，或者包含目的地的链路可能拥塞。

在CLI中查看丢弃统计信息

托管对象

安全外壳(SSH)到其中一个APIC并运行这些命令。

apic1# moquery -c l2IngrPktsAg15min

这为此类l2IngrPktsAg15min提供所有对象实例。

下面是一个使用过滤器查询特定对象的示例。在本示例中，过滤器仅显示具有属性dn的对象，其中包含tn-TENANT1/ap-APP1/epg-EPG1。

此外，此示例使用egrep仅显示所需的属性。

示例输出1:租户TENANT1、应用配置文件APP1、epg1的EPG计数器对象(l2IngrPktsAg15min)。

```
apic1# moquery -c l2IngrPktsAg15min -f 'l2.IngrPktsAg15min.dn*"tn-TENANT1/ap-APP1/epg-EPG1"' | egrep 'dn

dn               : uni/tn-TENANT1/ap-APP1/epg-EPG1/CDl2IngrPktsAg15min
dropPer          : 30                                <--- number of drop packet in the current periodic in
dropRate         : 0.050000                          <--- drop packet rate = dropPer(30) / periodic interv
repIntvEnd       : 2017-03-03T15:39:59.181-08:00     <--- periodic interval = repIntvEnd - repIntvStart
repIntvStart     : 2017-03-03T15:29:58.016-08:00                            = 15:39 - 15:29
                                                                            = 10 min = 600 sec
```

或者，如果您知道对象dn，则可以使用其他选项-d而不是-c来获取特定对象。

示例输出2:租户TENANT1、应用配置文件APP1、epg2的EPG计数器对象(l2IngrPktsAg15min)。

```
apic1# moquery -d uni/tn-TENANT1/ap-APP1/epg-EPG2/CDl2IngrPktsAg15min | egrep 'dn|drop[P,R]|rep'
dn               : uni/tn-jw1/BD-jw1/CDl2IngrPktsAg15min
dropPer          : 30
dropRate         : 0.050000
repIntvEnd       : 2017-03-03T15:54:58.021-08:00
repIntvStart     : 2017-03-03T15:44:58.020-08:00
```

硬件计数器

如果您看到故障，或希望使用CLI检查交换机端口上的丢包，最好通过查看硬件中的平台计数器来达到此目的。大多数计数器（但不是所有计数器）使用show interface显示。 只能使用平台计数器查看3个主要丢弃原因。 要查看这些信息，请执行以下步骤：

枝叶

通过SSH连接到枝叶并运行这些命令。

ACI-LEAF# vsh_lc
module-1# show platform internal counters port <X>
*其中X代表端口号

ethernet 1/31的输出示例：

<#root>

```
ACI-LEAF#

vsh_lc

vsh_lc
module-1#
module-1#

show platform internal counters port 31

Stats for port 31
(note: forward drops includes sup redirected packets too)
IF          LPort                    Input                         Output
                                Packets       Bytes          Packets    Bytes
eth-1/31    31  Total          400719   286628225      2302918    463380330
                Unicast        306610   269471065       453831     40294786
                Multicast           0           0      1849091    423087288
                Flood           56783     8427482            0            0
            Total Drops         37327                        0
                Buffer              0                        0
                Error               0                        0
                Forward         37327
                LB                  0
                AFD RED                                      0
                    ----- snip -----
```

主干

对于盒式主干(N9K-C9336PQ)，它与枝叶完全相同。

对于模块化主干（N9K-C9504等），必须先连接特定的线卡，然后才能查看平台计数器。
通过SSH连接到主干，并运行以下命令：

ACI-SPINE# vsh

ACI-SPINE#连接模块<X>

module-2# show platform internal counters port <Y>。

*其中X代表您要查看的线路卡的模块编号

Y表示端口号

ethernet 2/1的输出示例：

<#root>

ACI-SPINE#

vsh

Cisco iNX-OS Debug Shell
This shell can only be used for internal commands and exists
for legacy reasons. User can use ibash infrastructure as this
can be deprecated.
ACI-SPINE#
ACI-SPINE#

attach module 2

Attaching to module 2 ...
To exit type 'exit', to abort type '$.'
Last login: Mon Feb 27 18:47:13 UTC 2017 from sup01-ins on pts/1
No directory, logging in with HOME=/
Bad terminal type: "xterm-256color". can assume vt100.
module-2#
module-2#

show platform internal counters port 1

Stats for port 1
(note: forward drops includes sup redirected packets too)

IF	LPort		Input		Output	
			Packets	Bytes	Packets	Bytes
eth-2/1	1	Total	85632884	32811563575	126611414	25868913406
		Unicast	81449096	32273734109	104024872	23037696345
		Multicast	3759719	487617769	22586542	2831217061
		Flood	0	0	0	0
		Total Drops	0		0	

Buffer 0

0

Error 0

0

```
Forward          0

                   LB              0
                   AFD RED                                 0
                        ----- snip -----
```

故障

F112425 — 入口丢弃数据包速率(l2IngrPktsAg15min:dropRate)

描述:

此故障的常见原因之一是第2层数据包因转发丢弃原因而被丢弃。原因有很多，但最常见的是：

在某些平台(请参阅Cisco Bug ID CSCvo68407)上，存在需要重定向到CPU（例如，CDP/LLDP/UDLD/BFD等）的第2层数据包被记录为转发丢弃以及被复制到CPU的限制。这是由于这些型号中使用的ASIC的限制所致。

分辨率：

所描述的丢包纯粹是表面的，因此，最佳实践建议是增加故障的阈值，如“统计阈值”部分所示。为此，请参阅统计阈值中的说明。

F100264 — 入口缓冲区丢弃数据包速率(eqptIngrDropPkts5min:bufferRate)

描述:

当数据包在带有原因缓冲区的端口上丢弃时，此故障可能会增加。如前所述，当接口在入口或出口方向出现拥塞时，通常会发生这种情况。

分辨率：

此故障表示环境中由于拥塞而丢弃的实际数据包。丢弃的数据包可能导致在ACI交换矩阵中运行的应用程序出现问题。网络管理员可以隔离数据包流，并确定拥塞是由意外的流量流、低效的负载均衡等引起的，还是由这些端口上的预期利用率引起的。

F100696 — 入口转发丢弃数据包(eqptIngrDropPkts5min:forwardingRate)

注意：前面提到的F11245的ASIC限制可能会导致这些故障也增加。请参阅思科漏洞 ID CSCvo68407 以获取更多详细信息。

此故障由几种情况引起。最常见的是：

说明1 — 主干丢弃

如果在主干接口上发现此故障，则可能是由于流向未知终结点的流量。当ARP或IP数据包转发到主干进行代理查找时，并且终端在交换矩阵中未知时，将生成一个特殊收集数据包，并将其发送到相应BD（内部）组播组地址上的所有枝叶。这将触发来自网桥域(BD)中每个枝叶的ARP请求以发现终端。 由于限制，枝叶收到的收集数据包也会再次反射回交换矩阵，并触发连接到枝叶的主干链路上的转发丢弃。此场景中的转发丢弃仅在第1代主干硬件上递增。

决议1

由于已知问题是由向ACI交换矩阵发送不必要数量的未知单播流量导致的，因此需要找出导致此问题的设备，并查看是否可以阻止此问题。这通常是由出于监控目的扫描或探测子网上IP地址的设备引起的。为了找出发送此流量的IP，请通过SSH连接到主干接口的枝叶上，该枝叶显示了故障。

您可以在此处运行此命令以查看触发收集数据包的源IP地址(sip):

<#root>

```
ACI-LEAF# show ip arp internal event-history event | grep glean | grep sip | more
    [116] TID 11304:arp_handle_inband_glean:3035:

log_collect_arp_glean

;sip =

192.168.21.150

;dip =

 192.168.20.100

;info = Received glean packet is an IP packet
    [116] TID 11304:arp_handle_inband_glean:3035: log_collect_arp_glean;sip = 192.168.21.150;dip = 192.
```

在此示例输出中，收集数据包由192.168.21.150触发，建议查看是否可以缓解此问题。

说明2 — 枝叶丢弃

如果在枝叶接口上看到此故障，最可能的原因是提到了SECURITY_GROUP_DENY丢弃。

决议2

ACI枝叶会保留因违规而被拒绝的数据包日志。此日志不会捕获所有日志以保护CPU资源，但是，它仍会提供大量日志。

要获取所需的日志，如果发生故障的接口是port-channel的一部分，则需要对port-channel使用此命令和grep。否则，物理接口可能变灰。

根据合同删除的数量，此日志可以快速回滚。

<#root>

```
ACI-LEAF# show logging ip access-list internal packet-log deny | grep port-channel2 | more
[ Sun Feb 19 14:16:12 2017 503637 usecs]: CName: jr:sb(VXLAN: 2129921), VlanType: FD_VLAN, Vlan-Id: 59,

SIP: 192.168.21.150, DIP: 192.168.20.3

, SPort: 0, DPort: 0,

Src Intf: port-channel2

,

Pr
oto: 1

, PktLen: 98
[ Sun Feb 19 14:16:12 2017 502547 usecs]: CName: jr:sb(VXLAN: 2129921), VlanType: FD_VLAN, Vlan-Id: 59,
oto: 1, PktLen: 98
```

在本例中，192.168.21.150正在尝试将ICMP消息（IP协议号1）发送到192.168.20.3。但是，两个EPG之间没有允许ICMP的合同，因此数据包被丢弃。如果应允许ICMP，则可以在两个EPG之间添加合同。

统计信息阈值

本节介绍如何更改可能引发丢弃计数器故障的统计信息对象的阈值。

每个对象的统计信息阈值（例如，l2IngrPkts、eqptIngrDropPkts）通过监控策略针对各种对象进行配置。

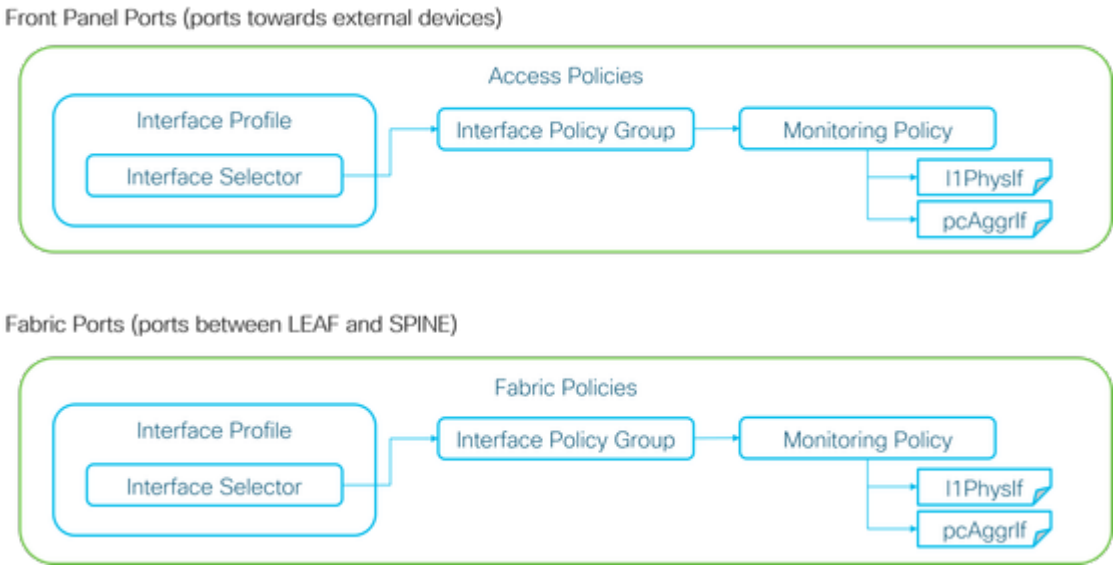
如开始处的表中所述，eqptIngrDropPkts通过监控策略在l1PhysIf对象下监控。

eqptIngrDropPkts中的转发丢弃数据包速率

这里有两个部分。

+访问策略(指向外部设备的端口。也称为前面板端口)。

+交换矩阵策略(枝叶和主干之间的端口。也称为交换矩阵端口)。



如上图所示，可以通过接口策略组为每个端口对象(l1PhysIf、pcAggrIf)分配自己的监控策略。

默认情况下，在APIC GUI中的Fabric > Access Policies和Fabric > Fabric Policies下都有默认监控策略。这些默认监控策略分别分配给所有端口。访问策略下的默认监控策略用于前面板端口，交换矩阵策略下的默认监控策略用于交换矩阵端口。

除非要求更改每个端口的阈值，否则可以直接修改每个部分的默认监控策略，以将更改应用于所有前面板端口和/或交换矩阵端口。

此示例将更改交换矩阵端口（交换矩阵策略）上eqptIngrDropPkts中的Forward Drop阈值。 在Fabric > Access Policies下对前面板端口执行相同操作。

1.导航至交换矩阵>交换矩阵策略>监控策略。

2.右键单击并选择创建监控策略。

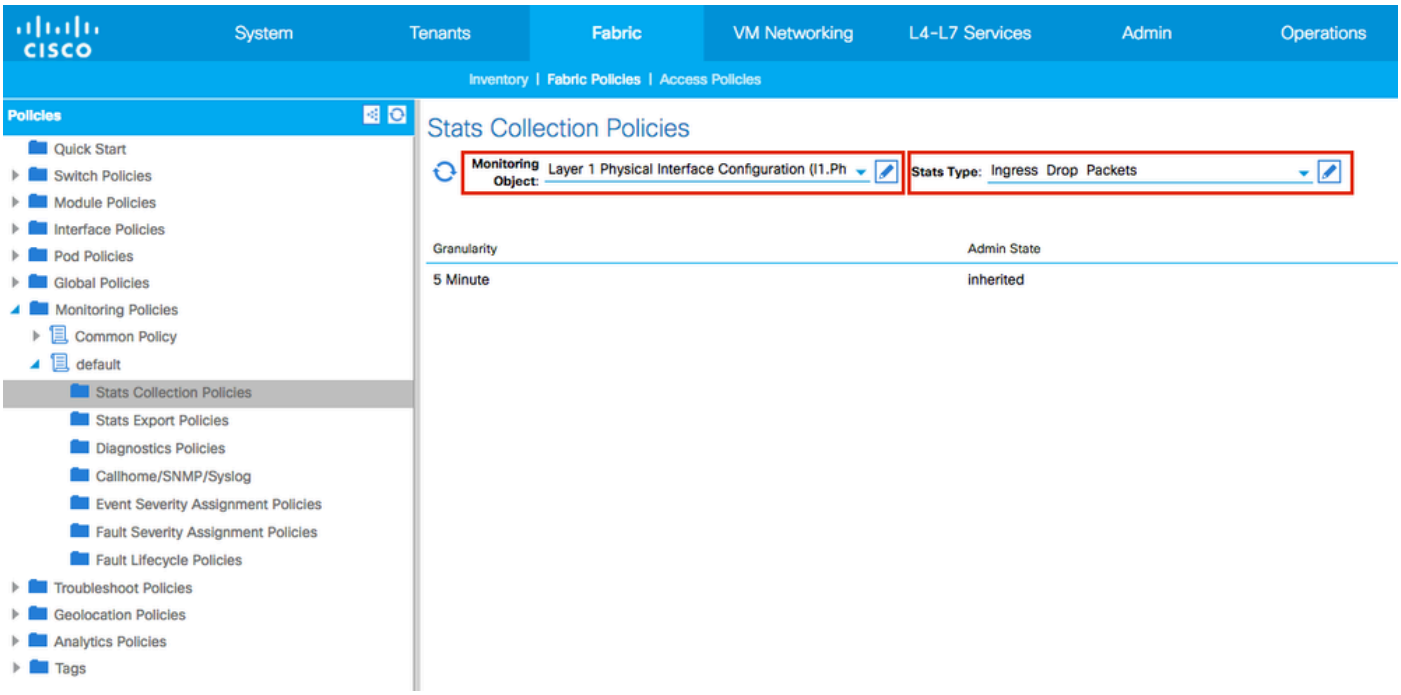
(如果阈值更改可以应用于所有交换矩阵端口，请导航到default，而不是创建新端口。)

3.展开新的Monitoring策略或默认策略，然后导航到Stats Collection Policies。

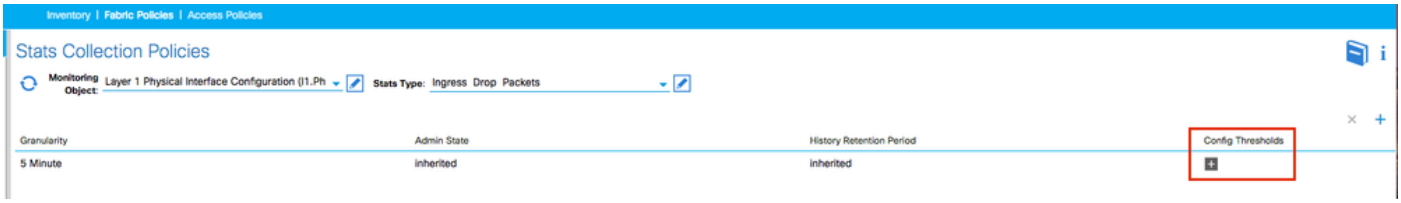
4.单击右侧窗格上Monitoring Object的铅笔图标，选择Layer 1 Physical Interface Configuration(l1.PhysIf)。

（使用默认策略时，可以跳过第4步。）

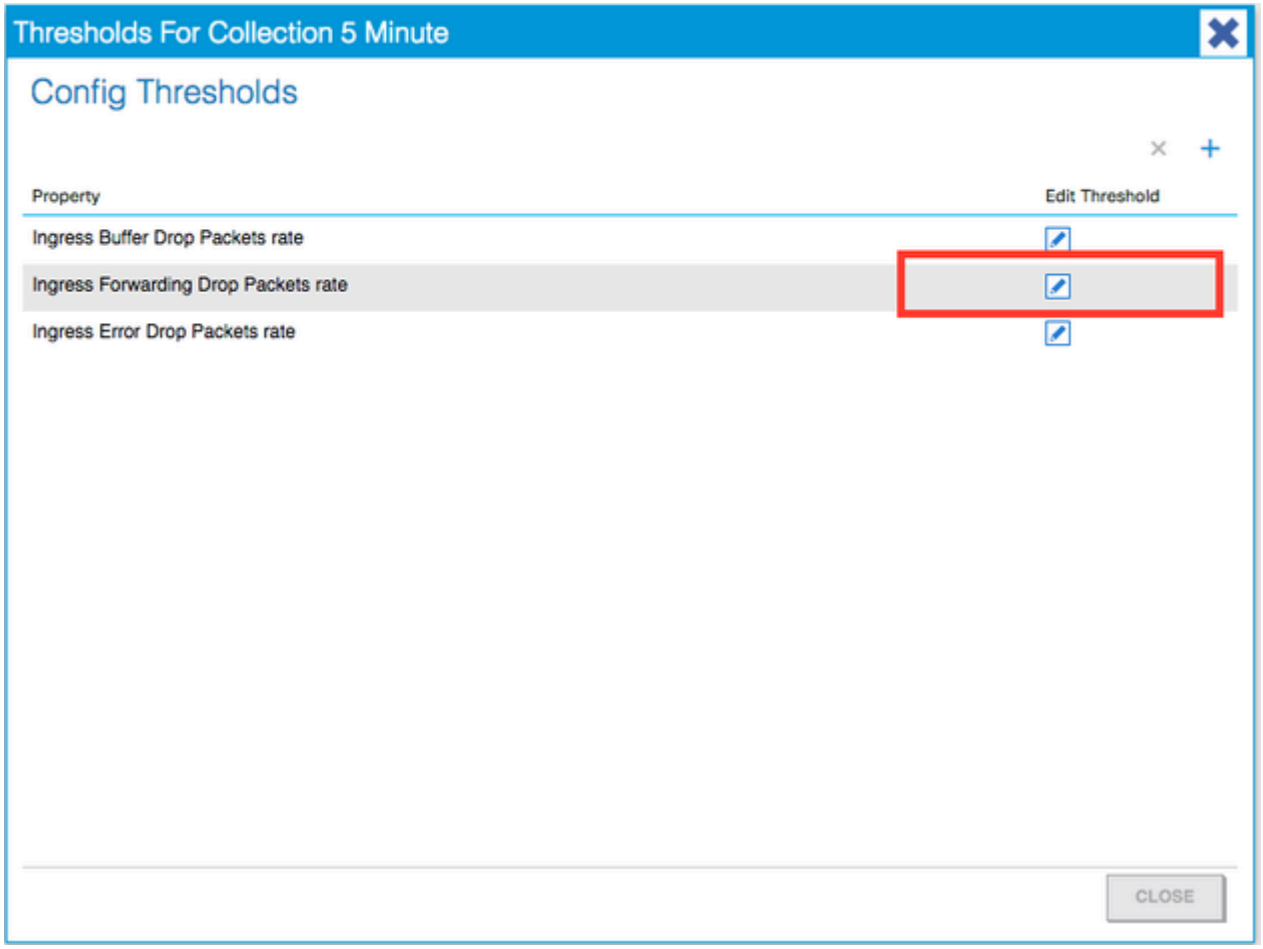
5.从右侧窗格的Monitoring Object（监控对象）下拉菜单，选择Layer 1 Physical Interface Configuration(l1.PhysIf)和Stats Type，然后选择Ingress Drop Packets。



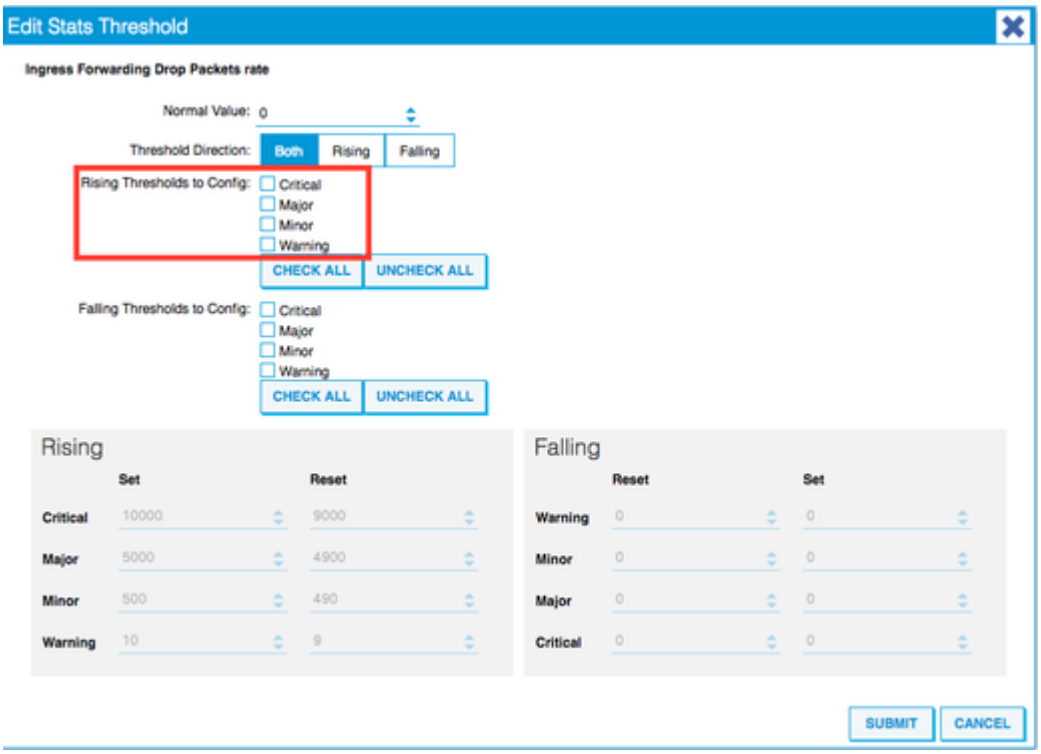
6.单击“配置阈值”旁边的+。



7.编辑Threshold for Forwarding Drop。

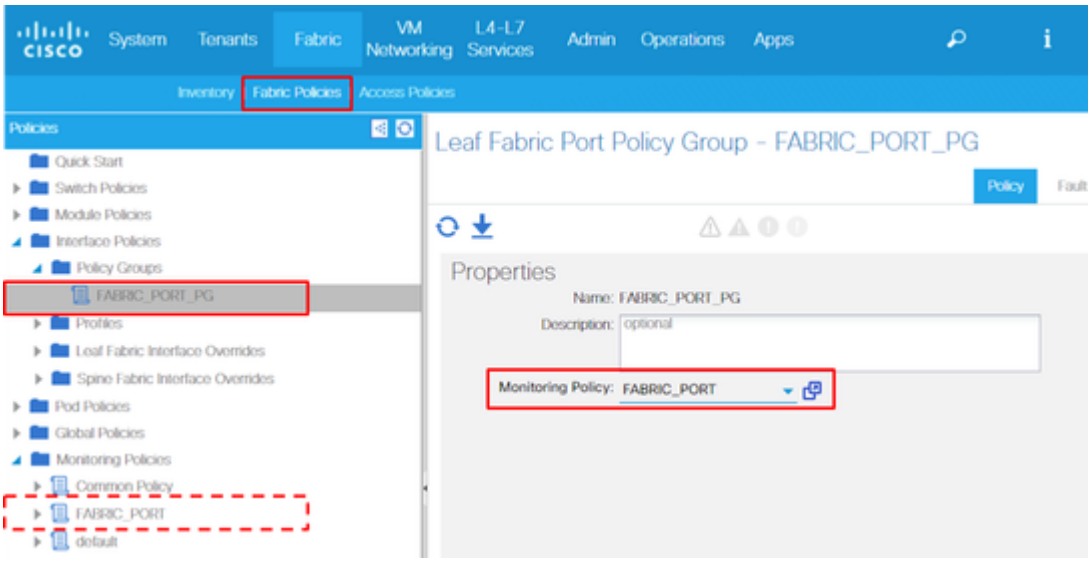


8.建议禁用递增阈值，以配置转发丢弃率的关键、主要、次要和警告。



9.将此新监控策略应用于所需端口的接口策略组。不要忘记在交换矩阵策略中相应地配置接口配置文件、交换机配置文件等等。

（使用默认策略时，可以跳过第9步。）

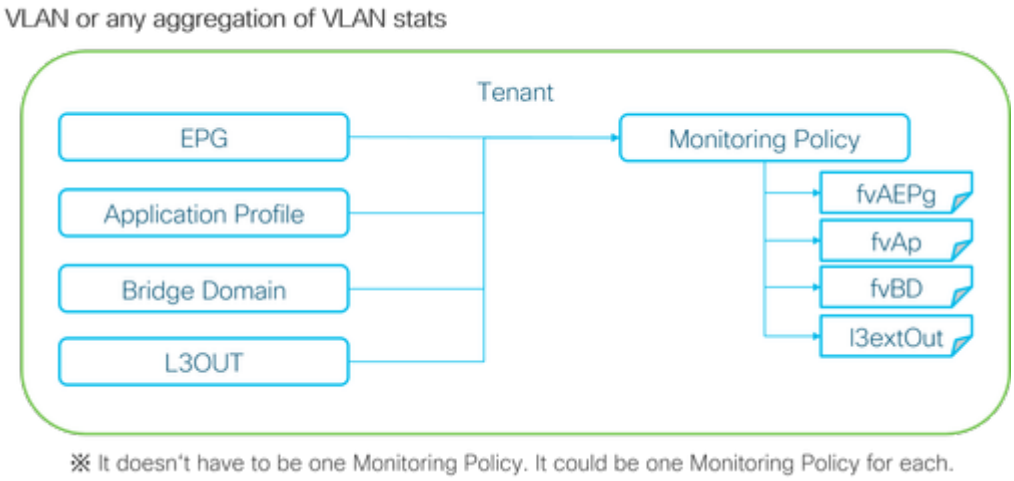


10.如果这是用于前面板端口（访问策略），请对汇聚接口(pc.AggrIf)执行与第1层物理接口配置(l1.PhysIf)相反的相同操作，以便此新的监控策略可以应用于端口通道和物理端口。

（使用默认策略时，可以跳过第10步。）

l2IngrPktsAg中的入口丢弃数据包速率

有多个部分。



如上一幅图片所示，l2IngrPktsAg在许多对象下被监控。此图片只显示一些示例，但不显示l2IngrPktsAg的所有对象。但是，统计信息的阈值是通过监控策略以及l1PhysIf或pcAggrIf下的eqptIngrDropPkts配置的。

如上图所示，可以为每个对象(EPG(fvAEPg)、网桥域(fvBD)等)分配自己的监控策略。

默认情况下，租户下的所有这些对象均使用Tenant > common > Monitoring Policies > default下的默认监控策略，除非另外配置。

除非要求更改每个组件的阈值，否则可以直接修改租户common下的默认监控策略，以将更改应用于所有相关组件。

本示例更改网桥域上l2IngrPktsAg15min中的入口丢弃数据包速率的阈值。

1.导航到租户>（租户名称）>监控策略。

（如果使用默认监控策略，或者需要在租户间应用新的监控策略，则租户需要通用）

2.右键单击并选择创建监控策略。

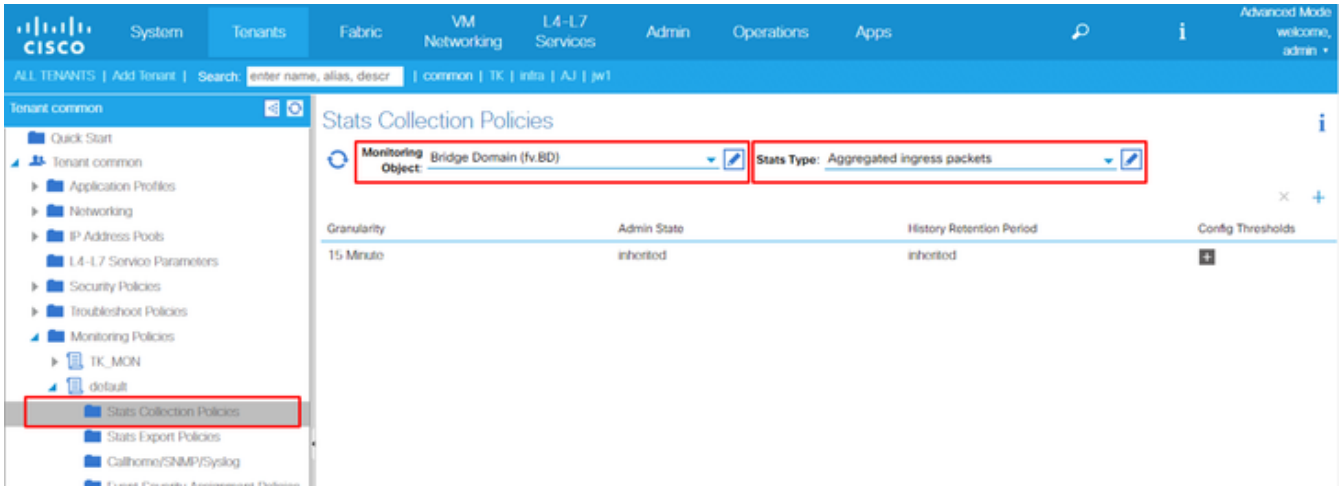
(如果阈值更改可应用于所有组件，请导航到default，而不是创建新组件。)

3.展开新的Monitoring策略或默认策略，然后导航到Stats Collection Policies。

4.单击右侧窗格上Monitoring Object的铅笔图标，选择Bridge Domain(fv.BD)。

（使用默认策略时，可以跳过第4步。）

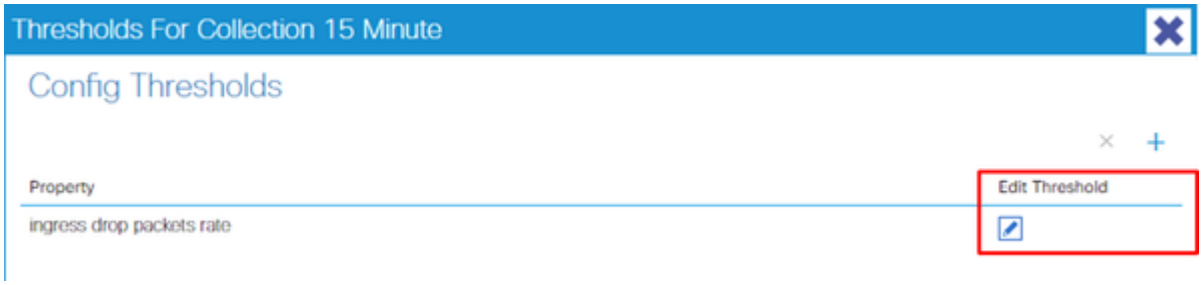
5.从右侧窗格的Monitoring Object下拉列表中选择Bridge Domain(fv.BD)和Stats Type，然后选择Aggregated ingress packets。



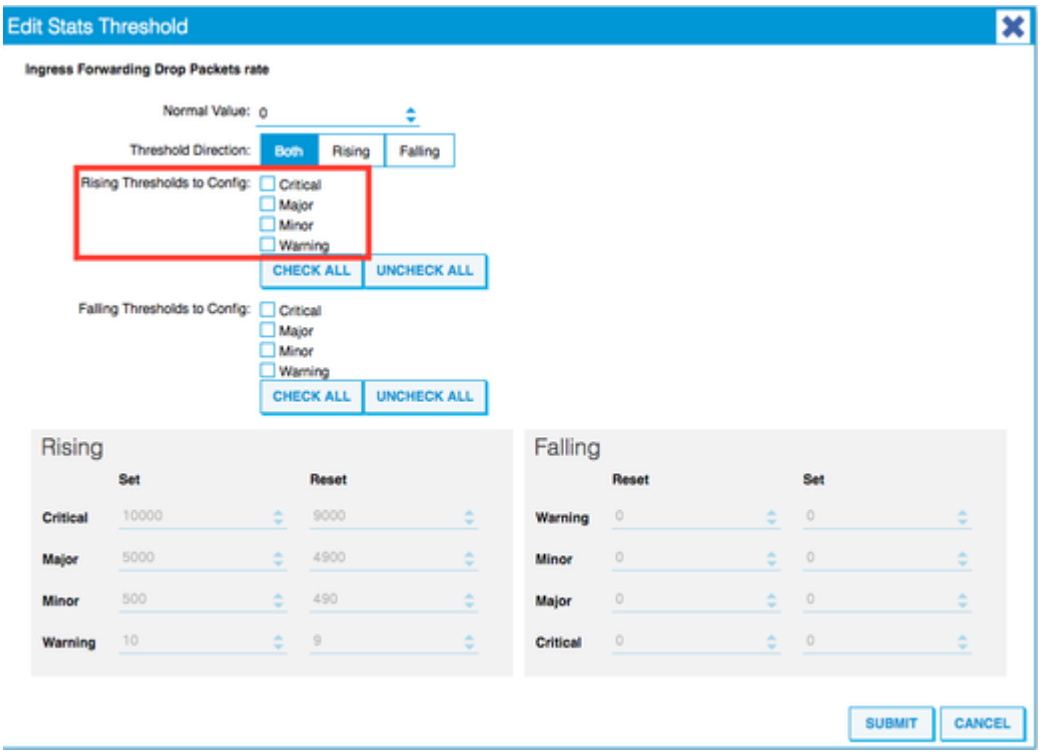
6.单击“配置阈值”旁边的+。



7.编辑Threshold for Forwarding Drop。

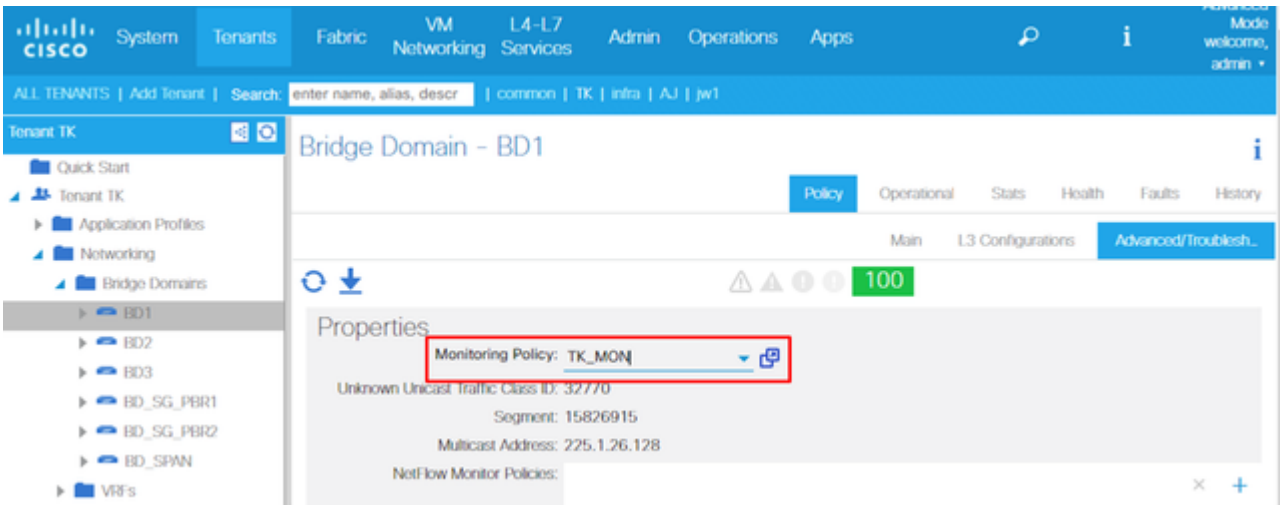


8.建议禁用递增阈值，以配置转发丢弃率的关键、主要、次要和警告。



9.将此新监控策略应用于需要更改阈值的网桥域。

（使用默认策略时，可以跳过第9步。）



NOTE:
非默认监控策略不能具有默认监控策略中存在的配置。如果需要保持这些配置与默认监控策略相同，用户需要检查默认监控策略配置，并在非默认监控策略上手动配置相同的策略。

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）。