

解释ACI中的数据包包丢弃故障

目录

[简介](#)

[托管对象](#)

[硬件丢弃计数器类型](#)

[转发](#)

[Error](#)

[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\(I2IngrPktsAg15min : dropRate\)](#)

[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](#)

[I2IngrPktsAg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。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故障。

托管对象

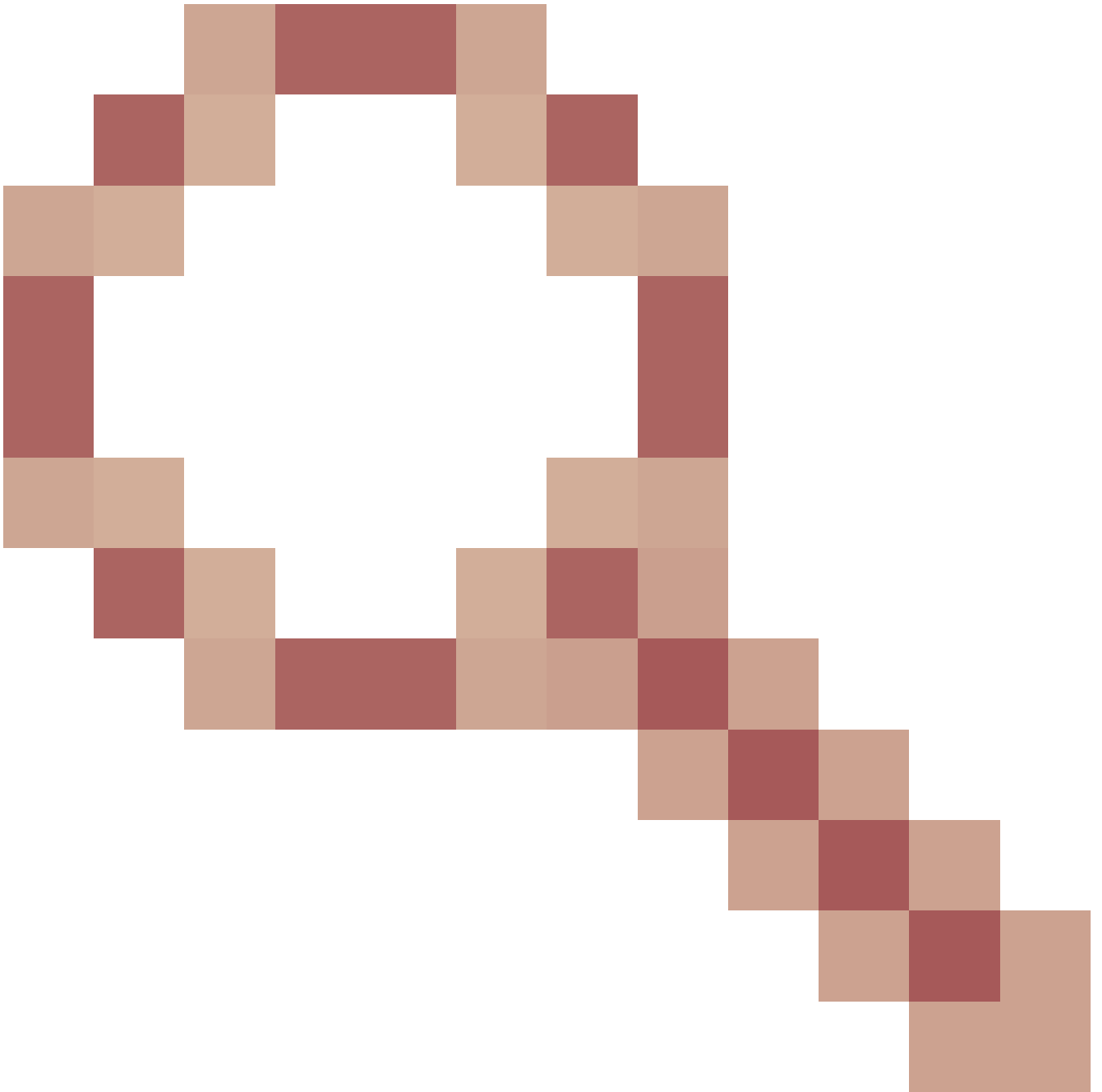
在思科ACI中，所有故障均在托管对象(MO)下引发。例如，故障F11245 - 入口丢弃数据包速率(I2IngrPktsAg15min : dropRate)与MO I2IngrPktsAg15min中的参数dropRate有关。

本节介绍一些与丢弃数据包故障相关的托管对象(MO)示例。

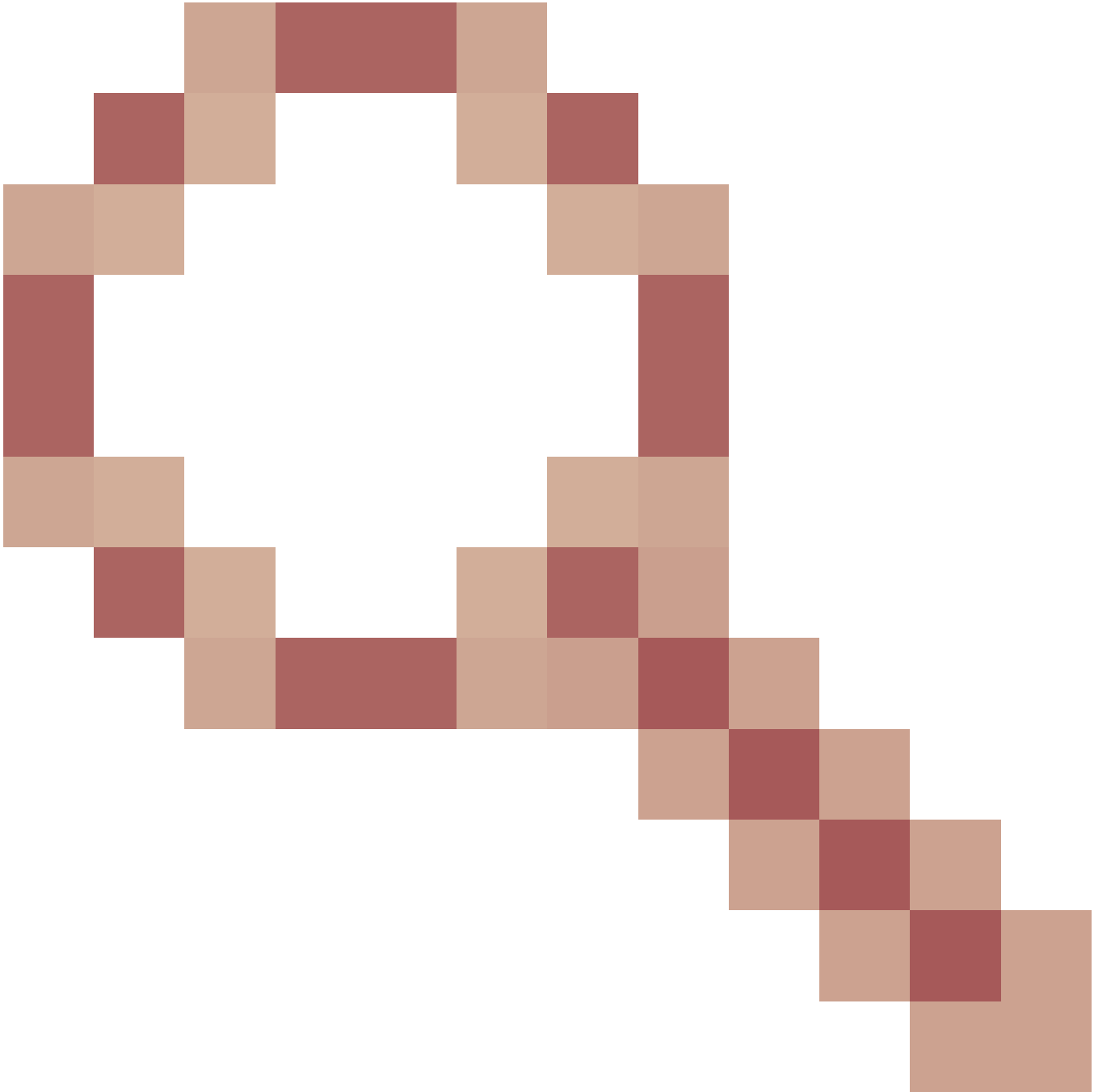
	示例	描述	示例参数	MO示例引起哪些故障
I2IngrPkts	I2IngrPkts5min I2IngrPkts15min	这表示每个VLAN在每个时间段内的入口数据包统计信息。	dropRate floodRate	vlanCktEp (VLAN)

	l2IngrPkts1h 等等。		multicastRate unicastRate	
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统计信息的汇聚。	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 forwardingRate *1 errorRate *1 bufferRate	l1PhysIf (物理端口) pcAggrIf (端口通道)

*1：由于Nexus 9000平台中的ASIC限制，eqptIngrDropPkts中的这些计数器可能会错误地引发，因为SUP_REDIRECT数据包被记录为转发丢弃。有关更多详细信息和修复的版本，另请参阅Cisco Bug ID [CSCvo68407](#)



和Cisco Bug ID [CSCvn72699](#)



。

硬件丢弃计数器类型

在以ACI模式运行的Nexus 9000交换机上，ASIC上有3个主要硬件计数器显示入口接口丢弃原因。

I2IngrPkts和I2IngrPktsAg中的dropRate包含这些计数器。表中用于eqptIngrDropPkts的三个参数(forwardingRate、errorRate、bufferRate)分别代表三个接口计数器。

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，则会计算转发丢弃。这有可能发生的原因有很多，但让我们谈谈主要原因：

SECURITY_GROUP_DENY

丢弃，因为缺少允许通信的合同。

当数据包进入交换矩阵时，交换机会查看源和目标EPG，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，则交换机会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不当而造成丢弃。

当数据包进入交换矩阵时，交换机查看该数据包，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。如果交换机在端口上有VLAN 10，它会检查内容，并根据目的MAC做出转发决策。但是，如果VLAN 10不在端口上，则会将其丢弃并将其标记为VLAN_XLATE_MISS。这会增加Forward Drop计数器。

使用XLATE或Translate的原因是，在ACI中，枝叶交换机采用带有802.1Q封装的帧，并将其转换为用于交换矩阵内部VXLAN和其他规范化的新VLAN。如果帧在未部署VLAN的情况下传入，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，不可由用户配置。Sup-tcam规则的目标主要是处理某些异常或控制平面流量，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ACI转发原则进行转发。

即使丢弃名称为ACL_DROP，此ACL也不与可在独立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相同。

SUP_REDIRECT

这不是一滴一滴的血。

即使数据包被正确处理并转发到CPU，管理引擎重定向的数据包（例如，CDP/LLDP/UDLD/BFD等）也可以算作转发丢弃。

这发生在-EX、-FX和-FX2平台中，例如N9K-C93180YC-EX或N9K-C93180YC-FX。这些不能计为丢弃，但这是因为-EX/-FX/-FX2平台中的ASIC限制。

Error

当交换机在一个前面板接口上收到无效帧时，该帧会作为错误丢弃。示例包括带有FCS或

CRC错误的帧。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，预计会在枝叶的上行/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已修剪的帧，且预期不会从接口发送出去。

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，并且没有可用于入口或出口的缓冲区信用时，该帧将随缓冲区一起丢弃。这通常提示网络中的某个位置存在拥塞。表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

使用安全外壳(SSH)连接到其中一个APIC并运行以下命令。

```
apic1# moquery -c l2IngrPktsAg15min
```

这提供了此类l2IngrPktsAg15min的所有对象实例。

以下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，其中包括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、epg 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 i
dropRate          : 0.050000         <--- drop packet rate = dropPer(30) / periodic inter
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代替-c获取特定对象。

示例输出2：租户TENANT1、应用配置文件APP1、epg 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/CD12IngrPktsAg15min
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

如果发现故障，或者想要使用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检查。大多数（但并非所有）计数器是使用show interface显示的。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连接到枝叶设备并运行以下命令。

```
ACI-LEAF# vsh_lc
module-1# show platform internal counters port <X>
* 其中X表示端口号
```

etherent 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代表您要查看的板卡的模块号

Y表示端口号

以太网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  
will 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". Will 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层数据包由于转发丢弃原因而被丢弃。原因有很多，但最常见的是：

在某些平台上(请参阅Cisco Bug ID [CSCvo68407](#))，存在需要重定向到CPU的L2数据包（例如，CDP/LLDP/UDLD/BFD等）、作为转发丢弃记录以及复制到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：

描述的丢包纯属表面效果，因此最佳实践建议是增加故障的阈值，如“统计信息阈值”部分所示。为此，请参阅统计信息阈值中的说明。

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


描述:

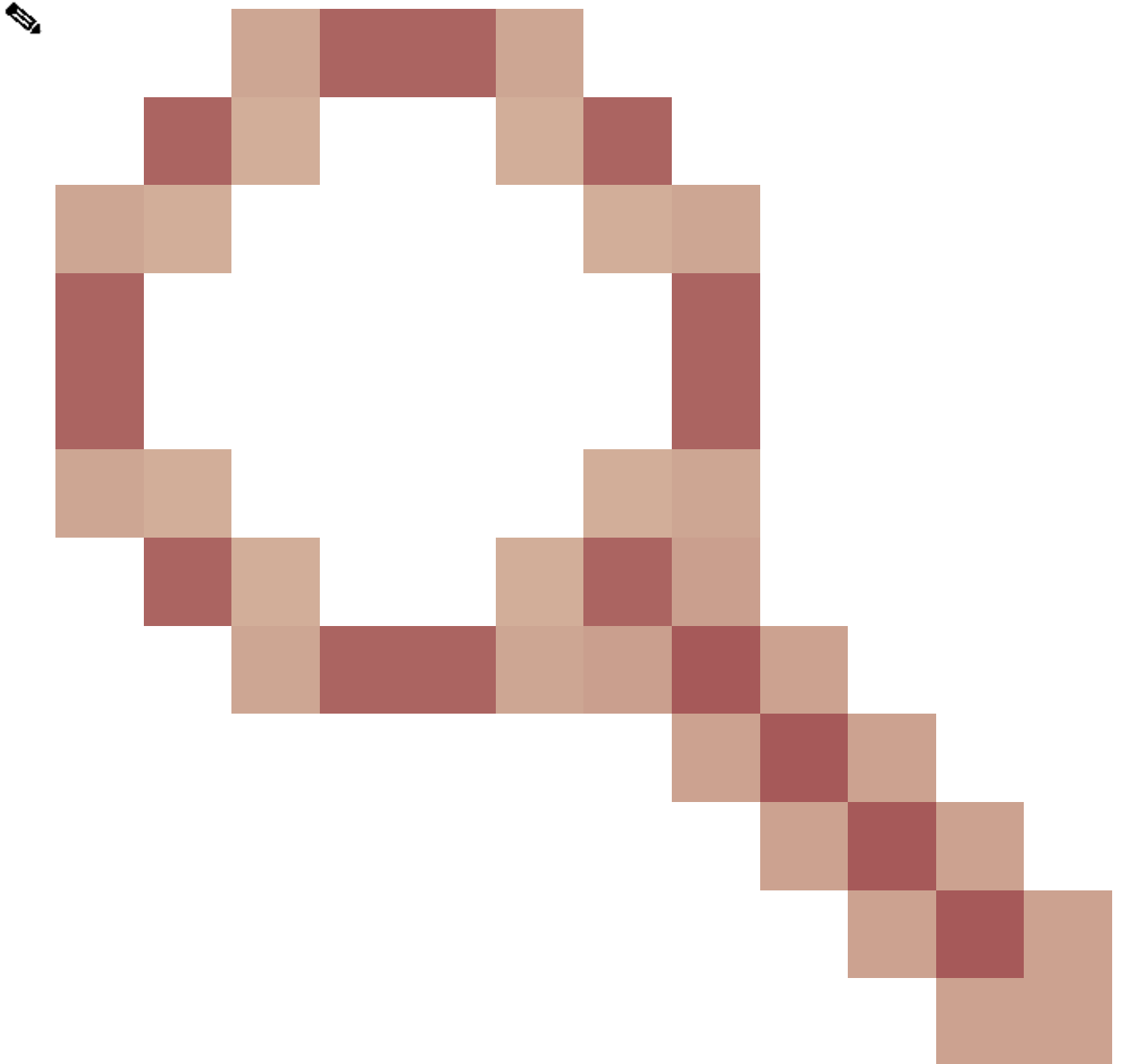
当数据包在带有原因缓冲区的端口上丢弃时，此故障可能会增加。如前所述，当接口在入口或出口方向出现拥塞时，通常会发生这种情况。

分辨率：

此故障表示因拥塞而实际丢弃的环境中数据包。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流量、低效负载均衡等原因还是由这些端口的预期利用率引起的。

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

 注意：之前提到的F11245 ASIC限制也可能导致这些故障增加。有关更多详细信息，请参阅Cisco Bug ID [CSCvo68407](#)



。

此故障由几种情况引起。最常见的是：

说明1) 脊柱丢弃

如果在主干接口上发现此故障，则可能是由于流向未知端点的流量。当ARP或IP数据包转发到主干进行代理查找，并且终端在交换矩阵中未知时，将生成一个特殊的收集数据包，并将其发送到相应BD（内部）组播组地址上的所有枝叶。这会触发来自网桥域(BD)中每个枝叶的ARP请求以发现终端。由于限制，枝叶接收的收集数据包也会再次反射回交换矩阵，并在连接到枝叶的主干链路上触发转发丢弃。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，因此需要确定导

致此问题的设备，并查看是否可以阻止此问题。这通常是由于出于监控目的扫描或探测子网上IP地址的设备引起的。要查找发送此流量的IP地址，请将SSH连接到主干接口的枝叶上，该枝叶将显示此故障。

从那里，您可以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.
```

在本示例输出中，收集数据包由192.168.21.150触发，建议查看是否能缓解此问题。

说明2)枝叶丢弃

如果在枝叶接口上发现此故障，最可能的原因是提到的SECURITY_GROUP_DENY丢包。

决议2)

ACI枝叶会保留由于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，物理接口可能会被损坏。

根据合同丢弃的数量，此日志可以快速滚动更新。

```
<#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 (例如，l2IngrPkts、eqptIngrDropPkts) 的阈值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在l1PhysIf对象下通过监控策略进行监控。

eqptIngrDropPkts中的转发丢弃数据包速率

这个有两个部分。

+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。又名交换矩阵端口)

Front Panel Ports (ports towards external devices)



Fabric Ports (ports between LEAF and SPINE)



可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，如上图所示。

默认情况下，APIC GUI中的Fabric > Access Policies和Fabric > Fabric Policies下都有默认监控

策略。这些默认监控策略将分别分配给所有端口。“访问策略”(Access Policies)下的默认监控策略用于前面板端口，而“交换矩阵策略”(Fabric Policies)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。

本示例是更改交换矩阵端口（交换矩阵策略）上eqptIngrDropPkts中的转发丢弃阈值。在Fabric > Access Policies下为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. 展开新的监控策略或默认值，然后导航到统计信息收集策略。

4. 在右窗格中单击Monitoring Object的铅笔图标，选择Layer 1 Physical Interface Configuration (I1.PhysIf)。

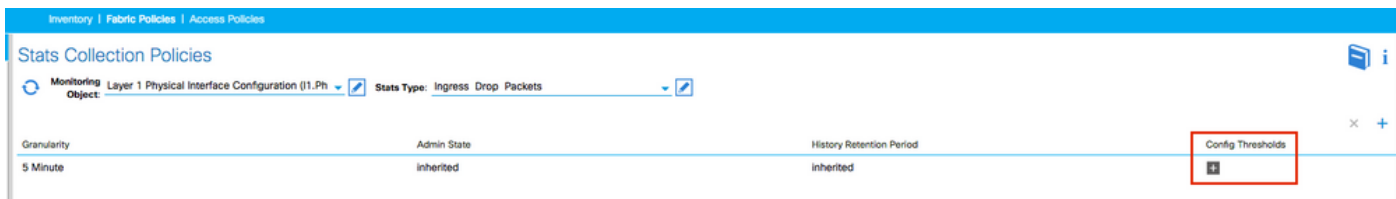
(使用默认策略时，可以跳过步骤4。)

5. 从右侧窗格中的Monitoring Object下拉菜单中选择Layer 1 Physical Interface Configuration (I1.PhysIf)和Stats Type，然后选择Ingress Drop Packets

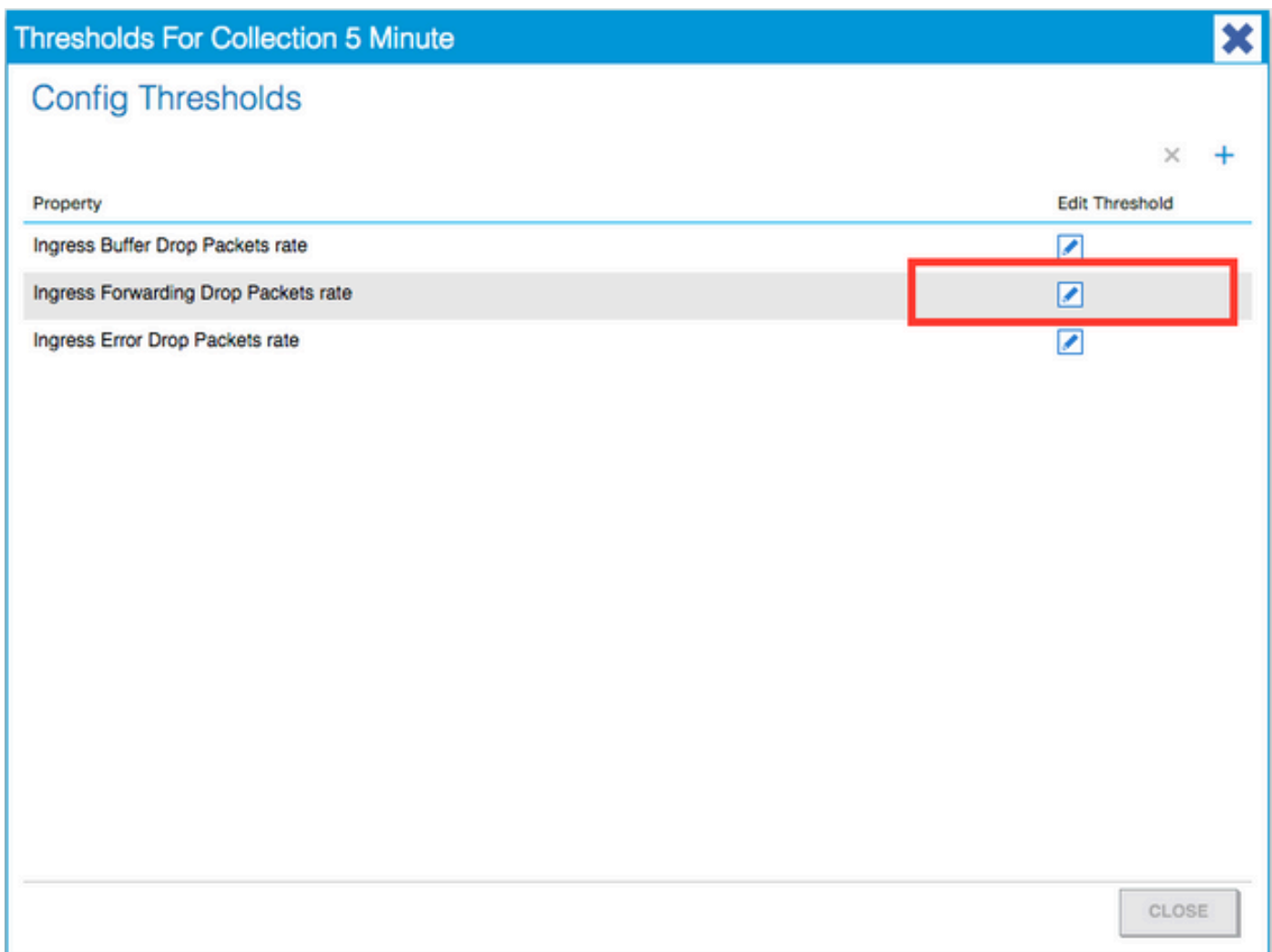
The screenshot shows the Cisco Fabric Policy configuration interface. The left sidebar displays a tree view of policies, with 'Stats Collection Policies' selected. The main content area shows the configuration for 'Stats Collection Policies' for the 'Monitoring Object: Layer 1 Physical Interface Configuration (I1.PhysIf)'. The 'Stats Type' is set to 'Ingress Drop Packets'. Below this, a table shows the configuration details:

Granularity	Admin State
5 Minute	inherited

6. 单击“配置阈值”旁边的+。



7. 编辑转发丢弃的阈值。



8. 建议禁用递增阈值，以配置转发丢包率的“严重”、“主要”、“次要”和“警告”。

Edit Stats Threshold
✕

Ingress Forwarding Drop Packets rate

Normal Value: 0 ↕

Threshold Direction: Both Rising Falling

Rising Thresholds to Config:

- Critical
- Major
- Minor
- Warning

CHECK ALL
UNCHECK ALL

Falling Thresholds to Config:

- Critical
- Major
- Minor
- Warning

CHECK ALL
UNCHECK ALL

Rising

	Set	↕	Reset	↕
Critical	10000	↕	9000	↕
Major	5000	↕	4900	↕
Minor	500	↕	490	↕
Warning	10	↕	9	↕

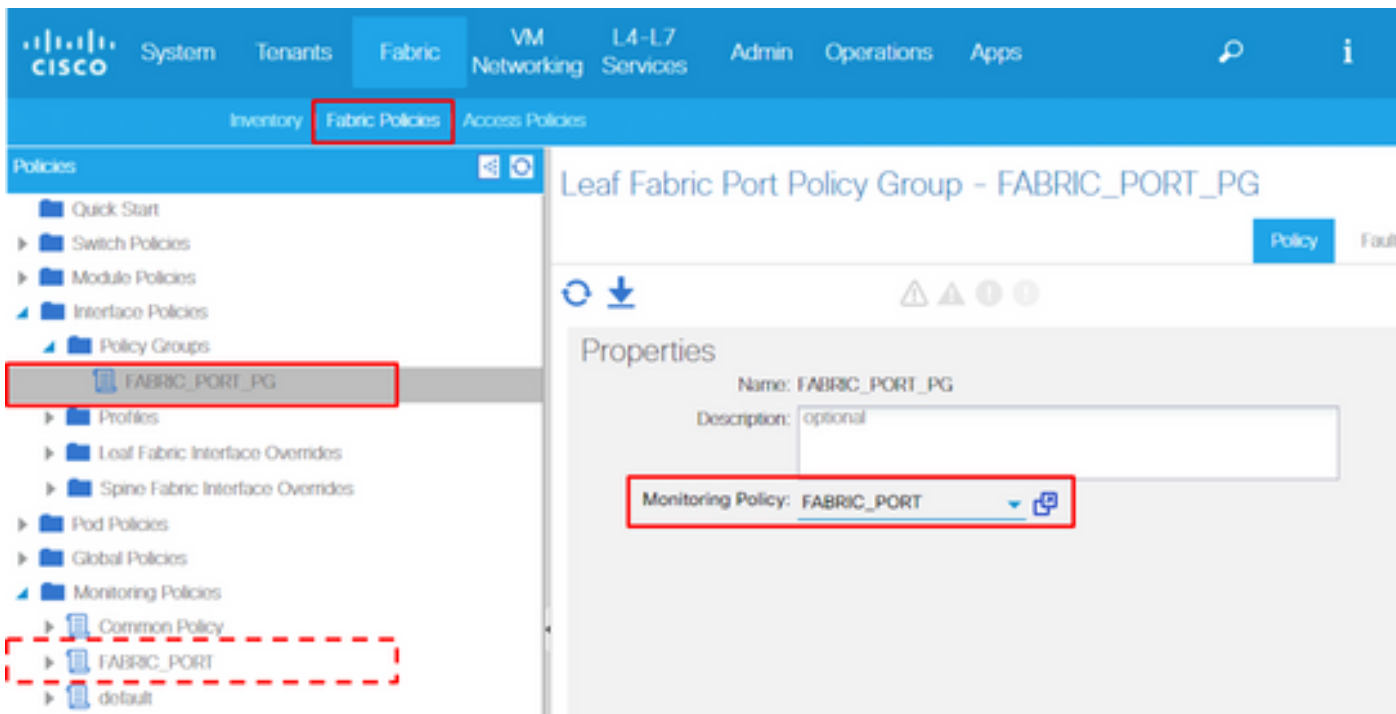
Falling

	Reset	↕	Set	↕
Warning	0	↕	0	↕
Minor	0	↕	0	↕
Major	0	↕	0	↕
Critical	0	↕	0	↕

SUBMIT
CANCEL

9. 将此新监控策略应用于所需端口的接口策略组。请勿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。

(使用默认策略时，可以跳过步骤9。)



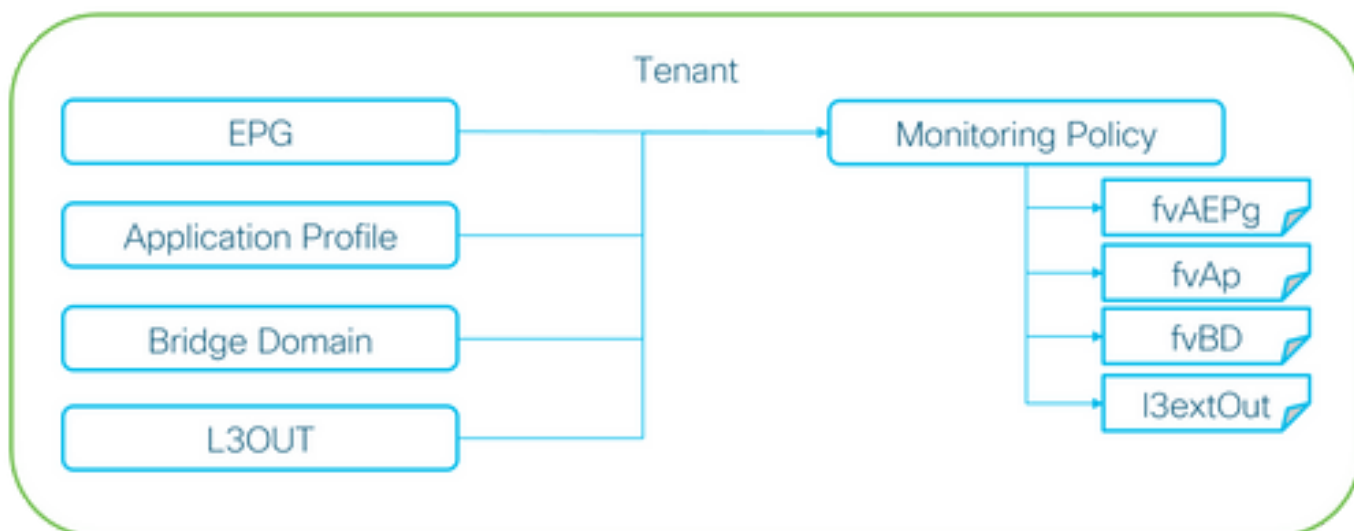
10. 如果这是针对前面板端口（访问策略），请对聚合接口(pc.AggrIf)执行与第1层物理接口配置(I1.PhysIf)相同的操作，以便将此新的监控策略应用于端口通道以及物理端口。

（使用默认策略时，可以跳过步骤10。）

I2IngrPktsAg中的入口丢弃数据包速率

有多个部分。

VLAN or any aggregation of VLAN stats



※ It doesn't have to be one Monitoring Policy. It could be one Monitoring Policy for each.

如上图所示，I2IngrPktsAg在许多对象下受到监控。上图只显示一些示例，但并非所有I2IngrPktsAg的对象。但是，统计信息的阈值是通过监控策略以及I1PhysIf或pcAggrIf下的eqptIngrDropPkts配置的。

每个对象(EPG(fvAEPg)、网桥域(fvBD)等)都可以分配其自己的监控策略，如上图所示。

默认情况下，除非另有配置，否则租户下的所有这些对象都使用租户>通用>监控策略>默认下的默认监控策略。

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。

本示例将更改桥接域上I2IngrPktsAg15min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。

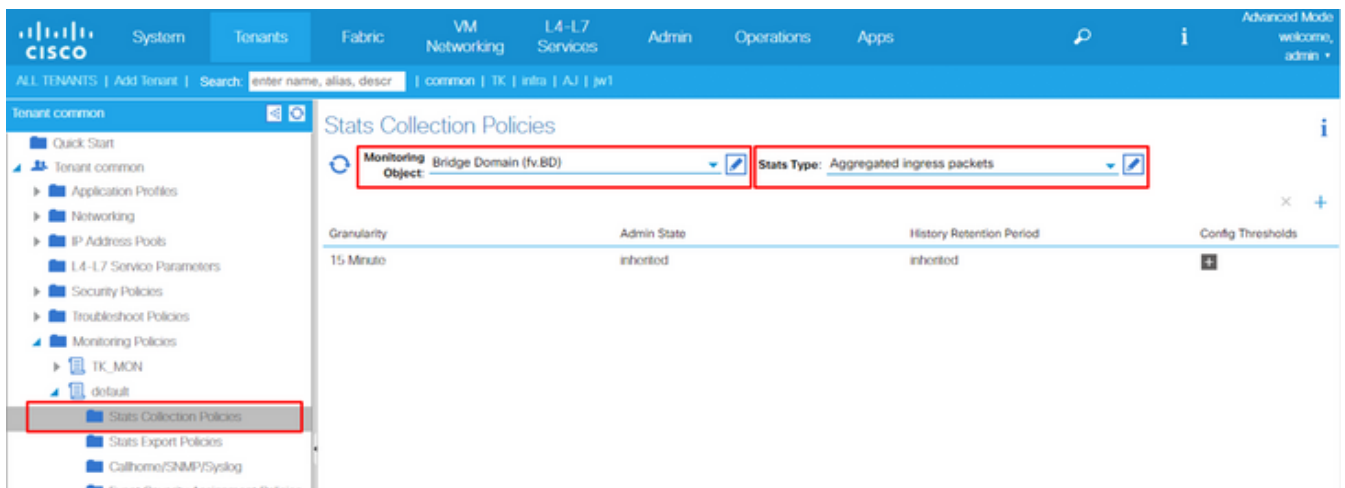
(如果阈值更改可应用于所有组件，请导航到default，而不是创建新组件。)

3. 展开新的监控策略或默认值，然后导航到统计信息收集策略。

4. 在右侧窗格中单击监控对象的铅笔图标，选择网桥域(fv.BD)。

(使用默认策略时，可以跳过步骤4。)

5. 从右侧窗格中的Monitoring Object下拉菜单中选择Bridge Domain (fv.BD)和Stats Type，然后选择Aggregated ingress packets。



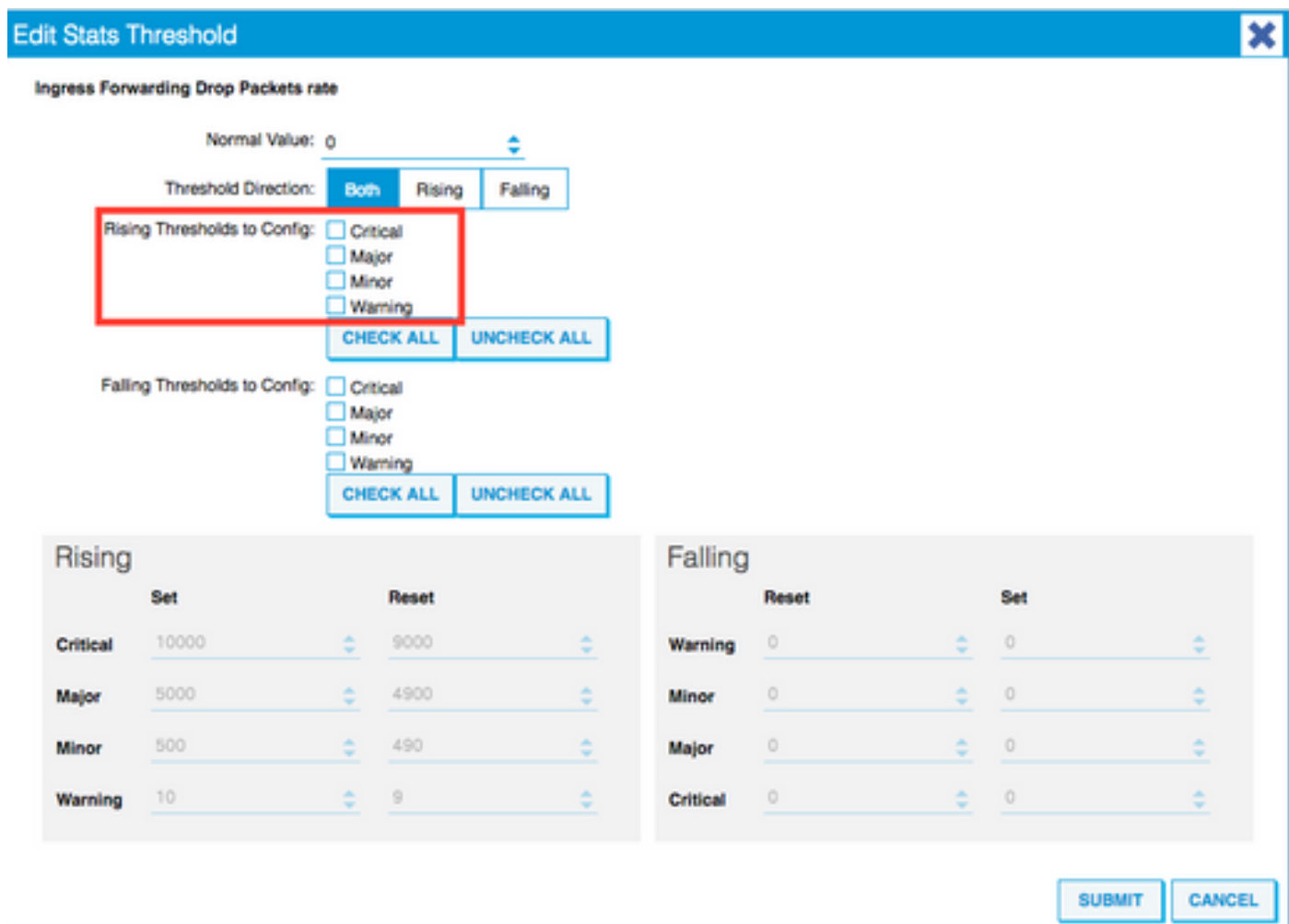
6.单击“配置阈值”旁边的+。



7. 编辑转发丢弃的阈值。




8. 建议禁用递增阈值，以配置转发丢包率的“严重”、“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，可以跳过步骤9。)

The screenshot displays the Cisco SD-WAN management interface for a Bridge Domain (BD1). The left sidebar shows a navigation tree with 'Bridge Domains' expanded to 'BD1'. The main content area shows the configuration for 'Bridge Domain - BD1' under the 'Policy' tab. A red box highlights the 'Monitoring Policy' dropdown menu, which is currently set to 'TK_MON'. Below this, the 'Properties' section lists: 'Unknown Unicast Traffic Class ID: 32770', 'Segment: 15826915', and 'Multicast Address: 225.1.26.128'. A green status indicator shows '100'.

 **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）。