

# 了解Nexus 9000 TAHUSD缓冲区系统日志和拥塞

## 目录

[简介](#)

[先决条件](#)

[要求](#)

[使用的组件](#)

[了解Cisco Nexus 9000云扩展ASIC缓冲架构](#)

[了解超订用和输出丢弃](#)

[了解BUFFER THRESHOLD EXCEEDED系统日志](#)

[了解输出丢弃接口计数器](#)

[超订用方案示例](#)

[后续步骤](#)

[其他信息](#)

[BUFFER THRESHOLD EXCEEDED系统日志配置选项](#)

[针对网络拥塞情况收集的日志](#)

[监控微爆发](#)

## 简介

本文档介绍在配备运行NX-OS软件的Cisco Cloud Scale ASIC（特定应用集成电路）的Cisco Nexus 9000系列交换机上进行排队和缓冲的机理。本文档还介绍了此平台上的端口超订用的症状，例如非零输出丢弃接口计数器和指示已超出缓冲区阈值的系统日志。

## 先决条件

### 要求

Cisco建议您了解共享介质网络上的以太网交换的基础知识，以及这些网络中排队/缓冲的必要性。思科还建议您了解Cisco Nexus交换机上的服务质量(QoS)和缓冲的基础知识。有关详细信息，请参阅以下文档：

- [Cisco Nexus 9000系列NX-OS服务质量配置指南，版本10.1\(x\)](#)
- [Cisco Nexus 9000系列NX-OS服务质量配置指南，版本9.3\(x\)](#)
- [Cisco Nexus 9000系列NX-OS服务质量配置指南，版本9.2\(x\)](#)
- [Cisco Nexus 9000系列NX-OS服务质量配置指南，版本7.x](#)

### 使用的组件

本文档中的信息基于运行NX-OS软件版本9.3(8)的Cisco Nexus 9000系列交换机以及云扩展ASIC。

本文档中介绍的过程仅适用于此处显示的硬件。

- **Nexus 9200/9300固定交换机** N9K-C92160YC-XN9K-C92300YCN9K-C92304QCN9K-

C92348GC-XN9K-C9236CN9K-C9272QN9K-C9332CN9K-C9364CN9K-C93108TC-EXN9K-C93108TC-EX-24N9K-C93180LC-EXN9K-C93180YC-EXN9K-C93180YC-EX-24N9K-C93108TC-FXN9K-C93108TC-FX-24N9K-C93180YC-FXN9K-C93180YC-FX-24N9K-C9348GC-FXPN9K-C93240YC-FX2N9K-C93216TC-FX2N9K-C9336C-FX2N9K-C9336C-FX2-EN9K-C93360YC-FX2N9K-C93180YC-FX3N9K-C93108TC-FX3PN9K-C93180YC-FX3SN9K-C9316D-GXN9K-C93600CD-GXN9K-C9364C-GXN9K-C9364D-GX2AN9K-C9332D-GX2B

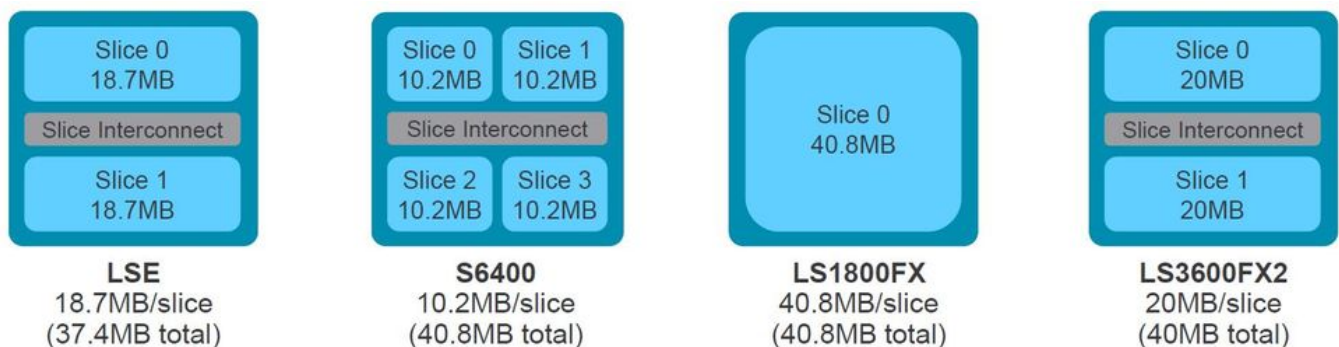
- **Nexus 9500模块化交换机线卡** N9K-X97160YC-EXN9K-X9732C-EXN9K-X9736C-EXN9K-X97284YC-FXN9K-X9732C-FXN9K-X9788TC-FXN9K-X9716D-GX

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。

## 了解Cisco Nexus 9000云扩展ASIC缓冲架构

采用Cisco Cloud Scale ASIC的Cisco Nexus 9000系列交换机实施了“共享内存”出口缓冲区架构。ASIC分为一个或多个“片”。每个片都有自己的缓冲区，只有该片中的端口可以使用该缓冲区。在物理上，每个片被划分为“单元格”，表示缓冲区的部分。切片被划分为“池组”。分配给每个池组的一定数量的信元，它们不在单独的池组之间共享。每个池组有一个或多个“池”，代表单播或组播流量的服务类别(CoS)。这有助于每个池组保证池组所服务的流量类型的缓冲区资源。

此处的图像直观地展示了各种思科云规模ASIC模型如何划分为多个切片。该图像还演示了如何通过单元为每个片分配一定量的缓冲区。



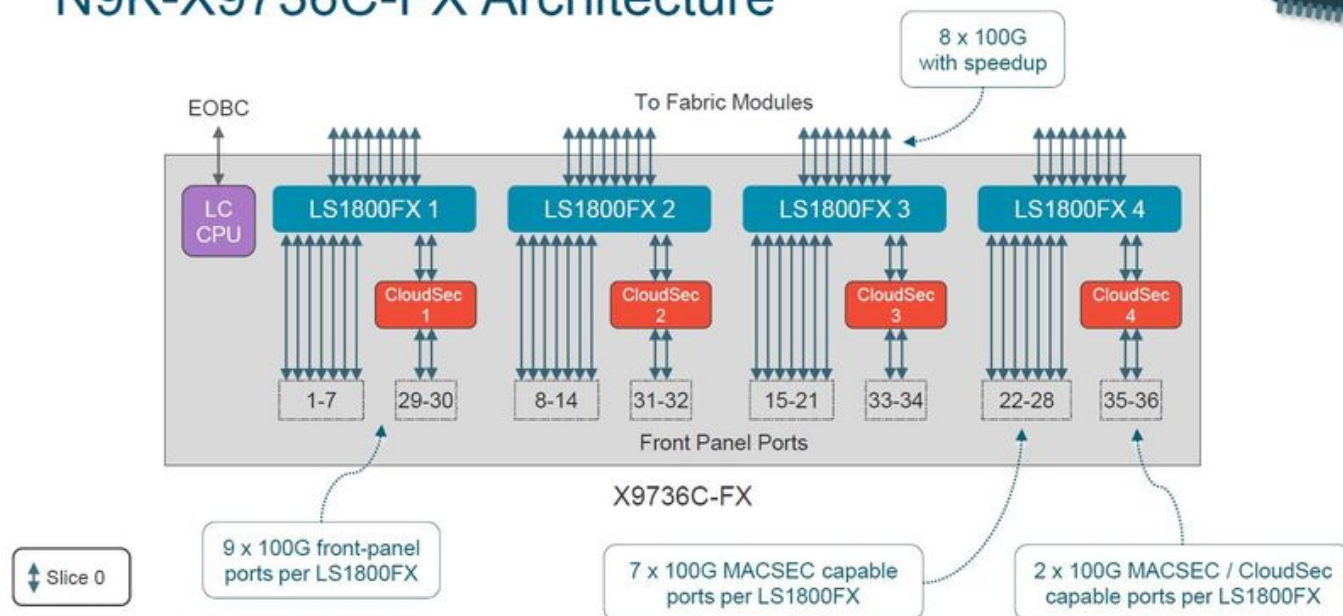
Nexus 9000系列交换机和Nexus 9500线卡的每个型号的内部具有不同数量的Cisco Cloud Scale ASIC，以及指示哪些前面板端口连接到哪个ASIC的不同布局。此处显示了使用N9K-X9736C-FX线卡和N9K-C9336C-FX2交换机的两个示例。

N9K-C9736C-FX线卡具有4个思科云扩展LS1800FX ASIC，每个ASIC一个片。在内部，每个ASIC都称为“单元”。每个扇区称为“实例”，并被分配一个从零开始的整数，用于唯一标识机箱内的该扇区。这将导致如下所示的置换：

- 单元0，片0称为实例0
- 单元1，片0称为实例1
- 单元2，片0称为实例2
- 单元3，片0称为实例3



## N9K-X9736C-FX Architecture

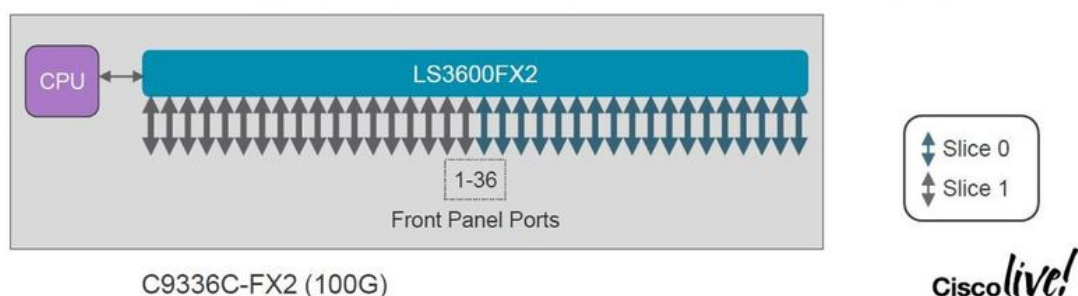


N9K-C9336C-FX2交换机具有一个Cisco Cloud Scale LS3600FX2 ASIC，每个ASIC有两个片。在内部，每个ASIC都称为“单元”。每个扇区称为“实例”，并被分配一个从零开始的整数，用于唯一标识机箱内的该扇区。这将导致如下所示的置换：

- 单元0，片0称为实例0
- 单元0，片1称为实例1



## Nexus 9300-FX2 Switch Architecture



每个线卡和交换机都具有不同的布局，从而产生了不同的实例编号。了解您使用的线卡或交换机的布局对于围绕带宽密集型流量设计网络非常重要。**show interface hardware-mappings**命令可用于将每个前面板端口关联到设备(ASIC)和扇区号。这里显示了一个示例，其中Nexus 9504交换机的接口Ethernet2/16与机箱插槽2中插入的N9K-X9736C-FX线卡映射到单元1，片0。

```
switch# show interface hardware-mappings
```

Legends:

```
SMod - Source Mod. 0 is N/A
Unit - Unit on which port resides. N/A for port channels
HPort - Hardware Port Number or Hardware Trunk Id:
HName - Hardware port name. None means N/A
FPort - Fabric facing port number. 255 means N/A
NPort - Front panel port number
```

VPort - Virtual Port Number. -1 means N/A  
 Slice - Slice Number. N/A for BCM systems  
 SPort - Port Number wrt Slice. N/A for BCM systems  
 SrcId - Source Id Number. N/A for BCM systems  
 MacIdx - Mac index. N/A for BCM systems  
 MacSubPort - Mac sub port. N/A for BCM systems

Name	Ifindex	Smold	Unit	HPort	FPort	NPort	VPort	Slice	SPort	SrcId	MacId	MacSP	VIF	Block	BlkSrcID
Eth2/1 32	1a080000	5	0	16	255	0	-1	0	16	32	4	0	145	0	
Eth2/2 24	1a080200	5	0	12	255	4	-1	0	12	24	3	0	149	0	
Eth2/3 16	1a080400	5	0	8	255	8	-1	0	8	16	2	0	153	0	
Eth2/4	1a080600	5	0	4	255	12	-1	0	4	8	1	0	157	0	8
Eth2/5	1a080800	5	0	0	255	16	-1	0	0	0	0	0	161	0	0
Eth2/6 40	1a080a00	5	0	56	255	20	-1	0	56	112	14	0	165	1	
Eth2/7 32	1a080c00	5	0	52	255	24	-1	0	52	104	13	0	169	1	
Eth2/8 32	1a080e00	6	1	16	255	28	-1	0	16	32	4	0	173	0	
Eth2/9 24	1a081000	6	1	12	255	32	-1	0	12	24	3	0	177	0	
Eth2/10 16	1a081200	6	1	8	255	36	-1	0	8	16	2	0	181	0	
Eth2/11	1a081400	6	1	4	255	40	-1	0	4	8	1	0	185	0	8
Eth2/12	1a081600	6	1	0	255	44	-1	0	0	0	0	0	189	0	0
Eth2/13 40	1a081800	6	1	56	255	48	-1	0	56	112	14	0	193	1	
Eth2/14 32	1a081a00	6	1	52	255	52	-1	0	52	104	13	0	197	1	
Eth2/15 32	1a081c00	7	2	16	255	56	-1	0	16	32	4	0	201	0	
<b>Eth2/16</b> 24	1a081e00	7	<b>2</b>	12	255	60	-1	<b>0</b>	12	24	3	0	205	0	
Eth2/17 16	1a082000	7	2	8	255	64	-1	0	8	16	2	0	209	0	
Eth2/18	1a082200	7	2	4	255	68	-1	0	4	8	1	0	213	0	8
Eth2/19	1a082400	7	2	0	255	72	-1	0	0	0	0	0	217	0	0
Eth2/20 40	1a082600	7	2	56	255	76	-1	0	56	112	14	0	221	1	
Eth2/21 32	1a082800	7	2	52	255	80	-1	0	52	104	13	0	225	1	
Eth2/22 32	1a082a00	8	3	16	255	84	-1	0	16	32	4	0	229	0	
Eth2/23 24	1a082c00	8	3	12	255	88	-1	0	12	24	3	0	233	0	
Eth2/24 16	1a082e00	8	3	8	255	92	-1	0	8	16	2	0	237	0	
Eth2/25	1a083000	8	3	4	255	96	-1	0	4	8	1	0	241	0	8
Eth2/26	1a083200	8	3	0	255	100	-1	0	0	0	0	0	245	0	0
Eth2/27 40	1a083400	8	3	56	255	104	-1	0	56	112	14	0	249	1	
Eth2/28 32	1a083600	8	3	52	255	108	-1	0	52	104	13	0	253	1	
Eth2/29 24	1a083800	5	0	48	255	112	-1	0	48	96	12	0	257	1	
Eth2/30	1a083a00	5	0	44	255	116	-1	0	44	88	11	0	261	1	

```

16
Eth2/31    1a083c00 6    1    48    255    120    -1    0    48    96    12    0    265    1
24
Eth2/32    1a083e00 6    1    44    255    124    -1    0    44    88    11    0    269    1
16
Eth2/33    1a084000 7    2    48    255    128    -1    0    48    96    12    0    273    1
24
Eth2/34    1a084200 7    2    44    255    132    -1    0    44    88    11    0    277    1
16
Eth2/35    1a084400 8    3    48    255    136    -1    0    48    96    12    0    281    1
24
Eth2/36    1a084600 8    3    44    255    140    -1    0    44    88    11    0    285    1
16

```

## 了解超订用和输出丢弃

连接到以太网络的接口一次只能传输一个数据包。当两个数据包需要同时出口以太网接口时，以太网接口开始传输一个数据包，同时缓冲另一个数据包。第一个数据包传输完成后，以太网接口开始从缓冲区传输第二个数据包。当需要出口接口的总流量超过接口的带宽时，接口将被视为超订用。例如，如果总共有15Gbps流量即时进入交换机并需要出口10Gbps接口，则10Gbps接口会超订用，因为它无法一次传输15Gbps流量。

具有云扩展ASIC的Cisco Nexus 9000系列交换机通过在与出口接口关联的ASIC分片的缓冲区中缓冲流量来处理此资源争用。如果需要出口接口的总流量超过接口带宽较长时间，则ASIC分片的缓冲区开始填充需要出口接口的数据包。

当ASIC分片的缓冲区达到90%的利用率时，交换机将生成类似于如下所示的系统日志：

```
%TAHUSD-SLOT2-4-BUFFER_THRESHOLD_EXCEEDED: Module 2 Instance 0 Pool-group buffer 90 percent threshold is exceeded!
```

当ASIC分片的缓冲区完全满时，交换机将丢弃需要传出接口的所有其他流量，直到缓冲区中的空间变为空闲。当交换机丢弃此流量时，交换机将增加出口接口上的Output Discards计数器。

生成的系统日志和非零输出丢弃计数器都是超订用接口的症状。每个症状都将在下面的子部分中详细介绍。

## 了解BUFFER\_THRESHOLD\_EXCEEDED系统日志

下面显示了BUFFER\_THRESHOLD\_EXCEEDED系统日志的示例。

```
%TAHUSD-SLOTX-4-BUFFER_THRESHOLD_EXCEEDED: Module X Instance Y Pool-group buffer Z percent threshold is exceeded!
```

此系统日志包含三个关键信息：

1. **模块X** — 包含超订用接口的线路卡插槽。
2. **实例Y** — 分配给包含超订用接口的ASIC和切片元组的实例编号。
3. **Pool-group buffer Z** — 生成系统日志之前受影响的池组的缓冲区阈值。此百分比是由“已使用的信元”除以“总信元”得出的百分比，如show hardware internal buffer info pkt-stats输出中所示，当连接到模块X时。

## 了解输出丢弃接口计数器

Output Discards接口计数器指示本应离开接口但由于ASIC片缓冲区已满且无法接受新数据包而无法丢弃的数据包数。输出丢弃计数器显示在show interface和show interface counters errors的输出中，如下所示。

```
switch# show interface Ethernet1/1
Ethernet1/1 is up
admin state is up, Dedicated Interface
  Hardware: 1000/10000/25000/40000/50000/100000 Ethernet, address: 7cad.4f6d.f6d8 (bia
7cad.4f6d.f6d8)
  MTU 1500 bytes, BW 40000000 Kbit , DLY 10 usec
  reliability 255/255, txload 232/255, rxload 1/255
  Encapsulation ARPA, medium is broadcast
  Port mode is trunk
  full-duplex, 40 Gb/s, media type is 40G
  Beacon is turned off
  Auto-Negotiation is turned on FEC mode is Auto
  Input flow-control is off, output flow-control is off
  Auto-mdix is turned off
  Rate mode is dedicated
  Switchport monitor is off
  EtherType is 0x8100
  EEE (efficient-ethernet) : n/a
    admin fec state is auto, oper fec state is off
  Last link flapped 03:16:50
  Last clearing of "show interface" counters never
  3 interface resets
  Load-Interval #1: 30 seconds
    30 seconds input rate 0 bits/sec, 0 packets/sec
    30 seconds output rate 36503585488 bits/sec, 3033870 packets/sec
    input rate 0 bps, 0 pps; output rate 36.50 Gbps, 3.03 Mpps
  Load-Interval #2: 5 minute (300 seconds)
    300 seconds input rate 32 bits/sec, 0 packets/sec
    300 seconds output rate 39094683384 bits/sec, 3249159 packets/sec
    input rate 32 bps, 0 pps; output rate 39.09 Gbps, 3.25 Mpps
RX
  0 unicast packets  208 multicast packets  9 broadcast packets
  217 input packets  50912 bytes
  0 jumbo packets  0 storm suppression bytes
  0 runts  0 giants  0 CRC  0 no buffer
  0 input error  0 short frame  0 overrun  0 underrun  0 ignored
  0 watchdog  0 bad etype drop  0 bad proto drop  0 if down drop
  0 input with dribble  0 input discard
  0 Rx pause
TX
  38298127762 unicast packets  6118 multicast packets  0 broadcast packets
  38298133880 output packets  57600384931480 bytes
  0 jumbo packets
  0 output error  0 collision  0 deferred  0 late collision
  0 lost carrier  0 no carrier  0 babble  57443534227 output discard      <<< Output discards
due to oversubscription
  0 Tx pause
```

```
switch# show interface Ethernet1/1 counters errors
```

```
-----
Port          Align-Err    FCS-Err    Xmit-Err    Rcv-Err    UnderSize  OutDiscards
-----
Eth1/1              0          0          0          0          0          57443534227
```

```

-----
Port          Single-Col  Multi-Col  Late-Col  Exces-Col  Carri-Sen  Runts
-----
Eth1/1        0           0          0         0          0          0

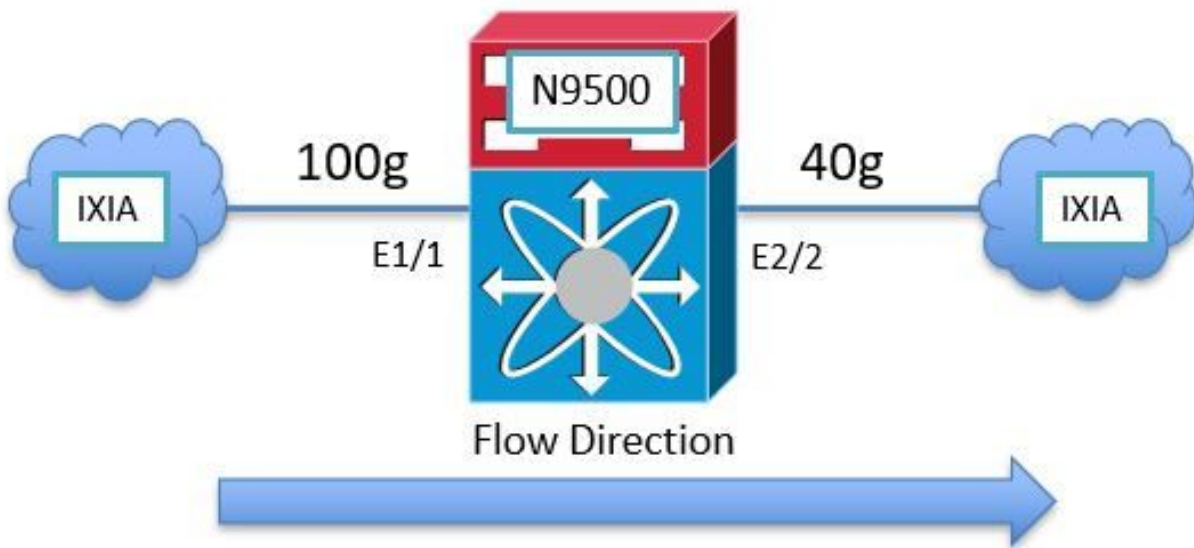
-----
Port          Giants SQETest-Err Deferred-Tx IntMacTx-Er IntMacRx-Er Symbol-Err
-----
Eth1/1        0           --         0         0          0          0

-----
Port          InDiscards
-----
Eth1/1        0

```

## 超订用方案示例

考虑以下场景：两个IXIA流量生成器之间的流量通过机箱插槽1和2中插入了两个N9K-X9736C-FX线卡的Nexus 9504交换机。100Gbps流量通过100Gbps接口Ethernet1/1进入交换机，需要出口40Gbps接口Ethernet2/2。因此，Ethernet2/2超订用。此方案的拓扑如下所示。



由于Nexus 9000云扩展ASIC使用共享内存出口缓冲区架构，您必须检查出口接口Ethernet2/2的缓冲区以查看拥塞。在本示例中，插入插槽2中的线卡是出口线卡，因此您必须使用**attach module 2**命令才能使用**show hardware internal tah buffer counters**命令查看内部硬件缓冲区。请注意，Unit 0、Slice 0 pool-group和关联池的非零“占用丢弃”计数器指示由于池组的缓冲区已满而丢弃的数据包数。

```

switch# attach module 2
module-2# show hardware internal tah buffer counters

```

```

Unit: 0 Slice: 0

```

```

=====
|-----

```

Output Pool-Group drops						
SPAN-PG	Drop-PG	No-drop	CPU--PG	LCPU-PG	RCPU-PG	
<b>Occupancy drops</b>						
0	<b>51152554987</b>	0	0	0	0	
AQM drops	0	0	N/A	N/A	N/A	
N/A						

Output UC Pool counters							
Pool 6	Pool 7	Pool 0	Pool 1	Pool 2	Pool 3	Pool 4	Pool 5
Dynamic Threshold (cells)							
93554	93554	93554	93554	93554	93554	93554	93554
<b>Occupancy drops</b>							
0	0	<b>51152555398</b>	0	0	0	0	0
AQM drops		0	0	0	0	0	0
0	0						

Output MC Pool counters							
Pool 6	Pool 7	Pool 0	Pool 1	Pool 2	Pool 3	Pool 4	Pool 5
Dynamic Threshold (cells)							
93554	93554	93554	93554	93554	93554	93554	93554
Dynamic Threshold (desc)							
93554	93554	93554	93554	93554	93554	93554	93554
Dynamic Threshold (inq thr)							
64035	64035	64035	64035	64035	64035	64035	64035
Occupancy drops							
0	0	0	0	0	0	0	0

Additional counters	
MEM cell drop reason	: 0
MEM descriptor drop reason	: 0
OPG cell drop reason	: 0
OPG descriptor drop reason	: 0
OPG CPU cell drop reason	: 0
OPG CPU descriptor drop reason	: 0
OPG SPAN cell drop reason	: 0
OPG SPAN descriptor drop reason	: 0
OPOOL cell drop reason	: 0
OPOOL descriptor drop reason	: 0
UC OQUEUE cell drop reason	: 51152556479
MC OQUEUE cell drop reason	: 27573307
OQUEUE descriptor drop reason	: 0
MC OPOOL cell drop reason	: 0
FWD DROP	: 15



```

SOD : 0
BMM BP : 0
No Drop : 0
Packets received : 87480806439
TRUNC MTU : 0
TRUNK BMM BP : 0
VOQFC messages sent : 0
SOD messages sent : 0
SPAN descriptor drop : 0

```

Unit: 1 Slice: 0

=====

```

|-----|
|-----|
|
|                               Output Pool-Group drops
|
|           Drop-PG           No-drop           CPU--PG           LCPU-PG           RCPU-PG
SPAN-PG           |-----|
|-----|
Occupancy drops           0           0           0           0           0
0           |
AQM drops           0           0           N/A           N/A           N/A
N/A           |

```

```

|-----|
|-----|
|
|                               Output UC Pool counters
|
|           Pool 0           Pool 1           Pool 2           Pool 3           Pool 4           Pool 5
Pool 6           Pool 7           |-----|
|-----|
Dynamic Threshold (cells)           93554           93554           93554           93554           93554           93554
93554           93554           |
Occupancy drops           0           0           0           0           0           0
0           0           |
AQM drops           0           0           0           0           0           0
0           0           |

```

```

|-----|
|-----|
|
|                               Output MC Pool counters
|
|           Pool 0           Pool 1           Pool 2           Pool 3           Pool 4           Pool 5
Pool 6           Pool 7           |-----|
|-----|
Dynamic Threshold (cells)           93554           93554           93554           93554           93554           93554
93554           93554           |
Dynamic Threshold (desc)           93554           93554           93554           93554           93554           93554
93554           93554           |
Dynamic Threshold (inq thr)           64035           64035           64035           64035           64035           64035
64035           64035           |
Occupancy drops           0           0           0           0           0           0
0           0           |

```

```

|-----+-----+-----+-----+-----+-----+-----+-----+
|
|                               Additional counters
|-----+-----+-----+-----+-----+-----+-----+-----+
MEM cell drop reason : 0
MEM descriptor drop reason : 0

```

```

OPG cell drop reason           :    0
OPG descriptor drop reason     :    0
OPG CPU cell drop reason       :    0
OPG CPU descriptor drop reason :    0
OPG SPAN cell drop reason      :    0
OPG SPAN descriptor drop reason:    0
OPOOL cell drop reason         :    0
OPOOL descriptor drop reason   :    0
UC OQUEUE cell drop reason     :    0
MC OQUEUE cell drop reason     :    0
OQUEUE descriptor drop reason  :    0
MC OPOOL cell drop reason      :    0
FWD DROP                       :    8
SOD                            :    0
BMM BP                         :    0
No Drop                        :    0
Packets received               : 45981341
TRUNC MTU                      :    0
TRUNK BMM BP                   :    0
VOQFC messages sent           :    0
SOD messages sent             :    0
SPAN descptor drop            :    0

```

回想一下，每个ASIC单元/片元组都通过唯一标识的“实例”表示。 **show hardware internal buffer info pkt-stats**命令的输出显示有关每个实例的拥塞池组（缩写为“PG”）的详细信息。该命令还显示缓冲区中已使用的信元的历史峰值/最大数量。最后，该命令显示流量被缓冲的端口的云扩展ASIC端口标识符的即时快照。此命令示例如下所示。

```

switch# attach module 2
module-2# show hardware internal buffer info pkt-stats

```

#### Instance 0

```

=====
|-----|
|-----|
|
|                                     Output Pool-Group Buffer Utilization (cells/desc)
|
|                                     Drop-PG      No-drop      CPU--PG      LCPU-PG      RCPU-PG
|-----|
| SPAN-PG      |
|-----|
|-----|
| Total Instant Usage (cells)      59992      0      0      0      0
| 0      |
| Remaining Instant Usage (cells) 33562      0      1500      250      1500
| 1500      |
| Peak/Max Cells Used              90415      0      N/A      N/A      N/A
| N/A      |
| Switch Cells Count              93554      0      1500      250      1500
| 1500      |
|
| Total Instant Usage (desc)          0      0      0      0      0
| 0      |
| Remaining Instant Usage (desc)      93554      0      1500      250      1500
| 1500      |
| Switch Desc Count                   93554      0      1500      250      1500
| 1500      |
|-----|
|-----|
|
|                                     Output UC Pool Buffer Utilization (cells/desc)

```

Pool 6	Pool 7	Pool 0	Pool 1	Pool 2	Pool 3	Pool 4	Pool 5
Total Instant Usage (cells)		60027	0	0	0	0	0
Total Instant Usage (desc)		0	0	0	0	0	0
Peak/Max Cells Used		62047	0	0	0	0	0

		Output MC Pool Buffer Utilization (cells/desc)					
Pool 6	Pool 7	Pool 0	Pool 1	Pool 2	Pool 3	Pool 4	Pool 5
Total Instant Usage (cells)		0	0	0	0	0	0
Total Instant Usage (desc)		0	0	0	0	0	0
Total Instant Usage (inq cells)		0	0	0	0	0	0
Total Instant Usage (packets)		0	0	0	0	0	0
Peak/Max Cells Used		60399	0	0	0	0	0

Instant Buffer utilization per queue per port  
Each line displays the number of cells/desc utilized for a given port for each QoS queue  
One cell represents approximately 416 bytes

ASIC Port	Q7	Q6	Q5	Q4	Q3	Q2	Q1	Q0
-----------	----	----	----	----	----	----	----	----

[12]  
<<< ASIC Port 12 in Unit 0 Instance 0 is likely the congested egress interface

UC->	0	0	0	0	0	0	0	59988
MC cells->	0	0	0	0	0	0	0	0
MC desc->	0	0	0	0	0	0	0	0

另请参阅命令的峰值。使用此命令将系统日志与特定池组、池或端口中的潜在峰值相关联

```
switch# show hardware internal buffer info pkt-stats peak
```

```
slot 1
=====
```

```
Instance 0
```

```
=====
```

Pool-Group Peak counters	
--------------------------	--

```
Drop PG      :      0
No-drop PG   :      0
```

```
-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Pool Peak counters                                     |
|-----+-----+-----+-----+-----+-----+-----+-----+
| MC Pool 0      :      0
| MC Pool 1      :      0
| MC Pool 2      :      0
| MC Pool 3      :      0
| MC Pool 4      :      0
| MC Pool 5      :      0
| MC Pool 6      :      0
| MC Pool 7      :      0
|
| UC Pool 0      :      0
| UC Pool 1      :      0
| UC Pool 2      :      0
| UC Pool 3      :      0
| UC Pool 4      :      0
| UC Pool 5      :      0
| UC Pool 6      :      0
| UC Pool 7      :      0
```

```
-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Port Peak counters                                     |
| classes mapped to count_0: 0 1 2 3 4 5 6 7
| classes mapped to count_1: None
|-----+-----+-----+-----+-----+-----+-----+-----+

```

```
[0]                                     <<< ASIC Port. This can be checked via "show
hardware interface-mappings"
count_0      :      0
count_1      :      0

[1]
count_0      :      0
count_1      :      0
```

**show interface hardware-mappings**命令可用于将云扩展ASIC端口标识符转换为前面板端口。在上述示例中，与片/实例0上的ASIC单元0关联的ASIC端口12（由show interface hardware-mappings输出中的HPort列表示）有59,988个占用的信元，每个信元为416个字节。此处显示了show interface hardware-mappings命令的示例，该命令将此接口映射到前面板端口Ethernet2/2。

```
switch# show interface hardware-mappings
Legends:
SMod  - Source Mod. 0 is N/A
Unit  - Unit on which port resides. N/A for port channels
HPort - Hardware Port Number or Hardware Trunk Id:
HName - Hardware port name. None means N/A
FPort - Fabric facing port number. 255 means N/A
NPort - Front panel port number
VPort - Virtual Port Number. -1 means N/A
Slice - Slice Number. N/A for BCM systems
SPort - Port Number wrt Slice. N/A for BCM systems
SrcId - Source Id Number. N/A for BCM systems
MacIdx - Mac index. N/A for BCM systems
MacSubPort - Mac sub port. N/A for BCM systems
```

```
-----
-----
Name      Ifindex  Smod  Unit  HPort  FPort  NPort  VPort  Slice  SPort  SrcId  MacId  MacSP  VIF  Block
```

BlkSrcID

```
-----  
-----  
Eth2/2    1a080200 5    0    12    255  4    -1    0    12    24    3    0    149  0  
24
```

我们可以使用show queuing interface命令进一步将接口Ethernet2/2的超订用与QoS队列丢弃相关联。这里显示了一个示例。

```
switch# show queuing interface Ethernet2/2
```

```
Egress Queuing for Ethernet2/2 [System]
```

```
-----  
QoS-Group# Bandwidth% PrioLevel          Shape          QLimit  
              Min           Max           Units  
-----  
7             -             1             -             -             9(D)  
6             0             -             -             -             9(D)  
5             0             -             -             -             9(D)  
4             0             -             -             -             9(D)  
3             0             -             -             -             9(D)  
2             0             -             -             -             9(D)  
1             0             -             -             -             9(D)  
0            100            -             -             -             9(D)  
-----  
+-----+  
|                               QOS GROUP 0                               |  
+-----+  
|                               | Unicast | Multicast |                               |  
+-----+  
|                               | Tx Pkts | 35593332351 | 18407162 |                               |  
|                               | Tx Byts | 53532371857088 | 27684371648 |                               |  
| WRED/AFD & Tail Drop Pkts | 53390604466 | 27573307 |                               |  
| WRED/AFD & Tail Drop Byts | 80299469116864 | 110293228 |                               |  
|                               | Q Depth Byts | 24961664 | 0 |                               |  
| WD & Tail Drop Pkts | 53390604466 | 27573307 |                               |  
+-----+  
|                               QOS GROUP 1                               |  
+-----+  
|                               | Unicast | Multicast |                               |  
+-----+  
|                               | Tx Pkts | 0 | 0 |                               |  
|                               | Tx Byts | 0 | 0 |                               |  
| WRED/AFD & Tail Drop Pkts | 0 | 0 |                               |  
| WRED/AFD & Tail Drop Byts | 0 | 0 |                               |  
|                               | Q Depth Byts | 0 | 0 |                               |  
|                               | WD & Tail Drop Pkts | 0 | 0 |                               |  
+-----+  
|                               QOS GROUP 2                               |  
+-----+  
|                               | Unicast | Multicast |                               |  
+-----+  
|                               | Tx Pkts | 0 | 0 |                               |  
|                               | Tx Byts | 0 | 0 |                               |  
| WRED/AFD & Tail Drop Pkts | 0 | 0 |                               |  
| WRED/AFD & Tail Drop Byts | 0 | 0 |                               |  
|                               | Q Depth Byts | 0 | 0 |                               |  
|                               | WD & Tail Drop Pkts | 0 | 0 |                               |  
+-----+  
|                               QOS GROUP 3                               |  
+-----+
```

	Unicast	Multicast
Tx Pkts	0	0
Tx Byts	0	0
WRED/AFD & Tail Drop Pkts	0	0
WRED/AFD & Tail Drop Byts	0	0
Q Depth Byts	0	0
WD & Tail Drop Pkts	0	0
QOS GROUP 4		
	Unicast	Multicast
Tx Pkts	0	0
Tx Byts	0	0
WRED/AFD & Tail Drop Pkts	0	0
WRED/AFD & Tail Drop Byts	0	0
Q Depth Byts	0	0
WD & Tail Drop Pkts	0	0
QOS GROUP 5		
	Unicast	Multicast
Tx Pkts	0	0
Tx Byts	0	0
WRED/AFD & Tail Drop Pkts	0	0
WRED/AFD & Tail Drop Byts	0	0
Q Depth Byts	0	0
WD & Tail Drop Pkts	0	0
QOS GROUP 6		
	Unicast	Multicast
Tx Pkts	0	0
Tx Byts	0	0
WRED/AFD & Tail Drop Pkts	0	0
WRED/AFD & Tail Drop Byts	0	0
Q Depth Byts	0	0
WD & Tail Drop Pkts	0	0
QOS GROUP 7		
	Unicast	Multicast
Tx Pkts	0	0
Tx Byts	0	0
WRED/AFD & Tail Drop Pkts	0	0
WRED/AFD & Tail Drop Byts	0	0
Q Depth Byts	0	0
WD & Tail Drop Pkts	0	0
CONTROL QOS GROUP		
	Unicast	Multicast
Tx Pkts	5704	0
Tx Byts	725030	0
Tail Drop Pkts	0	0
Tail Drop Byts	0	0
SPAN QOS GROUP		

	Unicast	Multicast
Tx Pkts	0	0
Tx Byts	0	0

Per Slice Egress SPAN Statistics

SPAN Copies Tail Drop Pkts	0
SPAN Input Queue Drop Pkts	0
SPAN Copies/Transit Tail Drop Pkts	0
SPAN Input Desc Drop Pkts	0

最后，您可以使用**show interface**命令检验出口接口Ethernet2/2是否具有非零输出丢弃计数器。这里显示了一个示例。

```
switch# show interface Ethernet2/2
Ethernet2/2 is up
admin state is up, Dedicated Interface
  Hardware: 1000/10000/25000/40000/50000/100000 Ethernet, address: 7cad.4f6d.f6d8 (bia
7cad.4f6d.f6d8)
  MTU 1500 bytes, BW 40000000 Kbit , DLY 10 usec
  reliability 255/255, txload 232/255, rxload 1/255
  Encapsulation ARPA, medium is broadcast
  Port mode is trunk
  full-duplex, 40 Gb/s, media type is 40G
  Beacon is turned off
  Auto-Negotiation is turned on FEC mode is Auto
  Input flow-control is off, output flow-control is off
  Auto-mdix is turned off
  Rate mode is dedicated
  Switchport monitor is off
  EtherType is 0x8100
  EEE (efficient-ethernet) : n/a
    admin fec state is auto, oper fec state is off
  Last link flapped 03:16:50
  Last clearing of "show interface" counters never
  3 interface resets
  Load-Interval #1: 30 seconds
    30 seconds input rate 0 bits/sec, 0 packets/sec
    30 seconds output rate 36503585488 bits/sec, 3033870 packets/sec
    input rate 0 bps, 0 pps; output rate 36.50 Gbps, 3.03 Mpps
  Load-Interval #2: 5 minute (300 seconds)
    300 seconds input rate 32 bits/sec, 0 packets/sec
    300 seconds output rate 39094683384 bits/sec, 3249159 packets/sec
    input rate 32 bps, 0 pps; output rate 39.09 Gbps, 3.25 Mpps
RX
  0 unicast packets  208 multicast packets  9 broadcast packets
  217 input packets  50912 bytes
  0 jumbo packets  0 storm suppression bytes
  0 runs  0 giants  0 CRC  0 no buffer
  0 input error  0 short frame  0 overrun  0 underrun  0 ignored
  0 watchdog  0 bad etype drop  0 bad proto drop  0 if down drop
  0 input with dribble  0 input discard
  0 Rx pause
TX
  38298127762 unicast packets  6118 multicast packets  0 broadcast packets
  38298133880 output packets  57600384931480 bytes
  0 jumbo packets
  0 output error  0 collision  0 deferred  0 late collision
```

```
0 lost carrier 0 no carrier 0 babble 57443534227 output discard <<< Output discards
due to oversubscription
0 Tx pause
```

## 后续步骤

如果您观察具有云扩展ASIC的Nexus 9000系列交换机上的输出丢弃，则可以使用以下一种或多种方法解决问题：

- 如果出现输出丢弃的接口是单个接口，并且未捆绑到端口通道中，则升级接口的带宽有助于缓解拥塞。例如，如果拥塞的出口接口是10Gbps接口，则升级到25Gbps、40Gbps接口或100Gbps接口有助于解决此问题。根据出口接口的收发器外形，这可以通过升级收发器（例如从插入QSFP端口内的CVR-QSFP-SFP10G中的10Gbps SFP+迁移至本地40Gbps QSFP收发器来实现）。这也可以通过将拥塞的出口接口配置从10Gbps端口迁移到25Gbps、40Gbps或100Gbps端口来实现。
- 如果出现输出丢弃的接口是单个接口，并且未捆绑到端口通道中，那么将拥塞接口与具有相同带宽的另一个接口一起配置为端口通道成员可以缓解拥塞。
- 如果出现输出丢弃的接口是端口通道接口，则向端口通道添加更多成员可以提高整个端口通道的带宽，并改善多个大型流量的负载均衡散列。
- 验证网络中主机之间的拥塞流量是否涉及速度降低的接口（例如，通过40Gbps接口进入交换机并通过10Gbps接口流出交换机的流量）。这可能成为导致网络拥塞的瓶颈。通过将低速接口（例如10Gbps）升级到高速接口（例如25Gbps、40Gbps等）来消除此瓶颈可以缓解网络拥塞。
- 如果无法增加拥塞出口接口上的可用带宽，请验证[端到端QoS](#)并为网络应用适当的排队操作。
- 如果微突发是间歇性拥塞的潜在原因，请参阅本文档的[监控微突发](#)部分，了解有关配置微突发监控的信息。

## 其他信息

本文档的这一部分包含有关在遇到BUFFER\_THRESHOLD\_EXCEEDED系统日志、网络拥塞/超订应用场景以及输出丢弃接口计数器递增时采取的后续步骤的其他信息。

### BUFFER\_THRESHOLD\_EXCEEDED系统日志配置选项

您可以修改系统缓冲区状态轮询间隔，该间隔控制系统轮询ASIC片缓冲区的当前利用率的频率。这通过**hardware profile buffer info poll-interval**全局配置命令完成。默认配置值为5,000毫秒。此配置可以全局修改或按模块修改。此配置命令的示例显示在此处，它被修改为值1,000毫秒。

```
switch# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
switch(config)# hardware profile buffer info poll-interval timer 1000
switch(config)# end
switch# show running-config | include hardware.profile.buffer
hardware profile buffer info poll-interval timer 1000
switch#
```

您可以修改端口出口缓冲区使用阈值，该值控制系统何时生成BUFFER\_THRESHOLD\_EXCEEDED系统日志，指示ASIC片缓冲区利用率已超过配置的阈值。这通过**hardware profile buffer info port-threshold**全局配置命令完成。默认配置值为90%。此配置可以



全局修改或按模块修改。此配置命令示例如下所示，其中将其修改为值80%。

```
switch# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
switch(config)# hardware profile buffer info port-threshold threshold 80
switch(config)# end
switch# show running-config | include hardware.profile.buffer
hardware profile buffer info port-threshold threshold 80
switch#
```

您可以修改交换机生成的BUFFER\_THRESHOLD\_EXCEEDED系统日志之间的最小间隔。您还可以彻底禁用BUFFER\_THRESHOLD\_EXCEEDED系统日志。这通过**hardware profile buffer info syslog-interval timer**全局配置命令完成。默认配置值为120秒。通过将值设置为0秒，可以完全禁用系统日志。此配置命令示例显示在此处，其中系统日志被完全禁用。

```
switch# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
switch(config)# hardware profile buffer info syslog-interval timer 0
switch(config)# end
switch# show running-config | include hardware.profile.buffer
hardware profile buffer info syslog-interval timer 0
switch#
```

## 针对网络拥塞情况收集的日志

除了本文档中列出的命令外，您还可以从受网络拥塞情况影响的交换机收集此处显示的日志，以确定拥塞的出口接口。

1. **show tech-support details**命令的输出。
2. **show tech-support usd-all**命令的输出。
3. **show tech-support ipqos all**命令的输出。
4. 使用插入了Cisco Cloud Scale线卡的Nexus 9500系列交换机时，**show system internal interface counters peak module {x}**命令的输出，其中{x}是托管拥塞出口接口的模块的插槽编号。

## 监控微爆发

如果在非常短的时间间隔（微突发）内发生拥塞或超订用，则需要其他信息来准确描述超订用对交换机的影响。

配备Cisco Cloud Scale ASIC的Cisco Nexus 9000系列交换机可监控微爆发流量，这些微爆发可导致您的环境中出现临时网络拥塞和流量丢失。有关微爆发以及如何配置此功能的详细信息，请参阅此处显示的文档：

- [Cisco Nexus 9000系列NX-OS服务质量配置指南10.1\(x\)版的“微突发监控”一章](#)
- [Cisco Nexus 9000系列NX-OS服务质量配置指南9.3\(x\)版的“微突发监控”一章](#)
- [Cisco Nexus 9000系列NX-OS服务质量配置指南9.2\(x\)版的“微突发监控”一章](#)
- [Cisco Nexus 9000系列NX-OS服务质量配置指南7.x版的“微突发监控”一章](#)

## 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）。